Research Frontiers in Cloud Forensics

A comprehensive survey on issues and frameworks

[1]Dr. B. RajalakshmiRavishankar,[2]Heyshanthini Pandiyakumari S,[3]Aishwarya Rajamani
[1]Head of Department,[2]Assistant Professor,[3]Student of Master of Technology
[1]Computer Science and Engineering
[1]New Horizon College of Engineering, Bangalore,India

Abstract : From the latter part of the last decade, the distributed paradigm of computing, that offers services and products on a demand basis known popularly as cloud computing, has been on the rise. Cost-effective and easy to work with, the cloud has been adopted by businesses small and big. With the proliferation of such a huge and networked paradigm of computing the need for security has also become imperative. Especially the field of forensics in relation to the cloud has to be developed in order to recreate an incident that has led to a crime or is a criminal activity in itself. The demands of cloudforensics are exponentially greater than the demands of digital forensics. This paper focuses on issues in cloud forensics, some important frameworks and a few inferences.

***Index Terms*- Cloud Computing, Cloud Forensics, Frameworks**

I. INTRODUCTION

Cloud Computing, the most cited buzzword since the start of this decade, is a paradigm of computing that offers services, computing infrastructure and platform in an on demand basis. The conveniences that the cloud brought with it are: the reduction of capital cost for purchasing expensive software; the absolute doing-away of memory shortage issues; the ability to access content from any place and at any time with a mere high-speed internet connection. The companies that shaped the rapid acceptance of this concept by the general public are Google, Amazon and Microsoft. Their cloud products such as Google Docs, Amazon Drive, Microsoft Office 365, has influenced the general class of people to embrace ubiquitous data access as a highly productive way to conduct everyday business. However, as with most scenarios in technological advancements where usability and efficiency override the factor of safety, here also the vulnerabilities in this paradigm been relegated.

This study focuses on vulnerabilities of cloud and probes into whether and how the cloud is able to provide substantial information for digital forensic investigations in the case of such eventualities.

II. CLOUD ARCHITECTURE AND ASSOCIATED VULNERABILITIES

The infrastructure of cloud setup can provide valuable insights into the origin and type of insecurity factor involved. A typical cloud contains: physical hardware assets such asnetworks of servers, data centres; co-ordination systems;software support for end-user access; auditing mechanismsand other resources suited to the application requirements.

The vulnerabilities thus begin with the interfaces or APIsoffered by cloud service providers. In case of weakauthorization checks or poor regression testing of the rathernaïve interfaces all the cloud services such SaaS, PaaS andIaaS are affected. Account hijacking is common in such casesand private user data can come under attack. Virtual Machinehopping where one Virtual Machine gains access to anotherVM by exploiting some infrastructure managementvulnerabilities may also occur. Some other VM related threatsare VM escaping where the hypervisor is exploited andbypassed. Data leakage occurs in the networking betweencloud components or between end user terminals and thecloud. Denial of Service is another threat that destabilizescloud system resources and services provided by it. Thuslegitimate users are affected severely.

III. INTRODUCTION TO CLOUD FORENSICS

Digital forensics is defined as the practice of assembling, scrutinizing, and reporting information found in computers and networks, in a legal milieu – as evidence in a criminal or civil investigation, or as proof in a commercial or private setting. The evidence of digital forensic analysis supports in incident response and consequent activities provided a successful digital-attack or data breach has been identified.

A caustic challenge that the cloud faces is the proliferation of malware and products created by criminals disguised as software as a service. These are often found to be low in cost and lethal in consequence [2]. With multiple tenants logging in from anywhere, anytime and possibly several times isolating a criminal incident is especially hard.

Cloud forensics is formally defined [2] as a cross-discipline between cloud computing and digital forensics. The model for cloud forensics is essentially multi-dimensional consisting of technical, legal and organizational aspects. The first aspect of technicality deals with facilitative tools and processes to carry out cloud forensic investigations. The second aspect of legality deals with the lawfulness of the conducted investigation. And the organizational aspect defines the hierarchy of person, designations and roles generally.

IV. CLOUD FORENSIC ISSUES

The process of cloud forensics has some challengingrequirements. Firstly identifying the origination of anunauthorized modification to the user's cloud environment isneeded. Legal rules that dictate the access rights forinvestigative purposes may be restrictive. Understanding thecloud's working mechanism and API interfaces is also veryvital. With millions of logs it is hard to discriminate betweenthe relevant and the ones out of scope.

Some of the major issues faced by cloud forensics are discussed in the sub-sections below.

4.1 Validity of Data Collected

Investigation of an attack in cloud setups cannot follow the same procedures that general digital forensic techniques employ. A study by Vincent Urias[5] et al, conducted experiments on VMWare's vSphere products as it is a popular virtualization solution utilized by several organizations across the globe. The authors identify the challenge that the hypervisors pose in establishing the validity of any data collected post an incident as this software layer has the ability to add or remove or modify data. This is risky because in electronic forensics data is primary evidence.

4.2 Complexity from Heterogeneity

Cloud computing in its essence is heterogeneous in terms of hardware and software. Owing to this it is especially hard to write code for forensic extraction procedures. Data and applications are highly customized and fast evolving. User satisfaction comes first in the cloud paradigm and consequently it is difficult to produce standard cloud forensic tools.

4.3 Multitenant Infrastructure

The users of cloud services share the infrastructure. And the order of accounts may range from thousands to millions. Thus retrieving the appropriate forensic evidence from particular accounts can be challenging.

4.4 The Legal and Jurisdictional Aspects

Service level agreements do not always contain clauses for forensic investigations in the cloud [6]. Each country has different clauses for handling real world criminals. But there are no legal procedures to prosecute cyber criminals especially if the attackers or misusers are located in a different continent/country.

4.5 Dearth of Expert Knowledge in Cloud Forensics

Cloud computing has to be perceived as a growing field in technology rather than one that has stabilized. Forensic activities are difficult to conduct in the cloud setup and the necessary expertise is yet to materialize in huge numbers.

4.6 Varied Format of Logs

Logs are records of events and transactions. Different cloud providers make use of proprietary or third-party logging services. Thus the representations of logs are varied making evidence processing an uphill task.

4.7 Data Preservation over time

Today the cloud is dealing with big data. The volume, variety and velocity that define this kind of data calls for humongous storage and complex distributed processing. Thus developing forensic techniques for big data in the cloud requires novel approaches.

V. ANTI-FORENSICS

The issue of anti-forensics is vital to the discussion at this stage. Harris [3] defined anti-forensics as attempts to wreck theavailability of evidence or to render it useless. The specific anti-forensic activities are: deletion, obscuration of evidencedata and trails, disabling of logs and so on. In certain cases retrieving the tampered data may require far too much time that the very worth of that piece becomes irrelevant to the forensic process. Harris, in his concluding remark rightly argues that while forensic technology has been the focus for academia, the training of personnel to tackle the rise of anti-forensics as well as the need to address the primary flaws that accommodate anti-forensic activities has to be brought to the forefront.

VI. FRAMEWORKS AND TECHNIQUES

6.1 Systematic Log Preservation

In a research paper by Shams Zawoad, et al [4] a concrete scheme for cloud forensic evidence from logs has been expounded. The main focus is on preserving user logs persistently and securely as unmodified log data is the life force in recreating incidents. There are two parts to the evidence establishment procedure here- integrity verification of the log is first done and then sequence verification is also performed. Log entries are first created and encrypted. Each of these encrypted entries is enchained into a log chain structure. Following this, all the log entries are stored in a persistent database which constitutes what they term as proof of past log. Thus this proof forms the test against which integrity and sequence checks are performed during investigations by the means of some APIs.

6.2 Observing Virtual Machines

Virtual Machine Monitor or Hypervisors offer limited knowledge to users about what proceeds in the layer beneath. In a research conducted by Alluri [7] et al, a trifold solution framework has been proposed with main emphasis on observing virtual machines. Swap space in general computing terminology refers to the memory to which the inactive pages of the system are shifted in order to free up space for other processes' pages. This framework suggests the monitoring of swap spaces of the virtual machine over time to reveal the kind of activities going on inside it. Any untoward activities can be detected and stored for future forensic activities. In addition to swap space observation the authors suggest the monitoring of terminated processes that were deemed to be suspicious previously.

6.3 Fuzzy Logic for Evidence Generation

In a research work by P. Santra et al [8], a well thought out framework based on data mining and fuzzy logic is expounded. A data repository containing knowledge required for inferring information about logs is created. This is associated with a classifier that categorizes attack and finally generates evidence. Sniffers are used to draw logs from cloud which transferred to the above discussed classifier. Thus the source and the classified evidence can be retrieved by the investigator and presented for legal procedures.

6.4 Utilizing Blockchain Concept

A commendable technique has been proposed in a research work by Songyang Wu [9] that emphasises on a tamper proof concept with strong privacy aspects. The methodology followed is: firstly, submitted forensic data is fortified into records for which Blockchain receipts are attached. This prevents further tampering. Also, the records are kept as hashed values rather than as explicit information. Thus privacy is established.

6.5 Thorough Analysis of Origin

Research conducted by J. Boucher [10] recommends a fourfold approach to differentiate between locally retrieved data and synched data. Such a distinction is helpful in forensic activities as every little detail can be a key to recreating the incident or identifying the arbitrators. Application, OS and timeline analysis is done following which an opinion is reached as to whether the piece of evidence originated from a local device or was synchronized from another.

VII. Recommended plan for solution development

Two opposing forces, cloud forensics and anti-forensicshave been laid out perpendicularly. The following are a few suggestions based on the current scenario with respect to cloud investigations:

a. Cloud investigations require standardization of technology and procedures. Thus a worldwide technical consortium for cloud security to develop international cloud laws is suggested

b. A huge body of cloud forensic system comprising of software and technical professionals has to be developed in parallel to the development of cloud

c. Certification bodies form one wing of the forensic equation in certain proposed frameworks. Thus legal accreditation of such bodies in order to establish trust is also suggested

VIII. BEST PRACTICES FOR CLOUD SECURITY

It is important for all stakeholders of cloud system to actively contribute to safe functioning of this mammoth digital infrastructure.

a. Users have to take care to encrypt personal data stored on the cloud in order to prevent incidents arising from exposure of personal data.

b. Cloud Service Providers have to take an active interest in establishing the foundation needed for adequate forensic software and hardware to operate.

IX. CONCLUSION AND FUTURE WORK

Thus cloud security challenges stand just as tall as thepromise of the technology itself. On the part of the user discretion in cloud usage follows as a direct consequence of the discussion presented above. While the cloud's facilities may be safely utilized to a certain extent in less critical fields while it's adoption in more critical fields has to be checked. Parallel to the deliberations on developing standards to evaluate cloud forensic tools and development of methods that eliminate the loopholes which facilitate anti-forensics, should the study of socio-cultural factors influencing the cloud usage be taken up scientifically. This is because the digital world for the most part replicates the real world.

In conclusion it would seem wise to pace cloud technology developments to the rate at which cloud security also evolves.

References

[1]V. R. Kebande, N. M. Karie, and H. S. Venter, "Cloud-Centric Framework for isolating Big data as forensic evidence from IoT infrastructures", 2017 1st International Conference on Next Generation Computing Applications (NextComp), 2017.

[2]R. Harris, "Arriving at an anti-forensics consensus: Examining how to define and control the anti-forensics problem," Digital Investigation, vol. 3, pp. 44–49, 2006.

[3]F. Marturana, G. Me, and S. Tacconi, "A Case Study on Digital Forensics in the Cloud," 2012 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, 2012.

[4]S. Zawoad, A. K. Dutta, and R. Hasan, "SecLaaS," Proceedings of the 8th ACM SIGSAC symposium on Information, computer and communications security - ASIA CCS 13, 2013.

[5]V. Urias, J. W. Young, and S. Hatcher, "Implications of Cloud Computing on Digital Forensics," PsycEXTRA Dataset.

[6]Rafael Accorsi and KeyunRuan, “Challenges of Cloud Forensics: A Survey of the Missing Capabilities”, ERCIM NEWS 90, July 2012.

[7]B K S P Kumar Raju Alluri and G. Geethakumari, “A digital forensic model for introspection of virtual machines in cloud computing,” 2015 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES), 2015.

[8]P. Santra, P. Roy, D. Hazra, and P. Mahata, “Fuzzy Data Mining-Based Framework for Forensic Analysis and Evidence Generation in Cloud Environment,” Advances in Intelligent Systems and Computing Ambient Communications and Computer Systems, pp. 119–129, 2018.

[9]Y. Zhang, S. Wu, B. Jin, and J. Du, “A blockchain-based process provenance for cloud forensics,” 2017 3rd IEEE International Conference on Computer and Communications (ICCC), 2017.

[10]J. Boucher and N.-A. Le-Khac, “Forensic framework to identify local vs synced artefacts,” Digital Investigation, vol. 24, 2018.