Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 6.017

IJCSMC, Vol. 6, Issue. 10, October 2017, pg.67 – 73

Digital Forensic: Issues and Challenges in Cloud Computing Environment

Sabavat Naresh[1] M.Tech, **R.Hari Singh[2]**

[1]Visvesvaraya College of Engg. & Technology, Ranga Reddy, India

[2]Research Scholar, Dept of CSE, Osmania University, Hyderabad, India

[1] nareshchowan47@gmail.com; [2] harisingh501@gmail.com

Abstract— Cloud computing is a relatively new computing paradigm that builds upon virtualisation technologies to provide hardware, platforms and software as services over the Internet. The huge popularity and utility of the cloud environment has made it the soft target of cloud crimes. In this paper, we have discussed about the basics of cloud computing, features, the emerging area of cloud forensics, and highlights its challenges and opportunities.

Keywords— Cloud computing, cloud forensics, Forensic, Digital Forensic, Cybercrime.

I. INTRODUCTION

Cloud computing is a recently developing paradigm of distributed computing. A cloud infrastructure is virtual, distributed in nature and usually spans over multiple jurisdictions. In cloud environments hardware, platforms and software are managed by a third party and they are not in the full control of the owners, as has been the case in traditional settings. Service providers may lease hardware storage services from third parties to store their enterprise and client data. Cloud computing is gaining wider acceptance despite the security concerns that still prevail. When security breaches occur in cloud environments, digital forensic investigations need to be carried out. In the cloud, critical data is more vulnerable and at the same time, difficult to acquire when an incident that requires a digital forensic investigation arises.

Characteristics

The features of cloud computing that have changed the face of physical computing and has urged traditional vendors to move to the cloud technology have been listed in this section.

- Resource pooling
- On demand service
- Scalability
- Virtualization
- Reliability
- Maintainability
- High performance
- Customizable
- Location independent
- Multitenancy
- Efficient resource utilization
- Cost pay-as-use

Types of Cloud

Cloud is of three types they are:

1. Private Cloud: This type of cloud is maintained within an organization and used solely for their internal purpose Security, network bandwidth are not critical issues for private cloud.
2. Public Cloud: In this type an organization rents cloud services from cloud providers on-demand basis.
3. Hybrid Cloud: This type of cloud is composed of multiple internal or external cloud.

II. CLOUD COMPUTING SERVICE MODELS

1. Infrastructure-as-a-Service (IaaS)
2. Platform-as-a-Service (PaaS)
3. Software-as-a-Service (SaaS)

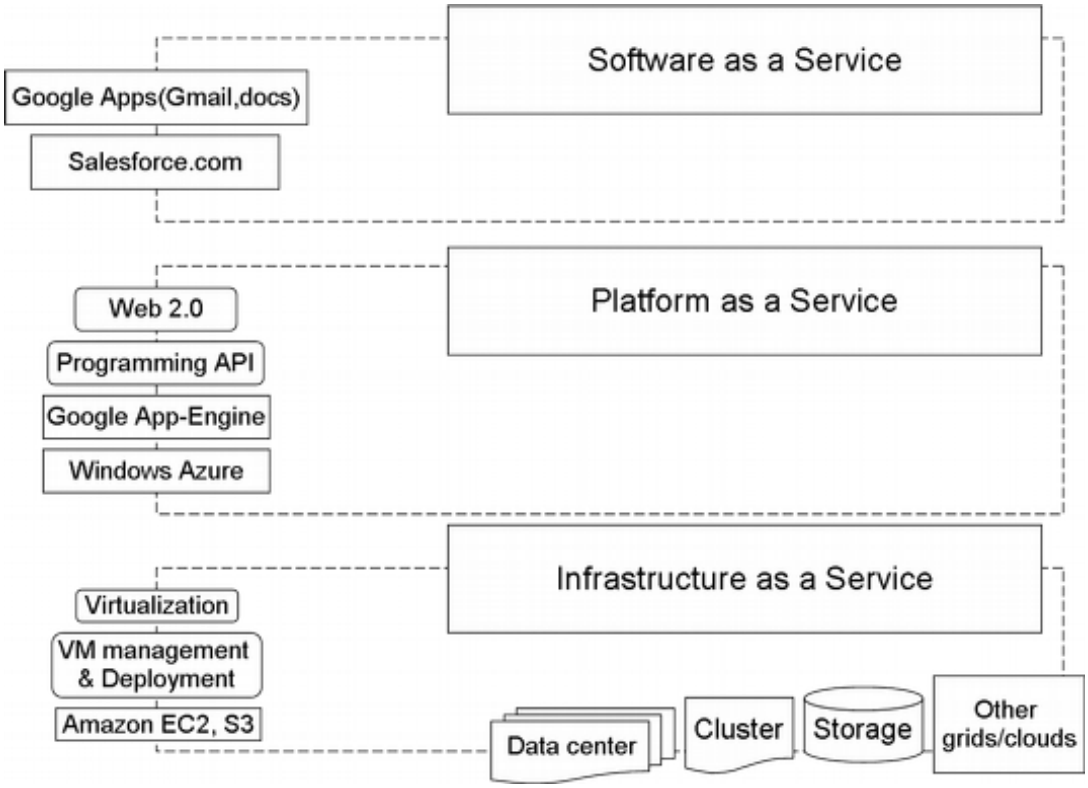

fig1: Cloud Service stack

Infrastructure-as-a-Service (IaaS):

This service model provides the user the facility of renting processing power and storage to run his or her own virtual machine in the cloud. A user can access the launched virtual machine through a thin client interface such as a web browser running in computers, mobiles, PDA's(Personal Digital Assistant), etc. devices. Users will be charged based on the resources the virtual machine consumes from the cloud.

Platform-as-a-Service (PaaS):

Through this model cloud owner provides the user the facility of renting a platform to develop and deploy the user applications in the cloud environment. It is basically an application middleware offered as a service to developers, integrators, and architects. Users will be charged according to the platform (e.g., Database, .Net, etc.) used and bandwidth consumed. The well-known example of PaaS is Google App Engine. There are a number of other PaaS providers like Windows Azure, Force.com, Drupal, IBM, etc., to name few.

Software as a Service (SaaS)

Through this model, a client can make use of software applications made available from the cloud provider. Typically, users interact with SaaS applications using a web browser. An example of SaaS is the Google Apps isuite offered by Google.

III. CLOUD FORENSICS

Cloud forensics is a cross discipline of cloud computing and digital forensics. Cloud computing is a shared collection of configurable networked resources that can be reconfigured quickly with minimal effort. Digital forensics is the application of computer science principles to recover electronic evidence for presentation in a court of law . In order to analyze the domain of cloud forensics more comprehensively, and to emphasize the fact that cloud forensics is a multi-dimensional issue instead of merely a technical issue, we discuss the technical, organizational and legal dimensions of cloud forensics.

Technical Dimension

The technical dimension encompasses the procedures and tools that are needed to perform the forensic process in a cloud computing environment. These include data collection, live forensics, evidence segregation, virtualized environments and proactive measures.

Legal Dimension

Traditional digital forensic professionals identify multi-jurisdictional and multi-tenancy challenges as the top legal concerns. The legal dimension of cloud forensics requires the development of regulations and agreements to ensure that forensic activities do not breach laws and regulations in the jurisdictions where the data resides. Also, the confidentiality of other tenants that share the same infrastructure should be preserved.

Organizational Dimension

A forensic investigation in a cloud computing environment involves at least two entities: the CSP and the cloud customer. However, the scope of the investigation widens when a CSP outsources services to other parties.

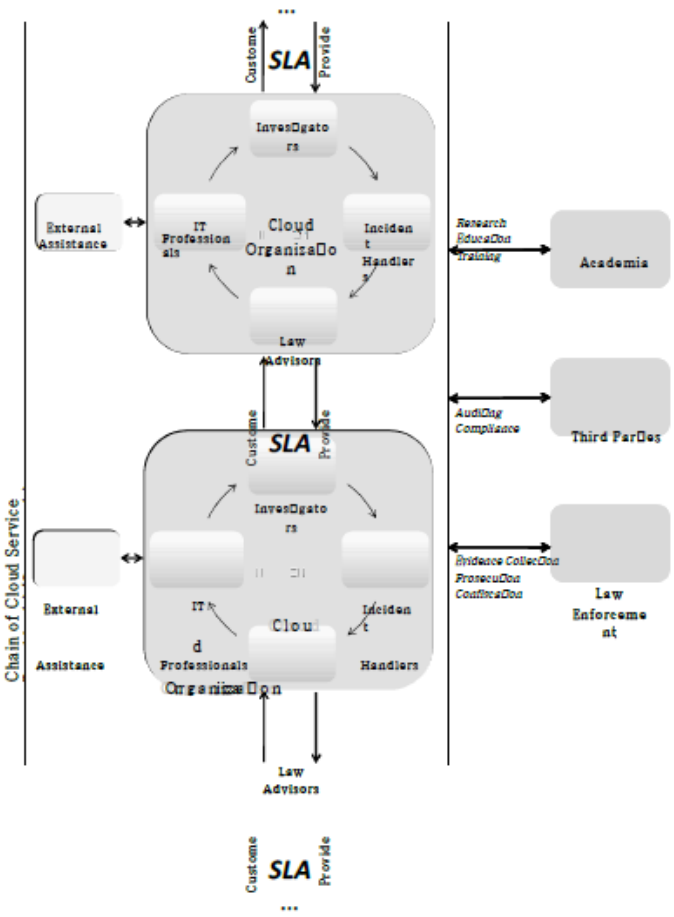

fig1: shows the various entities that may be involved in a cloud forensic investigation.

IV. CLOUD FORENSICS ISSUES AND CHALLENGES

Fourteen categories of issues and challenges were identified in the collection phase of cloud forensics and these fourteen categories were grouped into three groups of the Collection phase of digital forensics: identification of digital evidence; preservation of digital evidence and acquisition of digital evidence.

Digital Investigation Phase	Digital Investigation Process	Issue/Challenges
Collection Phase	Identification	Decentralization of data centres Decentralization of data logs Physical locations unknown or not accessible Specific logging volatile
	Preservation	Inaccessibility to virtual instance Dependency on CSP Metadata/Provenance protection Volatile data Evidence gathering
	Acquisition	Separation of Customer/Multi tenancy Protection of personal Information Layers of Trust Different jurisdictions/Bodies of Law Chain of Custody

Table1: Issues and Challenges

Identification

In the collection phase of digital evidence three categories were identified with issues and challenges. The categories are: decentralization of data centres, decentralization of data logs and physical location unknown or not accessible. The process of digital evidence identification is crucial in order to know where there the digital evidence is and how to access the digital evidence. If the digital evidence cannot be identified the first step, the chain of custody fails and there is no digital evidence admissible in a court of law.

Decentralization of data centre

Cloud computing's distributed architecture allows data to be created, stored, processed and distributed over several data centres and physical machines which are globally dispersed and also possibly dispersed into multiple geographical locations and jurisdictions. Data is replicated to other servers to ensure redundancy of data. The stored data within a data centre is replicated and distributed at physical level and the data can also be fragmented across multiple data centres. The distribution of data depends on the data centres' performance and availability.

Decentralization of data logs

In digital forensics some of the most useful information is stored within log files. In a cloud computing environment, these logs are decentralized, as data stored in the cloud is replicated to multiple server and data centres to confirm redundancy of data. Multiple cloud users' log information may be stored together or can be spread over multiple servers. Cloud architectures consist of several layers and tiers and logs are generated in each tier. All of these layers produce logs that are very valuable for digital forensic investigations. Even if the cloud customer specifies the location where the data should be stored, log files are decentralized as the replication of the data is decentralized and the cloud customer has very limited access to log files.

Physical locations unknown or not accessible

In digital forensics, access to the physical servers and actual data and to secure the evidential data is crucial to an investigation. For both the cloud customer and the investigator in cloud environments, physical access to the servers where the data is stored is in most of the cases technically not possible due to the remote location of the data or the fact that the location of the data cannot be determined. The remote location of the data can be distributed over several different geographical locations over different jurisdictions making it difficult to determine which legal framework and procedure to use in the evidence collection process. The challenge also exists in that there is a lack of physical proximity between the cloud customer and the cloud service provider. The cloud customer or investigator also has no access to the network routers, load balancers and other networking components within the cloud computing environment.In the case where the locations are distributed over locations in different jurisdictions with different legal requirements, the process to access the data can be

too slow due to serious cross border red tape and the adversary may have time to access the data and change or destroy evidential data.

Preservation

During the process of the coding and categorizing five categories emerged for the preservation of evidence. The categories are:

•Dependency on Cloud Service Provider (CSP)
•Inaccessibility to Virtual Instances
•Metadata/Provenance protection
•Specific logging volatile
•Volatile data

Dependency on CSP

In a cloud computing environment the cloud service provider has all the power over the environment and therefore controls the source of the evidential data. The process of preserving digital evidence in the cloud highly depends on the support that the investigator receives from the cloud service provider (CSP).The investigator or law enforcement agency (LEA) currently requires the cooperation of the CSP to collect evidential data. The CSP can also alter evidential data and logs if they themselves are under investigation.

Inaccessibility to virtual instances

The virtualized nature of cloud computing has an impact on the collection of evidential data due to highly limited or no access to the virtual instance. Even in the cloud service scenario of IaaS, the customer's virtual machine (VM) is controlled by the cloud service provider (CSP) and the CSP is responsible for the hypervisors, network infrastructure right down to the physical hardware of the data centre.

The virtualized data may be spread over different physical devices, different geographical locations and jurisdictions. In virtualized environments the collection of evidential data may need to occur through the virtualization software which can have an impact on the soundness of the evidence, chain of custody and the admissibility of the evidence in court.

Metadata/Provenance protection

Metadata, also known as data provenance, is the history of digital objects. Metadata describes ownership and the process history (create, modify and access) of data objects. Metadata is vital to the success of forensic investigations in order to determine the ownership of evidential data (who access the data) and the time -line of evidential data (when data accessed). The uncertainty about metadata and metadata availability are challenges for investigations in the cloud and who and when questions can remain unanswered if the supporting metadata is unavailable.

Specific logging volatile

Logs are very useful evidential data in an investigation. Logs include system logs, network logs, firewall logs and router logs. If the cloud service provider (CSP) does not run any logging application then no opportunity exists to collect specific logging information during an investigation. If the CSP does run logging applications, then the logs must be manageable sizes to be useful and to prevent wiping memory on hosting servers. Currently CSPs are not obligated to provide all logs and logs are not reasonably protected by the CSPs. Another challenge in cloud computing is that that the cloud customer cannot gather network logs or router logs due to the fact that the underlying cloud architecture is under the control of the CSP. Firewall logs may also be volatile as the logs can contain information of other cloud customers accessing the same CSP.

Volatile data

In cloud computing frequently used data may be stored in volatile memory or may be cached in the cloud customer's device during the interaction with the cloud. Volatile data is lost in cloud computing if the virtual machine is powered down or rebooted or if the incorrect preservation process was followed by the CSP in the process of evidence collection on behalf of the investing a-tor Acquisition.

Chain of custody

The chain of custody is the roadmap that shows how evidence was identified, preserved, acquired, examined and analyzed for the evidence to be admissible in court. The chain of custody refers to identification of devices, the physical control of devices, acquisition of data, whether the devices were running or powered down and also how the evidential data was preserved to prevent any further change to the evidential data. The chain of custody also documents all individuals who were in contact with that data.

Different Jurisdictions and bodies of law
The nature of cloud computing with decentralized data centres and virtualization makes it difficult to determine the physical location of the data, let alone the body of law which governs and also restrict the investigation scope.

Evidence gathering
Data is lost when an adversary cancels a cloud contract and a VM is powered down as a result. Real network and router logs cannot be gathered by the cloud customer for forensic purposes. Evidence is untrustworthy due to the cloud service provider (CSP) involvement in the collection of evidential data. Metadata is lost in the process of evidence gathering and the integrity of evidence data gathered cannot be verified.
Another challenge is that evidence gathering has to be timely to prevent an adversary of destroying or modifying data in the time that the law enforcement agency (LEA) has to wait for search warrants due red tape of cross border investigations due to the distributed nature of cloud computing.

Layers of trust
Cloud computing environment consists of several different layers and the different cloud services introduce several levels of trust. In SaaS the cloud customer is totally dependent on the cloud service provider (CSP) to provide evidential data from the applications down to the actual servers where data are processed and stored. In IaaS the cloud customer still needs to rely on CSP for evidential data from the physical host servers of the virtual machines (VM) of the cloud customer although the rest of the application layer is under control of the cloud customer. Thus, irrespective of service of SaaS, PaaS or IaaS there is a trust issue present when evidential data are collected by the CSP in the process of an investigation.

Protection of personal information
Local privacy and/or data protection legislation differ from jurisdiction to jurisdiction. In cloud computing environments, data can be processed and stored across the globe in different jurisdictions. Data can be stored on a cloud server in a country where privacy laws are not enforced or non-existent. In cloud computing, appropriate measures and controls should be applied to personal data in the process of investigation and unauthorized individuals may not access personal data. Detailed logs may contain sensitive and private information and access to this information should be controlled and authorized within the investigation team. The cloud service provider is also facing confidentially issues and is thus might be reluctant to provide raw data.

Separation of customers / multi tenancy
The resource pooling nature of cloud computing has the effect that multiple customers can share the same physical server. In the event of an investigation the cloud service provider (CSP) has to assure that the customer or investigator does not have access to other cloud customer's data.

V. Conclusion

Cloud forensics are still faced with a multitude of problems that have not been addressed yet by research to find solutions for the issues. Data acquisition in the cloud remains the biggest issue with many and varied problems. Much research is needed to develop procedures and tools that can be used by service providers to extract the data needed by investigators in a forensically sound way. Service Level Agreements between cloud customer and CSP's can only protect the cloud customer up to a point if a transparent investigation cannot be done.

References

1. J.E. Smith and R. Nair. An overview of virtual machine architectures. pages 1–20, October 2001. http:// www.ece.wisc.edu/~jes/902/papers/intro.pdf
2. N. Beebe, Digital forensic research: The good, the bad and the unaddressed, in Advances in Digital Forensics V, G. Peterson and S.Shenoi (Eds.), Springer, Heidelberg, Germany, pp. 17–36, 2009.
3. S. Liles, M. Rogers and M. Hoebich, A survey of the legal issues facing digital forensic experts, in Advances in Digital Forensics V,G. Peterson and S. Shenoi (Eds.), Springer, Heidelberg, Germany,pp. 267–276, 2009.
4. Abbadi, I. M., & Lyle, J. (2011). Challenges for provenance in cloud computing. TaPP.
5. Biggs, S., & Vidalis, S. (2009). Cloud computing: The impact on digital forensic investigations. In International Conference for nternet Technology and Secured Transactions (pp. 1–6). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5402561
6. Birk, D. (2011). Technical challenges of forensic investigations in cloud computing environments. In Workshop on Cryptography and Security in Clouds (pp. 1–6).
7. Dykstra, J., & Sherman, A. T. (2011). Understanding issues in cloud forensics: Two hypothetical case studies. Journal of Network Forensics, 3(1), 19–31.

8. Grispos, G., Storer, T., & Glisson, W. (2012). Calm before the storm: The challenges of cloud computing in digital forensics. International Journal of Digital Crime and Forensics, 4(2), 28–4
9. Reilly, D., Wren, C., & Berry, T. (2011). Cloud computing : Pros and cons for computer forensic investigations. International Journal Multimedia and Image Processing, 1(1), 26–34.
10. Taylor, M., Haggerty, J., Gresty, D., & Lamb, D. (2011). Forensic investigation of cloud computing systems. Network Security, 2011(3), 4–10.
11. Zargari, S., & Benford, D. (2012). Cloud forensics: Concepts, issues, and challenges. 2012 Third International Conference on Emerging Intelligent Data and Web Technologies, 236–243.

