

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/275021701>

Cloud Computing Virtualization and Cyber Attacks: Evidence Centralization

Conference Paper · March 2015

CITATIONS

0

READS

1,356

2 authors:



Yucel Turel

Polish-Japanese Academy of Information Technology

4 PUBLICATIONS 1 CITATION

[SEE PROFILE](#)



Romuald K. Kotowski

Polish-Japanese Academy of Information Technology

96 PUBLICATIONS 105 CITATIONS

[SEE PROFILE](#)

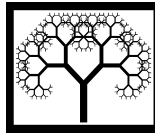
Some of the authors of this publication are also working on these related projects:



E-learning [View project](#)



Thermodynamics of irreversible processes [View project](#)



Cloud Computing Virtualization and Cyber Attacks: Evidence Centralization

Y. Turel and R. Kotowski

Department of Applied Informatics

Polish-Japanese Academy of Information Technology

Warsaw, Poland

Abstract

Cloud computing is the on-line resource management service which can be used on the pay-per-use base. Despite many advantages for the end-users, it has weaknesses, for example it is vulnerable against cyber-attacks. Cloud computing has a complex nature and it works on a number of physical servers running multiple virtual machines. A large number of log files from multiple servers and layers poses risks in cloud computing operations. Collecting digital evidence after the malicious attacks is important to understand and assess the criminal activity. In this paper the importance of centralization in the cloud computing virtual environment against the cyber-attacks is proposed and the virtualization, threats and centralized logging mechanism approach is introduced and discussed.

Keywords: cloud computing, digital forensics, centralization, virtualization, cyber-attacks, evidence data.

1 Introduction

The use of a collection of services, applications, information, and infrastructure comprising of pools of computer, network, information, and storage resources is what the cloud computing describes. Cloud computing provides many advantages for businesses such as elasticity of resources, reduction in cost of ownership (at least in theory). Nowadays, it is very common for enterprises to migrate to cloud from their in-house server farms. The main reason for this is that cloud computing offers virtual platforms based on pay-per-use to provide scalable access to data and applications. Thus, there is no physical infrastructure that users can reach but virtual environment that they can access via internet. Scalability and elasticity of the services

are the major advantages for the end-users. Using cloud computing, the users cut their costs in terms of upgrading hardware and software, in allocating space for the server farms, employing administrators and IT technicians. Cloud computing provides Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS) [1]. One of the main challenges of cloud computing are security and trust. Security is a big issue for the cloud clients as they have to ensure that their confidential data are secure and only accessible by them. Cloud computing has security weaknesses. Single security design cannot solve the cloud computing security problem; traditional and new technologies must be merged for protecting the total cloud computing integrity. There are severe threats against cloud computing. As a result of globalisation, the free movement of individuals and the widespread use of Information and Communication Technology (ICT), there are a wide number of scenarios on the intersection between cybercrime jurisdiction and the ubiquity of cloud computing services that raise confronting legal and policy issues. Data travelling in different jurisdiction and geographical distribution of cloud services always need more research and investigation on legal complications. Security, confidentiality and trust are the main key elements from the user's point of view. In this paper cloud computing security weaknesses, threats and evidence centralization are investigated. Monitoring cloud traffic helps to collect evidence after malicious intruders penetrate into the systems. The system can be hold idle for a long time, or the data can be stolen, changed or deleted. In Section 4 the new methodology is summarised briefly [2,3].

2 Virtualization of cloud computing

Virtualization is the mechanism, which cloud computing applies. Virtualization of servers gives mobility and reduces the overall costs. It is defined as a pool of virtualized computer resources. Virtualisation provide more efficient use of the hardware resources Virtualized cloud computing provides virtual machines to the end-users to execute their software and applications on remote sites.

It helps to install operating system and software directly on physical hardware. Virtualization provides possibility to add new services without investing in new infrastructures, training new personnel or licensing new software. Virtual storage is also offered to end users [2].

There are problems of virtualization such as some virtual machines can be dormant (not running) at some point in time [3]. These virtual machines cannot be easily kept up to-date, because typically this would require the virtual machines to be started, updated and shut down again, which is not only time-consuming but also is an irregular task to do on a permanent schedule. In Figure 1 basics of virtualized cloud computing are shown.

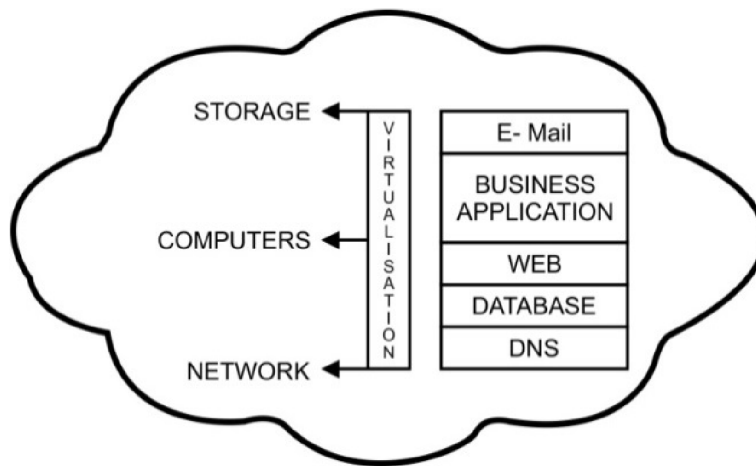


Figure 1: Cloud computing virtualization [2]

In cloud, virtual machine is created by a hypervisor or virtual machine monitor (VMM) which is a piece of computer software or hardware. Thus, cloud computing uses hypervisor to manage virtual devices. A computer on which a hypervisor is running one or more virtual machines is defined as a host machine. Each virtual machine is called a guest machine. The hypervisor presents the guest operating systems with a virtual operating platform and manages the execution of the guest operating systems. Multiple instances of a variety of operating systems may share the virtualized hardware resources [2]. The hypervisor lets us show the same application on lots of systems without having physically copy that application onto each system. Because of the hypervisor architecture, it can load any (or many) different operating system as though it were just another application (Figure 2). Therefore, the hypervisor is a very practical way of getting things virtualised quickly and efficiently [2].

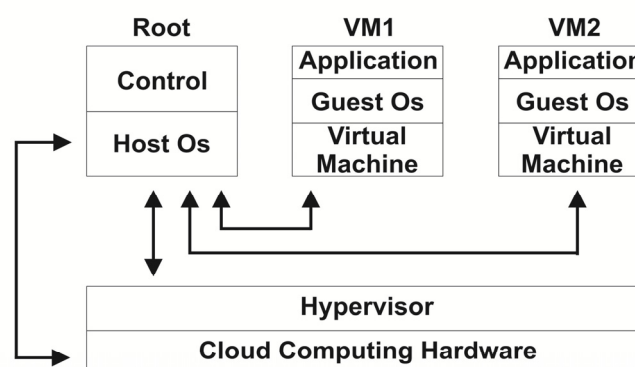


Figure 2: Cloud Virtualization [2]

VM can be dormant. Dormant VMs cause an important security risk because VMs that are offline do not receive security upgrades/patches. Dormant VMs should not be interpreted as if they were deleted. Inactive VMs present an important security threat and therefore must be identified and appropriate security mitigation should be applied. When a VM is dormant, the log files are absent [3]. Previously, researchers found a tool called ‘Update checker’ which provides a partial solution to the problem. The tool basically tracks the dormant VMs which missed latest updates and corrects them [4].

Virtualization weaknesses such as volatility of data are a major concern for the security issue. The result of it is absence of logs. Thus, identifying all those threats is necessary to bring solutions especially in case of a malicious attack it will be difficult to collect evidence.

3 Cyber threats

Malware is the reality of the digital world. Malicious hackers uses number of methods and tools and knowledge to hijack the systems. Thus, no single point of view will suffice to defence the systems. Cloud computing Virtual machine attacks are briefly introduced.

Understanding the enemy is the key to defence. The first step to protecting against malware is to understand what happens during an attack. From the initial overtures by the attacker in assessing a target to stealing information and sending it to the attacker’s home base.

Attacks, however, often follow predictable patterns in four steps: break in, latch on, expand, and gather and exfiltrate. The modern defences are able to recognize those patterns to more effectively combat malware threats. To achieve the full protection available, organizations need a comprehensive solution that goes beyond a special-function, point-product approach [5, 6].

How a guest VM can attack the host is shown in Figure 3. A VM can be captured by a malicious attacker or an attacker himself can organize the attack to damage the hypervisor or inject a malware. Attackers, besides attacking cloud infrastructure can base the cloud to launch attacks on other systems [7].

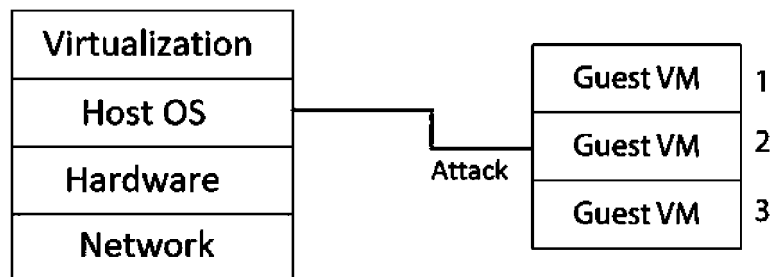


Figure 3: Guest VM attack [7]

4 Evidence Centralization

As mentioned before, cloud computing has complex nature such as virtualization of thousands machines spread all over, in case of an attack it is difficult to access evidence data. Thus, centralisation of log file as evidence is useful.

4.1 Evidence Log Data

The main purpose of collecting log data is to justify crime and document the proof of evidence in order to present in the courtroom if necessary. Thus, maximising performances for specific tasks that are to be carried out and increase security levels.

Through centralised and compressed data storage, it is possible to optimise the use of storage resources, as well as immediately identify relevant events in order to quickly obtain specific log fragments.

From the same centralised control interface it is also possible to define data retention and rotation policies for each monitored file (depending either on their current size, the periods of time elapsed, or a combination of the two), in order for the tool to be completely independent in handling operations that would turn out to be absolutely tedious and time-consuming if performed manually [8].

4.2 Log files as digital evidence

Log data are grouped into log files depending on the type of information registered. Logs can roughly be divided into two groups:

- logs regarding the status, processes and the operation of the device – System Logs;
- logs regarding the activity of the end users – Access Logs [9].

Log files can be very important sources of forensic evidence. A log file is a file that contains a list of events, which have been "logged" by a computer. Log files are often generated during software installations and are created by Web servers, but they can be used for many other purposes as well. Most log files are saved in a plain text format, which minimizes their file size and allows them to be viewed in a basic text editor.

4.3 Computer forensics

Cloud forensics refers to digital forensic investigations performed in cloud computing environments. It is basically concerned in gathering evidence from computing equipment and storage devices that can be used in a court of law as digital evidence. Computer forensics is the procedure of detecting computer system to determine valid evidence [10]. The goal of computer forensics is to examine digital media in a forensically sound manner with the aim of identifying, preserving, recovering, analysing and presenting facts and opinions about the information. The discipline involves similar techniques and principles to data recovery, but with additional guidelines and practices designed to create a legal audit trail.

The forensic requirements for the logging file are shown in Fig. 4. For the design and analysis of log files, it is important that in case of a malicious attack to answer questions such as ‘Who’, ‘What’, ‘When’, ‘Where’, and ‘How’. Sometimes, it is not possible to get answer for all these questions, depending on the nature of attacks and the layer of the cloud computing in which it is happening.

For the design and analysis of log files, it is important to find answers for questions such as ‘Who’, ‘What’, ‘When’, ‘Where’, ‘How’ in case of a malicious attack. Sometimes, it is not possible to get answers for all these questions, depending on the nature of attacks and the layer of the cloud computing in which the attack takes place. In terms of cloud computing, anomalies of VMs are seen as vulnerability and they are potentially inviting malicious attackers. Thus, VMs anomalies are important for the total security of cloud computing [10].

4.4 The need for log centralization

In cloud computing there are some difficulties with collecting evidence for digital forensic process. Several characteristics of cloud computing inhibit the process of cloud forensics data acquisition. Acquiring data from the virtual machine is complex because some of the clouds layers are not accessible and local and network storage are distant.

Cloud computing is a multi-tenant system, while traditional computing is a single owner system. In cloud computing, multiple VMs can share the same physical infrastructure. The multi-tenancy characteristic also brings novel side-channel attacks that are difficult to investigate [11].

In cloud infrastructure, log information is not located at any single centralized log server; rather logs are decentralized among several servers. Log information is spread to multiple servers in different locations. Bringing up together the log data is time consuming and not easily scalable. Thus, centralization of logs is recommended in this article.

There are some difficulties of acquiring log data from various tiers and layers of cloud such as PaaS and SaaS layers. There are dependencies between these cloud service models [12]. PaaS as well as SaaS is hosted on top of IaaS; thus, any breach in IaaS will impact the security of both PaaS and SaaS services, but it may also be true the other way round; any attack to any cloud service layer can compromise the upper layers [13]. In SaaS, customers can not get any log of their system, unless the Communicating Sequential Processes (CSP) provides the logs. In PaaS, it is only possible to get the application log from the customers. To get the network log, database log, or operating system log we need to depend on the CSP.

OS and VM logs are closed to outsiders, customers and investigators. However, in this article, a cooperative CSP is assumed. Hence, in case of severe crime, with a warrant from the court, an investigator can get access to VM logs from all layers. Virtual machines must have static IPs to be able to identify the crime case. Another issue is the volatile data case. When a virtual machine is turned off, the data will be lost unless the image is taken. In IaaS, to avoid the VM volatile data problem, persistent data storage could be applied [14].

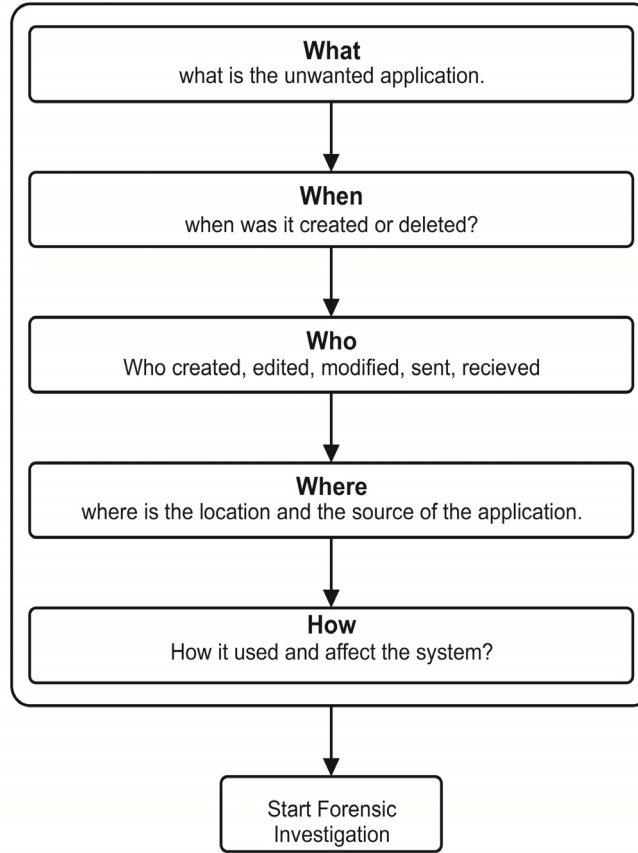


Figure 4: Digital forensic logging format

Volatile data cannot be sustained without power. After terminating a VM, all the data will be lost. In order to provide on demand computational and storage service, CSPs do not provide persistent storage to VM instances. The solution for VM data can be, by storing an image of the VM instance. However, migrating the image of the instances is sometimes not possible because minimum amount of data is approximately 2 GB [15].

4.5 Anomaly detection

Anomaly detection is the process of detecting something unusual, unwanted compared to something expected. In terms of cloud computing detecting unusual (or suspicious) behavior in order to identify malware or something unusual is the anomaly situation. Anomaly based detection approach is used in two phases. First phase is training phase which identifies the behavior of the system in absence of attacker and machine learning technique. Second phase compares the user behavior against the current user behavior. If there are any changes in current user behavior then it identifies whether it is malware or not [16]. In our case, VM behavior or anything unusual happens for instance dormant VM case can be considered as anomaly of VMs.

Examples of what anomaly detection could identify include:

- Accessing unusual location or unusual time of day.
- Using cloud services features not typically used.
- Changing personal information.
- Any of the VM dormant or terminated long time.
- Huge amount of bandwidth in one transfer.

For the testing platform, Eucalyptus cloud is chosen. Eucalyptus cloud is a Linux based OS and CentOS is a version OS of the Linux. The design of transferring logs to the forensic workstation shown in Figure 5.

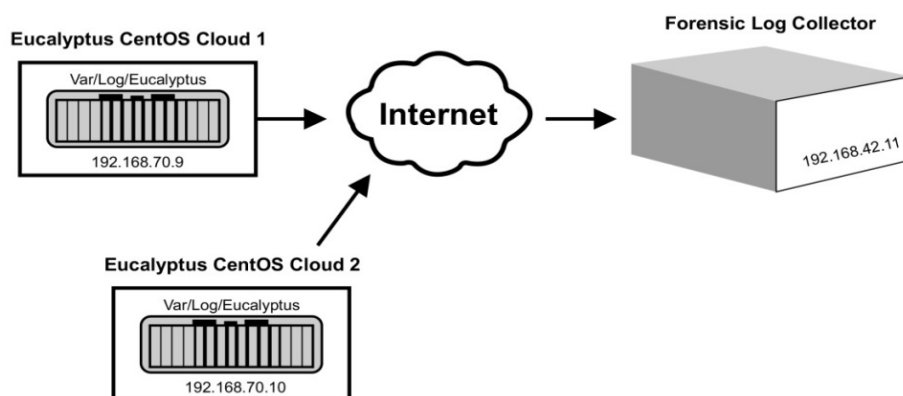


Figure 5: Log collection [17]

The current release of the Eucalyptus only works over GNU/Linux OSes; in this type of OS, important system log files and relevant services/applications selected logs are listed.

The log files should contain information listed below:

```
/var/log/messages: General and system related logs
/var/log/eucalyptus/cc.log: Shows the state of an instance log
/var/log/eucalyptus/cloud-debug.log: Debug message log
/var/log/eucalyptus/cloud-error.log: Error messages log
/var/log/eucalyptus/cloud-output.log: Hypervisor log
/var/log/eucalyptus/startup.log: Hypervisor log
/var/log/libvirt/: Libvirt virtualization API logs
/var/log/eucalyptus/postgresql: Cloud database log
/var/log/eucalyptus/jetty: Network log
/var/log/eucalyptus/nc.log: Guest OS log
```

4.6 An example of the results of acquiring evidence logs

As seen in Table 1 below, after collecting the specific log files, such results can be seen. An idle VM (or dormant VM) and unknown IP address are traced.

Log file	layer	Info	Location	Attack type	Anomalies	Severity
/var/log/Eucalyptus/local-net	1	VM IP address		N/A		Info-warning
/var/log/Eucalyptus/axis2c.log	3	Error logs		N/A		Info-Warning
/var/log/Eucalyptus/cc.log	3	Shows the state of an instance		N/A		Severe
/var/log/Eucalyptus/cloud-debug.log	3	Receive message from absent client		N/A		Info-Warning
/var/log/Eucalyptus/cloud-error.log	3	Error message		N/A		Error
/var/log/Eucalyptus/cloud-output.log	3			N/A		Info-Warning
/var/log/Eucalyptus/startup.log	3	Hypervisor		N/A		Info
/var/log/libvirt/ibvirt.log	4			N/A		Info-Warning
/var/log/Eucalyptus/postgresql		Cloud database log		N/A		Error
/var/log/Eucalyptus/jetty	1	Network		N/A	UnknownIPAddr ess:	Warning
					192.168.70.13	
/var/log/Eucalyptus/nc.log	5	Guest OS	VM6	MITM	Idle VM	Severe

Table 1: Results of collected logs [17]

4.7 Advantages of centralization

Advantages of centralization Log records have a major role in cloud computing. They help in the detection of anomalies in the system. Centralized logging brings two important benefits. Firstly, it gathers all of the log records in a single location. This process significantly facilitates log analysis and correlation tasks. Having the data all in the same place helps in forensic readiness resulting in quicker, coordinated response to incidents. Owing to centralized data IaaS providers can create a dedicated ready for use forensic server within the cloud. Secondly, it provides a long term secure storage area for the log data. After the establishment of a central log repository, the next step is the introduction of centralized analysis techniques [18].

5 Conclusions

In this article, cloud computing virtualization, risk factors and centralization of evidence data collection through a forensic analysis are discussed. Background and literature review is also given summarizing some of the previous work and other authors' points of views.

In the future, the authors plan to standardise and automate similar solutions to collect evidence of cyber-crime for the cloud forensic analysis.

Abbreviations

CSP – Communicating Sequential Processes

VM – Virtual Machine

IaaS – Infrastructure as a Service

PaaS – Platform as a Service

SaaS – Software as a Service

References

- [1] J. Leach, “The Rise of Service Oriented IT and the Birth of Infrastructure as a Service”, March 20, 2008. http://advice.ciocom/jim_leach/the_rise_of_service
- [2] Y. Turel, “Cloud Computing and Virtualization: A Comprehensive Survey”, *Przegląd Elektrotechniczny*, 89, 12/2013.
- [3] Y. Turel, “Survey on Cloud Computing Vulnerability and Cyber Attacks: A Defensive Approach”, *Task Quarterly*, 17, No 1–2, 95–107, 2014.
- [4] K. Bateman, “Don't let dormant virtual machines threaten data centre security”, 21/06/2011, <http://www.computerweekly.com/news/1370369>.
- [5] R. Schwarzkopf, M. Schmidt, C. Strack, S. Martin, B. Freisleben, “Increasing virtual machine security in cloud environments”, *Journal of Cloud Computing: Advances, Systems and Applications*, 1-12, 1:12 doi: 10.1186/2192-113X-1-12, 2012.
- [6] IBM Software Thought Leadership White Paper, “Defending against malware: A holistic approach to one of today's biggest IT risks”, 2014.
- [7] R. Meadows, “Cisco Router and Switch Forensics: Investigating and Analysing Malicious Network Activity”. Elsevier Science, 2009.
- [8] http://www.securelog.it/en/page/show/page_id/25.
- [9] I. Ivanović, M. Kukoleča, J. Palibrk, “Centralised web traffic filtering system”, (AMRES BPD 113), 2013.
- [10] http://en.wikipedia.org/wiki/Computer_forensics, accessed 30 January 2014.
- [11] N.M. Turab , A.A. Taleb, S.R. Masadeh, “Cloud Computing Challenges and Solutions”, *International Journal of Computer Networks & Communications (IJCNC)*, 5, 5, 209-216, 2013.
- [12] “Cloud Security Alliance Security guidance for critical areas of focus in Cloud Computing V3.0”, 2011, <https://cloudsecurityalliance.org/guidance/csaguide.v3.0.pdf>.
- [13] OWASP, “The Ten most critical Web application Security risks”, https://www.owasp.org/index.php/Category:OWASP_Top_Ten_Project, 2010.
- [14] R. Poisel, S Tjoa, “Discussion on the challenges and opportunities of cloud forensics”. *Proc. of the 2nd Multidisciplinary Research and Practice for Information Systems (CD-ARES'12)*, Prague, Czech, LNCS, vol. 7465. Springer-Verlag, 593–608, 2012.
- [15] J. Dykstra, A.T. Sherman, “Acquiring forensic evidence from infrastructure-as-a-service cloud computing: Exploring and evaluating tools, trust, and techniques”, *Elsevier Digital Investigation*, 9, 590-598, 2012.

- [16] “Cloud Security Alliance, Top Threats to Cloud Computing V1.0”, 2010, <https://cloudsecurityalliance.org/research/top-threats>.
- [17] Y.Turel, “Cloud Computing and Digital Forensics Investigation: Facade Logging As An Acquisition Method”, Springer, Computing, Comp-D-14-00047, 2014.
- [18] M. Chapple, “Is centralized logging worth all the effort?” Enterprise Compliance, 2008, <http://searchsecurity.techtarget.com/answer/Is-centralized-logging-worth-all-the-effort>.