Cloud computing and computer forensics for business applications

Ludwig Slusky
California State University

Parviz Partow-Navid
California State University

Mohit Doshi
California State University

ABSTRACT

The paper reviews issues related to teaching computer forensics with Cloud Computing. It examines the key issues that Computer Forensics is facing today and the challenges and opportunities that Cloud Computing brings for computer forensics. The impact of Cloud Computing can be seen in all basic tasks of Computer Forensic Investigation: Data Acquisition and Validation, Bookmaking data, File Signature Analysis and Hash Analysis, Analysis of Data, securing evidence files and reporting.

Colleges and university are beginning to identify and manage applications and services available through Cloud Computing. For example, Ohio State University and many other universities (Indiana University 2009) issued CC Guidelines, Key challenges and risks for Teaching, administrative support, and research. IBM and Google had started programs on college campuses to promote computer-programming techniques for clusters of processors known as "clouds"

Teaching Computer Forensics in Cloud Computing environment offers opportunities far beyond traditional computer forensics at the advance edge of technological innovations.

Key words: Cloud computing, computer forensics, computer security, teaching.

INTRODUCTION

Contrarily to traditional onsite application architecture where applications are residing in client machines or in a server accessible via client-server link of a Local Area Network, the Cloud Computing (CC) offers shared computer application resources installed on a central server and accessible via the Internet.

National Institute of Standards and technology (NIST) defined Cloud Computing as "a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models" (Mell 2009).

Four deployment models are: private cloud operated solely for an organization, community cloud shared by several organizations, public cloud available to the general public, and hybrid cloud as a composition of two or more clouds from the above. Many companies are already using cloud applications, and increasingly more companies are in transition to Cloud Computing (Frowen 2010).

Computer Forensics investigation, typically confined to computer components and media devices, is focused on recovery of legally admissible evidence in various forms of digital information. Thus, a Computer Forensics Analyst should have direct access to the components and devices that are subjects to investigation.

Computer Forensics

Computer forensics is considered to be the use of analytical and investigative techniques to identify, collect, examine and preserve evidence/information which is magnetically stored or encoded (Monica 2007). Evidence might be sought in a wide range of computer crime or misuse, including but not limited to theft of trade secrets, theft of or destruction of intellectual property, and fraud. Computer specialists can draw on an array of methods for discovering data that resides in a computer system, or recovering deleted, encrypted, or damaged file information. Any or all of this information may help during discovery, depositions, or actual litigation. Computer forensics provides digital evidence of a specific or general activity. The common scenarios include:

- Child pornography
- Industrial espionage
- Criminal fraud and deception cases
- Unauthorized access to and/or disclosure of protected information
- Cyber warfare attacks

Computers can be used for virtually any type of crime, including cyber stalking, pornography, identity theft and espionage. It is the job of computer forensics professionals to use technology to catch and stop these criminals. To do so effectively, however, requires periodic updates to stay current with the latest technology and trends in computer investigation. The key issues that Computer Forensics is facing today are:

1. New data breach techniques and new vulnerabilities are evolving: Attackers are increasingly superior at identifying and exploiting vulnerabilities such as SQL injection (SANS 2008). Once exploited, attackers utilize new and sophisticated malware to maintain their presence and infiltrate critical data. Both the attack and the malware evade perimeter and host defense tools such as firewalls, intrusion prevention/detection systems, and anti-virus products. The simplicity and effectiveness of these attacks allow for a full infrastructure compromise resulting in Personal Identifiable Information (PII) and credit card/debit account information compromised.

2. Insufficient forensic preparedness: In recent years, additional and expanded legal and regulatory compliance requirements revealed the lack of preparedness of many organizations for computer security incidents and data breaches. Organizations have been very reluctant to view external sources as real threats to critical data repositories. Recent advances in technology and techniques continue to move incident response and forensic analysis of those incidents forward. However, the responsibility to respond remains on the shoulders of relatively untrained first responders. As a result, market demand for qualified incident response personnel is higher. Thus, making it very difficult to recruit and retain these individuals. Companies must now invest significant time and training their personnel.

3. Insufficient forensic expertise: The PI license requirements of some states have made digital forensic examiners apprehensive in performing eDiscovery or Forensics in a few states. In addition, the ambiguous PI licensing requirements force digital forensic examiners and the organizations that hire them to consult with lawyers at high costs and with more complexity for doing business. State legislators are seeking a way to verify the qualifications of digital forensic professionals that will result ultimately in several national level certifications being recognized, a marked improvement over the initial plan of requiring a PI license.

4. Growing size of data requiring computer forensics: Ever-expanding file system sizes have forced investigators to rely increasingly on ephemeral information about the state of a system. About 95% of world's information is being generated and stored in digital format. Only about one-third of documentary evidence is printed out. Additionally, investigators have focused on gathering only pieces of the file system to help save acquisition and analysis time to reduce costs. There has been a price to pay for this transition, however, in that piecing together a complex series of technical activity becomes more difficult without full drive images.

5. Rising eDiscovery: eDiscovery legal changes to the Federal Rules of Civil Procedure in 2006 resulted in a new subsector in forensic products and services that have targeted responding to the overnight creation of a $1 billion dollar industry. The computer forensic industry has entered heavily into the eDiscovery landscape for litigation. With significant potential penalties for lack of compliance, there has been a focus on the timeliness of document/email collection for litigation purposes. Digital forensics has emerged as a viable option in eDiscovery due to the defensibility of the process, authenticity of collected evidence, and lower cost over the traditional methods.

6. Expanding mobile data forensics: The need for robust forensics capture and analysis tools for cell phones, iPods, and PDAs increased greatly during the past two years. An increasing number of investigations are dependent on data stored in these devices that often have different formats and means of access than the traditional computer hard-drive-centric techniques used by most investigators. The need is driving some investigators to create homegrown tools for such analysis, while others have simply avoided this potentially very useful source of case information.

7. Expanding volatile data forensics: There has been amazing progress in the area of memory forensics over the last few years. Volatile memory collection and analysis dramatically augmented digital investigations and helped address many new challenges such as encryption and recovering key evidence that might only exist for seconds on a computer.

Computer Forensic in Cloud Environment

Cloud applications operate in a computing environment that is different from the traditional onsite client applications environment. The additional complexity of forensics investigation of cloud applications arises from extensive use of Virtual Machines in a shared "cloud" environment and access to the VM servers over the Internet.

Cloud Computer Forensics investigation may be focused on two outcomes: evidence discovery for prosecution of computer crime and data recovery from malicious data manipulation or a crash caused by "cloud" environment vulnerabilities. With this new paradigm come challenges and opportunities for computer forensics. Let us review some of them:

- When an application is executed in the cloud, the current data content is also stored, accessible, and largely protected from intentional physical destruction in the cloud. That generally simplifies collection of evidence files. However, cloud environment makes evidence based on registry entries, temporary files and logs unavailable when the user exits the session.
- One way to deal with data availability shortcomings is to have assistance of the Cloud Computing provider in retaining and extracting data at the source. However, such approach runs into another obstacle: admissibility in the court of law the data that is collected and preserved by an employee who is not licensed Computer Forensics Investigator and who does not follow the required procedures that guarantee the collected data from contamination.
- Physical access to computer components is a simple procedure in traditional (non-cloud) environment. Confiscation of computer components, however, becomes increasingly complicated in the Cloud Computing environment requiring additional legal steps and much more time. In Computer Forensics, where the time is of the essence, such delays may be detrimental to the recovery of evidence.

According to Frowen (2010) currently "… there is no foolproof, universal method for extracting evidence in an admissible fashion from cloud-based applications, and in some cases, very little evidence is available to extract. As such, Cloud Computing represents just one of the

fast-paced technological developments that is presenting an ongoing challenge to legislators, law enforcement officials and computer forensic analysts."

Cloud Computing is particularly conducive to specific vulnerabilities arising from the use of Virtual Machines (VMs). Thus, an attacker can initiate a "cross-VM attack" against a target server in a shared Cloud Computing environment by implanting its own malicious VM in close proximity to the attack target, i.e., placing it on the same server (Wilson 2009). The attacker can increase probability of such close proximity by performing some preliminary detective investigation. For example, the research shows that a new VM is more likely to be created on the same server as a co-resident with another VM if both of them are created at the same (or near the same) time (Talbot 2009). If a victim has the capability of generating new VMs based on the service demands, then there is a way for the attacker to manipulate such close timing by flooding the victim with requests for services and forcing it to create a new VM, so the attacker can create its own malicious VM at the same time.

Such risk can be mitigated in part by enforcing access control procedures that permit access to VMs only from designated accounts. Such access control procedures (as well as other control measures arising from the use of VMs in the "cloud" environment) may in turn simplify computer forensics investigation. Ideally, one account per VM would be the closest simulation of the client computing environment. Fortunately, there is a trend toward underutilization of VMs that reduces the number of accounts with permitted access.

Another contributing factor that facilitates Computer Forensics investigation of evidence files collected in a "cloud" is possibly weaker encryption of data in VMs, which often have limited capabilities to generate random numbers used in encryption (Perez 2009).

Microsoft Security Development Lifecycle (SDL) – Version 5.0 (Microsoft 2010) addresses some of the security issues of cloud applications with solutions focused on development and implementation of security policies. Such adherence to strict policies can assist in computer forensics investigation.

Computer Forensics in Cloud

One of the first steps in Forensic investigation of cloud applications is cloud cartography, i.e., mapping the structure of cloud network. The cloud-internal IP address assigned to a VM is strongly correlated with creation parameters. Cloud computing as a separate environment offers several indisputable advantages for Computer Forensic investigation (Balding 2008). Infrastructure as a Service (IaaS) architecture used in clouds supports forensic readiness. A dedicated forensic server can be created but kept offline until it is needed. A copy of the VM can easily be distributed for use as new sources of evidence need forensic investigation.

Cloud computing significantly reduces time for data acquisition, data copying, transfer, and data cryptanalysis. Used in addition to onsite computing environment, cloud computer forensics decreases evidence acquisition time and onsite service downtime. It allows near instant imaging of the onsite data and placing it preserved in the cloud for investigation without any detrimental effect on the ongoing onsite operations (Morill 2008).

Copies between two co-resident VMs (residing in the same physical server) are made very fast. If one of these VMs is the object of investigation and the other one is the forensic investigation platform, then the evidence transfer rate between them is almost negligible. In addition, forensic image verification time is reduced if a Cloud Application (as some cloud application do, e.g., Amazon) generates cryptographic checksum or hash.

Cryptanalysis of data is the most complex and time consuming operation in computer forensics and requires sufficient CPU size and memory capacity. Typically, the computing power of cloud environment far exceeds capabilities of portable or stationary forensic computers, thus reducing the time needed for cryptanalysis and discovery of passwords.

Still, the prevailing thought according to Bigsey (2009) is that "CC vendors cannot ensure that data which could be used as evidence will be complete, retrievable or verifiable… Investigating inappropriate or illegal activity may be impossible in Cloud Computing, Gartner warns."

CC power can reach monstrous dimensions. Texsill (2010) sites sources describing a cloud consisting of 40000 VMs with 512 physical servers and roughly 1000 users; in addition, the cloud datacenter servicing VMworld 2009 contained 48TBs memory. Trying to identify evidence data in this massive data storage among multiple VMS and users could be a very complex project.

The basic tasks of Computer Forensic Investigation are Data Acquisition and Validation, Bookmaking data, File Signature Analysis and Hash Analysis, Analysis of data, securing evidence files and reporting. The CC environment posses several complicated legal issues for Computer Forensics. Data acquisition from VMs is not straight forward as it could be in the onsite client computing. The forensically sound bit-by-bit methods of data collection, typical in onsite forensic investigation, may not be fully applicable or court-admissible when used in the CC environment.The data itself is dynamic and may not reside permanently in one VM or physical location. So, there is an additional task of determining the location where a specific VM resided in the cloud.

Privacy issues are also more complex as the users of various VMs in the cloud may belong to different organizations with different policy constraints. Data belonging to one VM co-mingles with data belonging to a different VM on a particular server. Computer Forensic tools are well suited to go after a physical server. Therefore, separation of data on the server that belongs to the target VN from data that belongs to other VMs on the same server could be problematic. That, in turn, affects the privacy issues relate to multiple VMs and tenants on the server.

Cloud Forensics is already shaping up as a business. At least one cloud provider, Terramark, is developing a cloud for forensic investigative work for the FBI using forensics tools from Netwitness and Guidance Software (Texiwill April 2010).

Teaching Cloud Computer Forensics

Colleges and university are beginning to identify and manage applications and services available on the Web as Cloud Computing. Within the university, the confidentiality, integrity, availability, use control, and accountability of institutional data and services are expected to be ensured by a suite of physical, technical, and administrative safeguards proportional to the sensitivity and criticality (i.e., risk) of those information assets and services. These safeguards help protect the reputation of the university and reduce institutional exposure to legal and compliance risks. Thus, Ohio State University and many other universities (Indiana University 2009) issued CC Guidelines, Key challenges and risks for Teaching, administrative support, and research (OSU 2010).

Since Cloud Computing made its way into highly scalable business applications, it is gaining more and more attention from researchers, software engineers, and IT media. Software companies like IBM and Google had started programs on college campuses to promote computer-programming techniques for clusters of processors known as "clouds" (Bulkeley 2007). IBM continues to work with universities and educational institutions worldwide, giving students access to Cloud Computing technologies to help them complete research projects that aid in the development of remote regions and socio-economic conditions all over the globe.

Yahoo! Inc. expanded its successful research partnership with Carnegie Mellon, University of California at Berkeley, Cornell University and the University of Massachusetts at Amherst, to begin using Yahoo's M45 computing cluster to advance Cloud Computing research (Carnegie Mellon 2009). Yahoo! cluster, also known as M45, has been operational since November 2007 and in use by Carnegie Mellon. The cluster has approximately 4,000 processor-cores and 1.5 petabytes of disks.

Yahoo!'s M45 cluster runs Hadoop, an open source distributed file system and parallel execution environment that enables its users to process massive amounts of data. Apache Hadoop is an open source project of the Apache Software Foundation, to which Yahoo! Engineers have been the primary contributors to date.

Recently, Intel's Open Cirrus, a global open-source test bed for the advancement of Cloud Computing partnered Carnegie Mellon's School of Computer Science for research and education (Carnegie Mellon 2010). "The phenomenal success of world wide web (www) was built on open-source (the Linux-Apache-MySQL-Python stack) in partnership between universities and academia. At Intel, we hope that Open Cirrus will become both enable academic research on cloud software from infrastructure to applications, but also recreate that dynamic partnership to power the development of the cloud as a global, open, interoperable, information and compute platform for the 21st century," said Andrew Chien, vice president of Intel Labs and director of Future Technologies Research.

Clouds are being integrated into all sectors of business, from law, health, and education that have a large amount of data to store, to science, forensics, and engineering firms that have large amounts of data to analyze and process. Computer forensic investigations involve the scientific analysis of computer equipment to recover the evidence. The content of each drive has to be carefully replicated so that original evidence cannot be contaminated. The information could be stored within the cloud, and then a simple click of the mouse could potentially produce an exact image of the current state of the firm's data, allowing the investigation to progress fast, as replicating the complete drive could be time consuming (Frowen 2009). But still there is less evidence available how cloud-based applications help to extract forensic data. Thus, teaching Cloud Computing forensics could be very helpful to legislators and law enforcement officials.

Forensics Lab

In Cloud Forensics Lab, the actual forensics applications are run on several physical servers. Each student operates one dedicated VM. Virtual machines of different students may sit on the same or different physical servers. The number of VMs in cloud envenomed can be expanded or contracted dynamically to meet students' demand and, thus, providing tremendous efficiencies in utilizing the available Cloud Computing resources. Hardware Tools frequently used for computer forensics are as follows:

1. FRED Systems: FRED system is the complete forensic hardware and hardware solution. The 19" LCD monitor is included. The available OSs on this system is MS-DOS 6.22, Wins 98, Wins XP Pro, and Linux 9.1 Pro. Some software such as Norton GHOST 10.0 & 2003, Nero DVD/CD Authoring Software, DriveSpy, Image, PDWipe, PDBlock, and PART are included. Moreover, the toolbox is also included which contains all the necessary cables, adapters, digital camera, and security screwdriver set, etc.

2. Forensic Network: A Forensic Network is a series of processing and imaging computers connected and integrated directly with a high-speed, high-capacity server to share resources. The file server operates as the core of the Forensic Network and can be used as a central storage facility for Forensic Images as well as applications software for use by the client processing and imaging stations. Workstation clients on the network perform the actual imaging and processing tasks, while the central file server stores the images and case work.

3. Forensic Write Blockers**:** Digital Intelligence designs and offers parallel IDE, serial ATA and SCSI hardware write blockers, as well as other custom solutions, to effectively address specific write blocking requirements. Learn how our UltraKit, UltraBlock, FireFly, FireBlock, SCSIBlock and FireChief devices can maintain the integrity of your evidence.

4. Forensic Devices:

- Fannie - forensic area network numerous imaging enclosure
- Rack-a-Tacc password decryption
- Tacc1441 Hardware Accelerator
- Modular Accessories
- Forensic Duplicator
- Hardcopy 3 & Hardcopy 2
- Shadow 2

Software Tools frequently used for computer forensics are as follows:

1. Digital Intelligence Software: Digital Intelligence has created several forensic software tools in-house specifically for forensic use. These tools include DriveSpy, Image, Part, PDBlock and PDWipe.

2. Accessdata: Since 1987, AccessData has been a leader in password recovery and applied cryptography. The forensic tools available are Ultimate Toolkit, Forensic Toolkit, Password Recovery Toolkit and Registry Viewer.

3. Guidance Software**:** EnCase Forensic Edition, by Guidance Software, is the world's leading solution for computer investigations and forensics.

4. Paraben Forensic Tools: Paraben has forensic software for PDAs, password recovery, text searching, data acquisition, e-mail examination and more.

5. Hot Pepper Technology: Authors of EMail Detective, a dedicated software solution for recovering and reconstructing AOL email. EMD is the most comprehensive AOL extraction tool available to forensic agencies.

6. Stepanet DataLifter: Stepanet's DataLifter is a suite of products built on years of nvestigative experience. These tools have been specifically designed to assist with Computer Forensics, Information Auditing, Information Security and Data Recovery.

CONCLUSION

Cloud Computing and its impact on computer forensics will continue to grow. The ability to access and analyze massive data sets is becoming increasingly important and critical to the advancement of computer forensic research. Cloud computing has matured to a point where it is considered a mainstream technology service; it has become a way to reduce costs, improve service, and much more agile computing. People touch the cloud every day without knowing it -- by sending instant messages and sharing files easily over the Internet, and staying connected with social networking tools.

Cloud Computing is offering utility-oriented IT services to users all over the world. Clouds help to drive the design of data centers by architecting them as networks of virtual services so that users can access and run applications from anywhere in the world. Cloud computing offers significant advantages to IT firms by setting up basic hardware and software infrastructures, and making it possible for creating and innovating business values for their services.

Cloud Computing and legal expertise must be combined in order to discover, develop and utilize digital evidence for forensic experts. The process or technology used must conform to both law and science and must be robust to ensure that all probative information is recovered. Cloud computing could be helpful to explore stored digital information but risks and challenges has to be considered while using Cloud Computing.

References

Balding, Craig (2008). Assessing the Security Benefits of Cloud Computing. Retrieved June 4, 2010 from http://cloudsecurity.org/blog/2008/07/21/assessing-the-security-benefits-of-cloud-computing.html.

Bigsey (2009). Cloud Computing & The Impact On Digital Forensic Investigations. Retrieved June 4, 2010 from http://www.zdnet.co.uk/blogs/cloud-computing-and-the-impact-on-digital-forensic-investigations-10012285/cloud-computing-and-the-impact-on-digital-forensic-investigations-10012286/.

Bulkeley, William (2007). IBM, Google, Universities Combine 'Cloud' Forces. Wall Street Journal. Retrieved August 18, 2010, from http://online.wsj.com/article/SB119180611310551864.html.

Carnegie Mellon (2009). Three Universities Join SCS in Yahoo! Cloud Computing Research. Retrieved August 18, 2010, from http://www.cmu.edu/news/blog/2009/Spring/cloud-computing-at-scs.shtml.

Carnegie Mellon (2010). Fulfilling the Vision of Essential Computing. Retrieved August 18, 2010, from http://www.cmu.edu/corporate/news/2010/intel.shtml.

Frowen, Andrew (2010). Cloud Computing and Computer Forensics. Retrieved August 19, 2010, from http://www.artipot.com/articles/384511/cloud-computing-and-computer-forensics.htm.

Indiana University (2009). Use of Cloud Computing. Retrieved August 17, 2010 from http://informationpolicy.iu.edu/resources/articles/cloud_computing.

Mell, Peter and Grance, Tim (2009). The NIST Definition of Cloud Computing. Version 15. NIST. Retrieved June 4, 2010 from http://csrc.nist.gov/groups/SNS/cloud-computing/index.html.

Microsoft (2010). Microsoft Security Development Lifecycle (SDL) – Version 5.0. Retrieved June 4, 2010 from http://www.microsoft.com/downloads/details.aspx?FamilyID=7d8e6144-8276-4a62-a4c8-7af77c06b7ac&displaylang=en.

Monica (2007). Computer Forensics World. Retrieved May 13, 2010 from http://www.computerforensicsworld.com/.

Morill, Dan (2008). Cloud Computing Making Forensics Easier. Retrieved June 4, 2010 from http://www.cloudave.com/link/Cloud-computing-making-forensics-easier.

OSU (2010). Cloud Computing Guidelines for Teaching, Administrative Support, and Research. Ohio State University. Retrieved June 4, 2010 from http://cio.osu.edu/policies/ccg_V62.pdf.

Perez, Sarah (2009). The Cloud Isn't Safe?! Retrieved June 3, 2010 from http://www.readwriteweb.com/archives/the_cloud_isnt_safe_or_did_blackhat_just_scare_us.php.

SANS (2008). SANS Top 7 Forensic Trends, Retrieved May 13, 2010, from http://computer-forensics.sans.org/community/top7_forensic_trends.php.

Talbot, David (2009). Vulnerability Seen in Amazon's Cloud-Computing. Retrieved June 4, 2010 from http://www.technologyreview.com/computing/23792/.

Texiwill (April 2010). Virtualization of Forensics: How Different is It? Retrieved June 4, 2010 from http://www.virtualizationpractice.com/blog/?p=5126.

Texiwill (February 2010). Virtualization and Cloud Missing Key Features: Auditing and Forensics. Retrieved June 4, 2010 from http://www.virtualizationpractice.com/blog/?p=4110.

Wilson, Tim (2009). University Research Exposes Potential Vulnerabilities in Cloud Computing. Retrieved June 3, 2010 from http://www.darkreading.com/securityservices/security/management/showArticle.jhtml?articleID=219700098.