Cloud Forensics

Keyun Ruan, Joe Carthy, Tahar Kechadi, Mark Crosbie

Chapter 3

CLOUD FORENSICS

Keyun Ruan, Joe Carthy, Tahar Kechadi and Mark Crosbie

Abstract Cloud computing may well become one of the most transformative technologies in the history of computing. Cloud service providers and customers have yet to establish adequate forensic capabilities that could support investigations of criminal activities in the cloud. This paper discusses the emerging area of cloud forensics, and highlights its challenges and opportunities.

Keywords: Cloud computing, cloud forensics

1. Introduction

Cloud computing has the potential to become one of the most transformative computing technologies, following in the footsteps of mainframes, minicomputers, personal computers, the World Wide Web and smartphones [15]. Cloud computing is radically changing how information technology services are created, delivered, accessed and managed. Spending on cloud services is growing at five times the rate of traditional on-premises information technology (IT) [9]. Cloud computing services are forecast to generate approximately one-third of the net new growth within the IT industry. Gartner [8] predicts that the worldwide cloud services market will reach $150.1 billion in 2013.

Just as the cloud services market is growing, the size of the average digital forensic case is growing at the rate of 35% per year – from 83 GB in 2003 to 277 GB in 2007 [7]. The result is that the amount of forensic data that must be processed is outgrowing the ability to process it in a timely manner [16].

The rise of cloud computing not only exacerbates the problem of scale for digital forensic activities, but also creates a brand new front for cyber crime investigations with the associated challenges. Digital forensic practitioners must extend their expertise and tools to cloud computing

environments. Moreover, cloud-based entities – cloud service providers (CSPs) and cloud customers – must establish forensic capabilities that can help reduce cloud security risks. This paper discusses the emerging area of cloud forensics, and highlights its challenges and opportunities.

2. Cloud Forensics

Cloud forensics is a cross discipline of cloud computing and digital forensics. Cloud computing is a shared collection of configurable networked resources (e.g., networks, servers, storage, applications and services) that can be reconfigured quickly with minimal effort [12]. Digital forensics is the application of computer science principles to recover electronic evidence for presentation in a court of law [10].

Cloud forensics is a subset of network forensics. Network forensics deals with forensic investigations of networks. Cloud computing is based on broad network access. Therefore, cloud forensics follows the main phases of network forensics with techniques tailored to cloud computing environments.

Cloud computing is an evolving paradigm with complex aspects. Its essential characteristics have dramatically reduced IT costs, contributing to the rapid adoption of cloud computing by business and government [5]. To ensure service availability and cost-effectiveness, CSPs maintain data centers around the world. Data stored in one data center is replicated at multiple locations to ensure abundance and reduce the risk of failure. Also, the segregation of duties between CSPs and customers with regard to forensic responsibilities differ according to the service models being used. Likewise, the interactions between multiple tenants that share the same cloud resources differ according to the deployment model being employed.

Multiple jurisdictions and multi-tenancy are the default settings for cloud forensics, which create additional legal challenges. Sophisticated interactions between CSPs and customers, resource sharing by multiple tenants and collaboration between international law enforcement agencies are required in most cloud forensic investigations. 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.

2.1 Technical Dimension

The technical dimension encompasses the procedures and tools that are needed to perform the forensic process in a cloud computing environ-

ment. These include data collection, live forensics, evidence segregation, virtualized environments and proactive measures.

Data collection is the process of identifying, labeling, recording and acquiring forensic data. The forensic data includes client-side artifacts that reside on client premises and provider-side artifacts that are located in the provider infrastructure. The procedures and tools used to collect forensic data differ based on the specific model of data responsibility that is in place. The collection process should preserve the integrity of data with clearly defined segregation of duties between the client and provider. It should not breach laws or regulations in the jurisdictions where data is collected, or compromise the confidentiality of other tenants that share the resources. For example, in public clouds, provider-side artifacts may require the segregation of tenants, whereas there may be no such need in private clouds.

Rapid elasticity is one of the essential characteristics of cloud computing. Cloud resources can be provisioned and deprovisioned on demand. As a result, cloud forensic tools also need to be elastic. In most cases, these include large-scale static and live forensic tools for data acquisition (including volatile data collection), data recovery, evidence examination and evidence analysis.

Another essential characteristic of cloud computing is resource pooling. Multi-tenant environments reduce IT costs through resource sharing. However, the process of segregating evidence in the cloud requires compartmentalization [4]. Thus, procedures and tools must be developed to segregate forensic data between multiple tenants in various cloud deployment models and service models.

Virtualization is a key technology that is used to implement cloud services. However, hypervisor investigation procedures are practically non-existent. Another challenge is posed by the loss of data control [4]. Procedures and tools must be developed to physically locate forensic data with specific timestamps while taking into consideration the jurisdictional issues.

Proactive measures can significantly facilitate cloud forensic investigations. Examples include preserving regular snapshots of storage, continually tracking authentication and access control, and performing object-level auditing of all accesses.

2.2 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

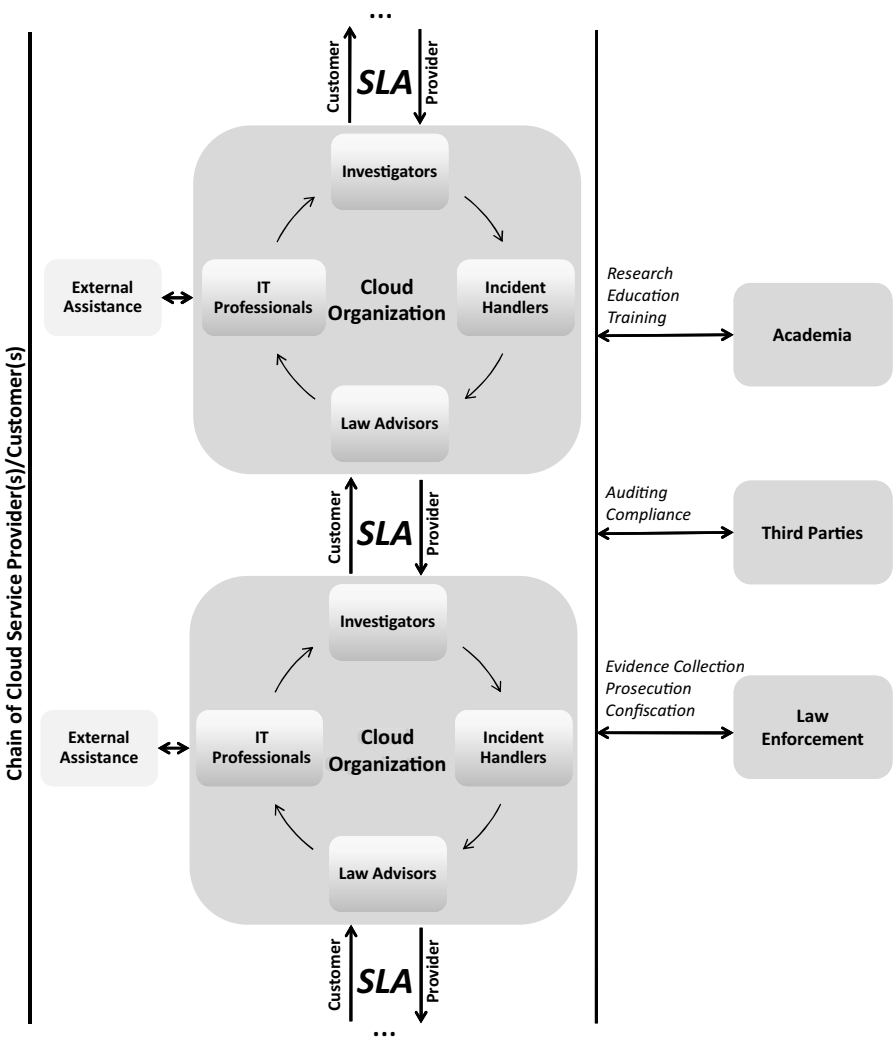

Figure 1. Entities involved in a cloud forensic investigation.

other parties. Figure 1 shows the various entities that may be involved in a cloud forensic investigation.

CSPs and most cloud applications often have dependencies on other CSPs. The dependencies in a chain of CSPs/customers can be highly dynamic. In such a situation, the cloud forensic investigation may depend on investigations of each link in the chain. Any interruption or corruption in the chain or a lack of coordination of responsibilities between all the involved parties can lead to serious problems.

Organizational policies and service level agreements (SLAs) facilitate communication and collaboration in forensic activities. In addition to law enforcement, the chain of CSPs must communicate and collaborate with third parties and academia. Third parties can assist with auditing and compliance while academia can provide technical expertise that could enhance the efficiency and effectiveness of investigations.

To establish a cloud forensic capability, each cloud entity must provide internal staffing, provider-customer collaboration and external assistance that fulfill the following roles:

- **Investigators:** Investigators are responsible for examining allegations of misconduct and working with external law enforcement agencies as needed. They must have sufficient expertise to perform investigations of their own assets as well as interact with other parties in forensic investigations.

- **IT Professionals:** IT professionals include system, network and security administrators, ethical hackers, cloud security architects, and technical and support staff. They provide expert knowledge in support of investigations, assist investigators in accessing crime scenes, and may perform data collection on behalf of investigators.

- **Incident Handlers:** Incident handlers respond to security incidents such as unauthorized data access, accidental data leakage and loss, breach of tenant confidentiality, inappropriate system use, malicious code infections, insider attacks and denial of service attacks. All cloud entities should have written plans that categorize security incidents for the different levels of the cloud and identify incident handlers with the appropriate expertise.

- **Legal Advisors:** Legal advisors are familiar with multi-jurisdictional and multi-tenancy issues in the cloud. They ensure that forensic activities do not violate laws and regulations, and maintain the confidentiality of other tenants that share the resources. SLAs must clarify the procedures that are followed in forensic investigations. Internal legal advisors should be involved in drafting the SLAs to cover all the jurisdictions in which a CSP operates. Internal legal advisors are also responsible for communicating and collaborating with external law enforcement agencies during the course of forensic investigations.

- **External Assistance:** It is prudent for a cloud entity to rely on internal staff as well as external parties to perform forensic tasks. It is important for a cloud entity to determine, in advance, the actions

that should be performed by external parties, and ensure that the relevant policies, guidelines and agreements are transparent to customers and law enforcement agencies.

2.3 Legal Dimension

Traditional digital forensic professionals identify multi-jurisdictional and multi-tenancy challenges as the top legal concerns [3, 11]. Performing forensics in the cloud exacerbates these challenges.

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.

SLAs define the terms of use between a CSP and its customers. The following terms regarding forensic investigations should be included in SLAs: (i) the services provided, techniques supported and access granted by the CSP to customers during forensic investigations; (ii) trust boundaries, roles and responsibilities between the CSP and customers regarding forensic investigations; and (iii) the process for conducting investigations in multi-jurisdictional environments without violating the applicable laws, regulations, and customer confidentiality and privacy policies.

3. Challenges

This section discusses eight challenges to establishing a cloud forensic capability that cover the technical, organizational and legal dimensions.

3.1 Forensic Data Collection

In every combination of cloud service model and deployment model, the cloud customer faces the challenge of decreased access to forensic data. Access to forensic data varies considerably based on the cloud model that is implemented [1]. Infrastructure as a service (IaaS) customers enjoy relatively unfettered access to the data required for forensic investigations. On the other hand, software as a service (SaaS) customers may have little or no access to such data.

Decreased access to forensic data means that cloud customers generally have little or no control – or even knowledge – of the physical locations of their data. In fact, they may only be able to specify location at a high level of abstraction, typically as an object or container. CSPs intentionally hide data locations from customers to facilitate data movement and replication.

Additionally, SLAs generally neglect to mention the terms of use that would facilitate forensic readiness in the cloud. Many CSPs do not provide services or interfaces for customers to gather forensic data. For example, SaaS providers may not provide their customers with the IP logs of client accesses, and IaaS providers may not provide recent virtual machine and disk images. Indeed, cloud customers have very limited access to log files and metadata at all levels, as well as a limited ability to audit and conduct real-time monitoring on their own.

3.2 Static, Elastic and Live Forensics

The proliferation of endpoints, especially mobile endpoints, is a challenge for data discovery and evidence collection. Because of the large number of resources connected to the cloud, the impact of a crime and the workload of an investigation can be massive.

Constructing the timeline of an event requires accurate time synchronization. Time synchronization is complicated because the data of interest resides on multiple physical machines in multiple geographical regions, or the data may be in flow between the cloud infrastructure and remote endpoint clients.

The use of disparate log formats is already a challenge in traditional network forensics. The challenge is exacerbated in the cloud due to the sheer volume of data logs and the prevalence of proprietary log formats.

Deleted data is an important source of evidence in traditional digital forensics. In the cloud, the customer who created a data volume often maintains the right to alter and delete the data [1]. When the customer deletes a data item, the removal of the mapping in the domain begins immediately and is typically completed in seconds. Remote access to the deleted data is not possible without the mapping. Also, the storage space occupied by the deleted data is made available for write operations and is overwritten by new data. Nevertheless, some deleted data may still be present in a memory snapshot [1]. The challenges are to recover the deleted data, identify the ownership of the deleted data, and use the deleted data for event reconstruction in the cloud.

3.3 Evidence Segregation

In the cloud, different instances running on a single physical machine are isolated from each other via virtualization. The neighbors of an instance have no more access to the instance than any other host on the Internet. Neighbors behave as if they are on separate hosts. Customer instances have no access to raw disk devices, instead they access virtualized disks. At the physical level, system audit logs of shared resources

collect data from multiple tenants. Technologies used for provisioning and deprovisioning resources are being improved [4]. It is a challenge for CSPs and law enforcement agencies to segregate resources during investigations without breaching the confidentiality of other tenants that share the infrastructure.

Another issue is that the easy-to-use feature of cloud models contributes to a weak registration system. This facilitates anonymity, which makes it easier for criminals to conceal their identities and harder for investigators to identify and trace suspects.

CSPs use encryption to separate data hosting and data use; when this feature is not available, customers are encouraged to encrypt their sensitive data before uploading it to the cloud [1]. The chain of separation must be standardized in SLAs and access to cryptographic keys should formalized in agreements between CSPs, customers and law enforcement agencies.

3.4 Virtualized Environments

Cloud computing provides data and computational redundancy by replicating and distributing resources. Most CSPs implement redundancy using virtualization. Instances of servers run as virtual machines, monitored and provisioned by a hypervisor. A hypervisor is analogous to a kernel in a traditional operating system. Hypervisors are prime targets for attack, but there is an alarming lack of policies, procedures and techniques for forensic investigations of hypervisors.

Data mirroring over multiple machines in different jurisdictions and the lack of transparent, real-time information about data locations introduces difficulties in forensic investigations. Investigators may unknowingly violate laws and regulations because they do not have clear information about data storage jurisdictions [6]. Additionally, a CSP cannot provide a precise physical location for a piece of data across all the geographical regions of the cloud. Finally, the distributed nature of cloud computing requires strong international cooperation – especially when the cloud resources to be confiscated are located around the world.

3.5 Internal Staffing

Most cloud forensic investigations are conducted by traditional digital forensic experts using conventional network forensic procedures and tools. A major challenge is posed by the paucity of technical and legal expertise with respect to cloud forensics. This is exacerbated by the fact that forensic research and laws and regulations are far behind the rapidly-evolving cloud technologies [2]. Cloud entities must ensure

that they have sufficient trained staff to address the technical and legal challenges involved in cloud forensic investigations.

3.6 External Dependency Chains

As mentioned in the organizational dimension of cloud forensics, CSPs and most cloud applications often have dependencies on other CSPs. For example, a CSP that provides an email application (SaaS) may depend on a third-party provider to host log files (i.e., platform as a service (PaaS)), who in turn may rely on a partner who provides the infrastructure to store log files (IaaS). A cloud forensic investigation thus requires investigations of each individual link in the dependency chain. Correlation of the activities across CSPs is a major challenge. An interruption or even a lack of coordination between the parties involved can lead to problems. Procedures, policies and agreements related to cross-provider forensic investigations are virtually nonexistent.

3.7 Service Level Agreements

Current SLAs omit important terms regarding forensic investigations. This is due to low customer awareness, limited CSP transparency and the lack of international regulation. Most cloud customers are unaware of the issues that may arise in a cloud forensic investigation and their significance. CSPs are generally unwilling to increase transparency because of inadequate expertise related to technical and legal issues, and the absence of regulations that mandate increased transparency.

3.8 Multiple Jurisdictions and Tenancy

Clearly, the presence of multiple jurisdictions and multi-tenancy in cloud computing pose significant challenges to forensic investigations. Each jurisdiction imposes different requirements regarding data access and retrieval, evidence recovery without breaching tenant rights, evidence admissibility and chain of custody. The absence of a worldwide regulatory body or even a federation of national bodies significantly impacts the effectiveness of cloud forensic investigations.

4. Opportunities

Despite the many challenges facing cloud forensics, there are several opportunities that can be leveraged to advance forensic investigations.

4.1 Cost Effectiveness

Security and forensic services can be less expensive when implemented on a large scale. Cloud computing is attractive to small and medium enterprises because it reduces IT costs. Enterprises that cannot afford dedicated internal or external forensic capabilities may be able to take advantage of low-cost cloud forensic services.

4.2 Data Abundance

Amazon S3 and Amazon Simple DB ensure object durability by storing objects multiple times in multiple availability zones on the initial write. Subsequently, they further replicate the objects to reduce the risk of failure due to device unavailability and bit rot [1]. This replication also reduces the likelihood that vital evidence is completely deleted.

4.3 Overall Robustness

Some technologies help improve the overall robustness of cloud forensics. For example, Amazon S3 automatically generates an MD5 hash when an object is stored [1].

IaaS offerings support on-demand cloning of virtual machines. As a result, in the event of a suspected security breach, a customer can take an image of a live virtual machine for offline forensic analysis, which results in less downtime. Also, using multiple image clones can speed up analysis by parallelizing investigation tasks. This enhances the analysis of security incidents and increases the probability of tracking attackers and patching weaknesses. Amazon S3, for example, allows customers to use versioning to preserve, retrieve and restore every version of every object stored in an S3 bucket [1]. An Amazon S3 bucket also logs access to the bucket and objects within it. The access log contains details about each access request including request type, requested resource, requester's IP address, and the time and date of the request. This provides a wealth of useful information for investigating anomalies and incidents.

4.4 Scalability and Flexibility

Cloud computing facilitates the scalable and flexible use of resources, which also applies to forensic services. For example, cloud computing provides (essentially) unlimited pay-per-use storage, allowing comprehensive logging without compromising performance. It also increases the efficiency of indexing, searching and querying logs. Cloud instances can be scaled as needed based on the logging load. Likewise, forensic activities can leverage the scalability and flexibility of cloud computing.

4.5 Policies and Standards

Forensic policies and standards invariably play catch-up to technological advancements, resulting in brittle, *ad hoc* solutions [13]. However, cloud computing is still in the early stage and a unique opportunity exists to lay a foundation for cloud forensic policies and standards that will evolve hand-in-hand with the technology.

4.6 Forensics as a Service

The concept of security as a service is emerging in cloud computing. Research has demonstrated the advantages of cloud-based anti-virus software [14] and cloud platforms for forensic computing [16]. Security vendors are changing their delivery methods to include cloud services, and some companies are providing security as a cloud service. Likewise, forensics as a cloud service could leverage the massive computing power of the cloud to support cyber crime investigations at all levels.

5. Conclusions

Cloud computing is pushing the frontiers of digital forensics. The cloud exacerbates many technological, organizational and legal challenges. Several of these challenges, such as data replication, location transparency and multi-tenancy, are unique to cloud forensics. Nevertheless, cloud forensics brings unique opportunities that can significantly advance the efficacy and speed of forensic investigations.

References

[1] Amazon, AWS Security Center, Seattle, Washington (aws.amazon.com/security).

[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] R. Broadhurst, Developments in the global law enforcement of cyber crime, *Policing: International Journal of Police Strategies and Management*, vol. 29(2), pp. 408–433, 2006.

[4] Cloud Security Alliance, Security Guidance for Critical Areas of Focus in Cloud Computing V2.1, San Francisco, California (www.cloudsecurityalliance.org/csaguide.pdf), 2009.

[5] EurActiv, Cloud computing: A legal maze for Europe, Brussels, Belgium (www.euractiv.com/en/innovation/cloud-computing-legal-maze-europe-linksdossier-502073), 2011.

[6] European Network and Information Security Agency, Cloud Computing: Benefits, Risks and Recommendations for Information Security, Heraklion, Crete, Greece (www.enisa.europa.eu/act/rm/files/deliverables/cloud-computing-risk-assessment), 2009

[7] Federal Bureau of Investigation, Regional Computer Forensics Laboratory, Annual Report for Fiscal Year 2007, Washington, DC (www.rcfl.gov/downloads/documents/RCFL_Nat_Annual07.pdf), 2007.

[8] Gartner, Gartner says worldwide cloud services revenue will grow 21.3 percent in 2009, Stamford, Connecticut (www.gartner.com/it/page.jsp?id=920712), March 26, 2009.

[9] F. Gens, IT cloud services forecast – 2008 to 2012: A key driver of new growth (blogs.idc.com/ie/?p=224), October 8, 2008.

[10] K. Kent, S. Chevalier, T. Grance and H. Dang, Guide to Integrating Forensic Techniques into Incident Response, Special Publication 800-86, National Institute of Standards and Technology, Gaithersburg, Maryland, 2006.

[11] 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.

[12] P. Mell and T. Grance, The NIST Definition of Cloud Computing (Draft), Special Publication 800-145 (Draft), National Institute of Standards and Technology, Gaithersburg, Maryland, 2011.

[13] M. Meyers and M. Rogers, Computer forensics: The need for standardization and certification, *International Journal of Digital Evidence*, vol. 3(2), 2004.

[14] J. Oberheide, E. Cooke and F. Jahanian, CloudAV: N-version antivirus in the network cloud, *Proceedings of the Seventeenth USENIX Security Conference*, pp. 91–106, 2008.

[15] R. Perry, E. Hatcher, R. Mahowald and S. Hendrick, Force.com cloud platform drives huge time to market and cost savings, IDC White Paper, International Data Corporation, Framingham, Massachusetts (thecloud.appirio.com/rs/appirio/images/IDC_Force.com_ROI_Study.pdf), 2009.

[16] V. Roussev, L. Wang, G. Richard and L. Marziale, A cloud computing platform for large-scale forensic computing, in *Advances in Digital Forensics V*, G. Peterson and S. Shenoi (Eds.), Springer, Heidelberg, Germany, pp. 201–214, 2009.