cloud security alliance®

Incident Management and Forensics Working Group

Mapping the Forensic Standard ISO/IEC 27037 to Cloud Computing

June 2013

Contents

Acknowledgments

Working Group Co-Chairs

Dominik Birk
Michael Panico

Contributors

Aaron Alva, University of Washington
Bernd Jaeger, Colt Technology
Dominik Birk, Zurich Insurance Company
Josiah Dykstra, University of Maryland Baltimore County
Keyun Ruan, University College Dublin
Michael Panico, Stroz Friedberg
Richard Austin, Hewlett-Packard

CSA Global Staff

Alex Ginsburg, Copyeditor
Brianna Lichtenauer, Copyeditor
Luciano JR Santos, Global Research Director
Evan Scoboria, Webmaster
Kendall Scoboria, Graphic Designer
John Yeoh, Research Analyst

Introduction

Cloud computing has become a dominant paradigm in information technology, but with its many promising features and cost advantages for both enterprises and governments come unique security challenges.

In addition to the security challenges inherent in these multi-tenant, highly virtualized environments, processes for conducting forensic investigations and electronic discovery (eDiscovery) are immature.

The purpose of this document is to survey the issues related to forensic investigation in cloud environments, to describe, in detail, the international standards for cloud forensics, and to summarize the current integration of cloud forensic requirements into service level agreements (SLAs).

1.0 Forensic Science and Traditional Digital Forensics

According to the American Academy of Forensic Sciences (AAFS):

"Forensic Science is the application of scientific principles and technological practices to the purposes of justice in the study and resolution of criminal, civil, and regulation issues."

– AAFS Board of Directors, 1993

The inaugural Digital Forensic Research Workshop (DFRWS) provided another widely adopted definition:

"The use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation and preservation of digital evidence derived from digital sources for the purpose of facilitating or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations."

– DFRWS, 2001

Under real world circumstances, the practice of digital forensics is fundamentally related to the legal system and its rules of evidence as established for a particular jurisdiction [1]. These legal systems and their rules of evidence provide the context for the practice of digital forensics and place restrictions on how the process is carried out in a particular location. For this reason, it is critical to understand local legal systems when considering the practice of digital forensics in a specific investigation.

1.1 The Notion of Cloud Forensics

The history of information technology has revealed that data stored on systems and within applications is never fully immune to illicit access or compromise. The risks to corporate data will not diminish in multi-tenant, highly

virtualized cloud computing environments. In some cases, cloud environments will exacerbate security challenges for cloud consumers due to the distributed, virtualized nature of the cloud [2, 3, 15]. Furthermore, the practice of digital forensics is also challenged by the migration to more complex, highly-virtualized cloud computing environments [4, 5, 6, 7].

Hence, in an increasingly cloud-oriented society, the ability to identify, obtain, preserve, and analyze potential digital evidence is a critical business capability. Whether responding to a security incident, data breach, or in support of litigation, the ill-prepared organization will find itself at a severe (and potentially costly) disadvantage.

The CSA Trusted Cloud Reference Architecture [8] emphasizes the criticality of forensic readiness by including it in both the *"Business Operation Support Services (BOSS)"* and *"Security and Risk Management"* domains. Forensic readiness also plays an important role in the security incident response processes specified in the *"Information Technology Operation and Support (ITOS)"* domain.

2.0 Forensic Requirements for CSPs

Customers and law enforcement agencies will increasingly ask cloud service providers (CSPs) for forensic support. The CSP's forensic support obligations depend on the service model [9] that is offered by the CSP and used by the customer. Different service models provide different capabilities for the customer in terms of digital forensics [5].

1. **SaaS Environments**
   From a customer perspective, the software as a service (SaaS) model is one in which the capabilities of the customer are most restricted. The customer possesses no control over the underlying operating infrastructure such as the network, servers, operating systems or source code of the application in use, thus limiting customers' forensic capabilities. In most cases, SaaS environments demand that the forensic examiner rely on high-level application logs provided by the application and therefore on the CSP's support for forensic functionality. As such, required forensic functionality must be specified in service level objectives (SLOs) incorporated into the service level agreement between the customer and the CSP.

   SLOs may include requirements for notification, identification, preservation, and access to potential evidence sources.

   SLOs may specify potential evidence sources under CSP control, including:
   a. Webserver logs
   b. Application server logs
   c. Database logs
   d. Guest operating system logs
   e. Host access logs

f. Virtualization platform logs and SaaS portal logs
g. Network captures
h. Billing records

2. **PaaS Environments**

One of the main advantages of the platform as a service (PaaS) model is that the customer controls the developed software application and the source code of the application does not have to leave the local development environment. Given these circumstances, the customer maintains the power to implement forensic capabilities within the application. Automatic logging functionalities [10] could be implemented that push logs to external logging servers implementing the write-once, read-many (WORM [11]) principle. However, the PaaS model still necessitates coordination between the customer and the CSP. Although the customer controls the functionality of the application, the actual operation of the application will occur within the CSP's infrastructure. As a result, the customer must clearly identify the responsibilities of the CSP when the need for a forensic investigation arises. These responsibilities should take the form of SLOs documented in the SLA between the customer and the CSP.

SLOs may include requirements for notification, identification, preservation, and access to potential evidence sources.

SLOs may specify potential evidence sources under CSP control, including:

a. Webserver logs
b. Application server logs (see SaaS)
c. Guest operating system logs
d. Host access logs
e. Virtualization platform logs
f. Network captures
g. Billing records
h. Management portal logs

3. **IaaS Environments**

Compared with SaaS and Paas, the infrastructure as a service (IaaS) deployment model offers a greater range of potential evidence sources under control of the customer. However, some (perhaps essential) data might only exist in the CSP infrastructure. This requires that the customer clearly document the responsibilities of the CSP when the need for a forensic investigation arises. These responsibilities should take the form of SLOs memorialized in the contract between the customer and the CSP.

SLOs may specify potential evidence sources under CSP control, including:

a. Cloud or network provider perimeter network logs
b. Logs from DNS servers
c. Virtual machine monitor (VMM) logs
d. Host operating system logs
e. API logs

f. Management portal logs
g. Packet captures
h. Billing records

In addition to forensic requirements that vary with the service model, forensic requirements might also depend on the specific capabilities of the customer and CSP. For example:

1. A smaller company without an IT department has a SaaS application offered and hosted by the CSP. The company is informed by an external party that their website is leaking customer data. The CSP provides a full "forensic service" including incident response (IR), reporting and re-building the system in a secure way. Without the help of the CSP, the customer does not have the access to data necessary to perform a comprehensive forensic investigation.

2. An experienced customer with a large internal forensic department has detected "strange" behavior of a VM hosted by the CSP. In this case the CSP may have to provide some information, for example VM snapshots and some firewall/router logs. The customer may need to pull and check utilization statistics, weblogs and real-time guest OS kernel events and file system checksums or hashes.

2.1 Importance of the Sla

The need for SLOs embedded in the SLA is essential for specifying CSP responsibilities associated with forensic investigations. SLAs are a legally binding agreement between a cloud consumer and a cloud provider. SLOs determine the way that CSPs address forensic investigations, including the process for identification and preservation of potential evidence and access to data. For example, the importance of SLOs in determining the accessibility of potential evidence is illustrated in the following figure where the proportion of relevant evidence available as "accessible sources of evidence" is determined by the terms of the SLA.

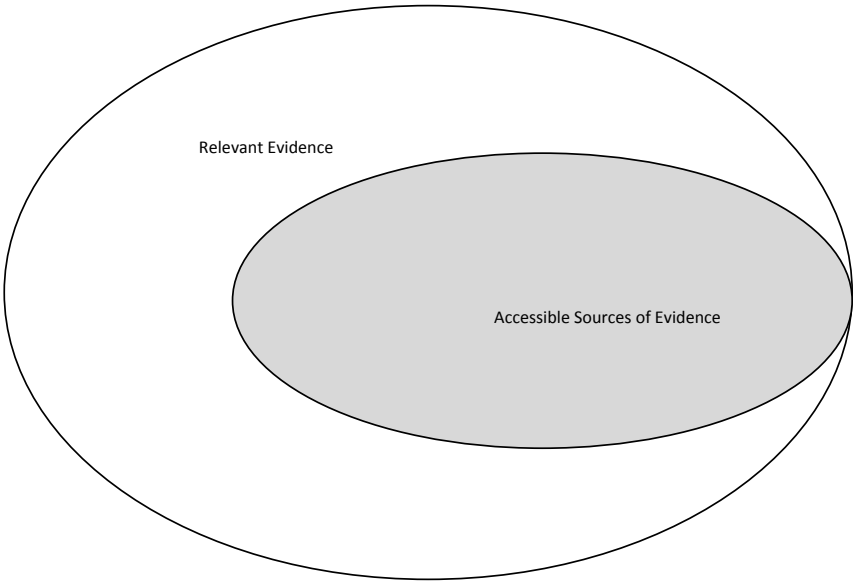

Figure 1: Availability of Potential Digital Evidence

Within a cloud provider infrastructure, there may be many sources of relevant evidence. However, the customer may only have access to the restricted evidence set (the shaded area) provided by the CSP. In a cloud environment, it may be difficult to identify all instances of relevant data (and therefore what should be memorialized as "accessible"). For example, virtual instances used by a particular customer may migrate transparently between various physical instances with little recordkeeping. What records do exist may be very transitory and only available for a short period of

time. The amount of accessible evidence may also be severely constrained by cost, technology (e.g., available storage space), multi-tenancy, privacy implications and other factors relevant to a particular CSP's infrastructure.

For these reasons, it is critical that the customer understand the sources of potential digital evidence that will be available from the CSP, limitations on volumes of data, and retention periods. To avoid misunderstandings and potential litigation, these understandings should be documented in SLOs within the SLA.

In [6], the authors identify a list of key terms that can be included in the SLA in order to support forensic investigations. These key terms are organized under four categories: technical key terms, organizational key terms, legal key terms and auditing key terms.

2.2 General High-Level Requirements

The latest release of Cloud Control Matrix (CCM) [12] has modified and added the following security principles (CO-04, DG-05, IS-24, SA-12) that cover forensic investigations.

CO-04 Compliance – Contract/Authority Maintenance: Points of contact for applicable regulatory authorities, national and local law enforcement and other legal jurisdictional authorities shall be maintained and regularly updated as per the business need (i.e., change in impacted-scope and/or a change in any compliance obligation) to ensure direct compliance liaisons have been established and to be prepared for a forensic investigation requiring rapid engagement with law enforcement.

DG-05 Data Governance – Secure Disposal: Policies and procedures shall be established with supporting business processes and technical measures implemented for the secure disposal and complete removal of data from all storage media, ensuring data is not recoverable by any computer forensic means.

IS-24 Information Security – Incident Response Legal Preparation: In the event a follow-up action concerning a person or organization after an information security incident requires legal action, proper forensic procedures including chain of custody shall be required for preservation and presentation of evidence to support potential legal action subject to the relevant jurisdictions. Upon notification, impacted customers (tenants) and/or other external business relationships of a security breach shall be given the opportunity to participate as is legally permissible in the forensic investigation.

SA-12 Security Architecture – Audit Logging/Intrusion Detection: Higher levels of assurance are required for protection, retention, and lifecycle management of audit logs, adhering to applicable legal, statutory or regulatory compliance obligations and providing unique user access accountability to detect potentially suspicious network behaviors and/or file integrity anomalies through to forensic investigative capabilities in the event of a security breach.

3.0 ISO 27037

ISO 27037 [14] is the first of a developing family of international standards that seek to create a common baseline for the practice of digital forensics. It is not intended to replace local laws or usurp local and national governments' authority to regulate the practice of digital forensics. Rather, its intent is to facilitate the usability of evidence obtained in one jurisdiction by a legal process operating in another jurisdiction.

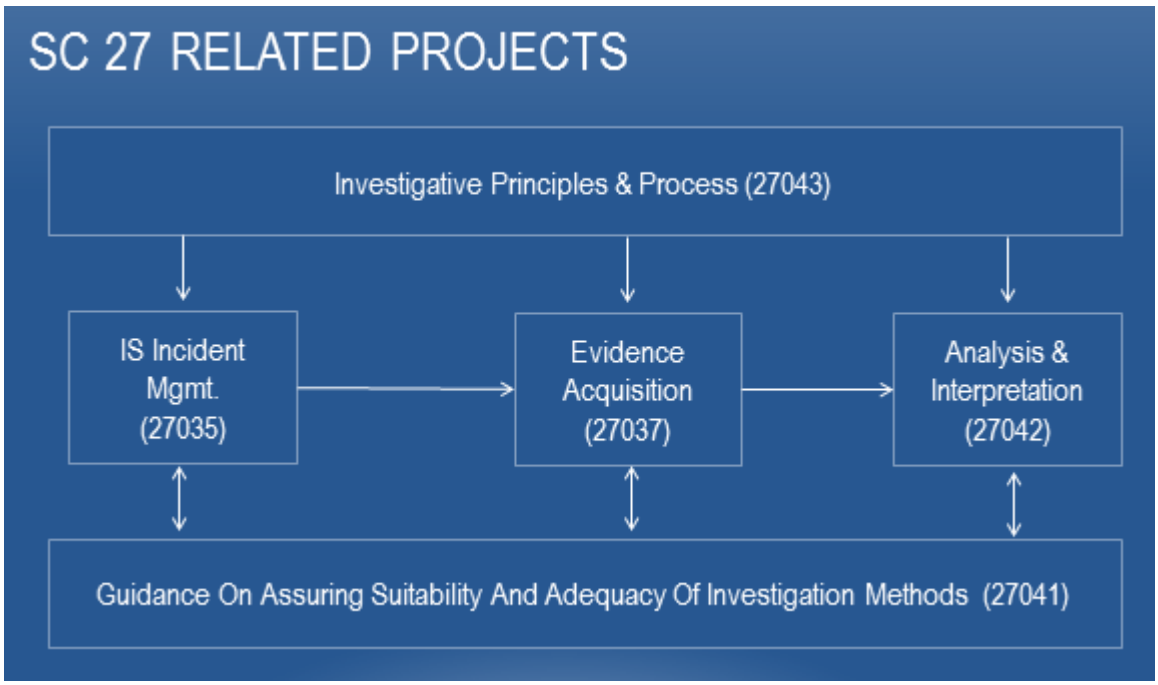

Figure 1: Developing International Standards[1]

As its title suggests, ISO 27037 only addresses the initial steps of the forensics process: identifying, obtaining and preserving potential digital evidence[2]. Other steps in the forensics process are the subject of additional standards currently under development.

[1] Diagram courtesy of Mr. Eric Hibbard of HDS and is used with permission

[2] The term "potential digital evidence" is used to recognize that evidence must be accepted by a court or other judicial.

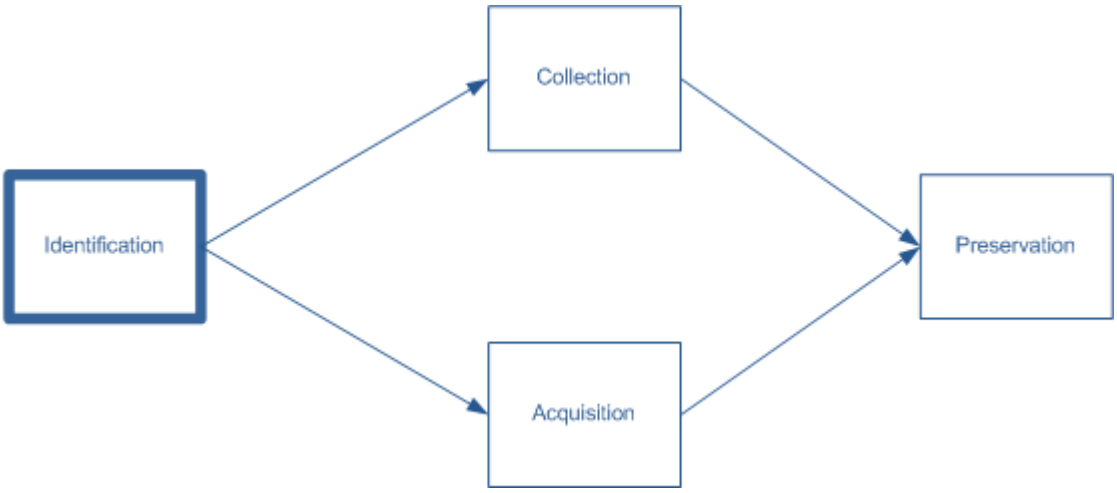

Figure 2: Evidence-Handling Processes According to ISO 27037

3.1 Identification

The forensics process begins with the identification of items that may be or may contain potential digital evidence. Formally, identification is the *"process involving the search for, recognition and documentation of potential digital evidence"* [14].

Although the identification of potential digital evidence sounds simple in principle, there are subtle complexities. For example, digital evidence has both a physical and virtual representation. Consider a hard drive containing potential digital evidence. The physical location of the evidence is the hard drive, but the evidence itself is the data contained within the drive. Furthermore, it also may not be at all obvious where potential digital evidence is housed. A server may have very few directly attached disks and have a significant part of its storage within a SAN or NAS. As will be discussed later, these aspects of the cloud environment compound the difficulties in identifying relevant evidence.

3.2 Collection and Acquisition

After potential digital evidence is identified, it must either be collected or acquired:

- Collection – *"Process of gathering items that contain potential digital evidence."* [14]
- Acquisition – *"Process of creating a copy of data within a defined set."* [14]

Collection is roughly equivalent to the standard law enforcement practice of seizing items containing potential digital evidence under authority of a legal order (i.e., search warrant) and removing them to a forensics lab or other facility for processing and analysis. Acquisition is more common in the private sector due to the need to minimize business impact of an ongoing investigation. Similar concerns with reducing the impact on other applications and customers will make acquisition the more likely process in the cloud environment as well.

It should be noted that the copy created during acquisition can range from the forensic image of a hard drive to a copy of the contents of a server's memory to the logical contents of an individual user's email box depending

on the purpose and scope of the investigation. In all cases, the requirements for the copy are very similar: it must be made using a well-understood, defensible, well-documented process. Furthermore, the process must include integrity measures to ensure that the copy has not been modified since acquisition. The wide variety of potential digital evidence to be copied, and the requirements on the copying process, make acquisition a more complex and challenging process than collection.

3.3 Preservation

Once potential digital evidence has been collected or acquired, it must be preserved. ISO 27037 defines preservation as the *"process to maintain and safeguard the integrity and/or original condition of the potential digital evidence"* [14]. The preservation of potential digital evidence is a complex and important process. Evidence preservation helps assure admissibility in a court of law. However, digital evidence is notoriously fragile, and is easily changed or destroyed. Given that the backlog in many forensic laboratories ranges from six months to a year (and that delays in the legal system might create further delays), potential digital evidence may spend a significant period of time in storage before it is analyzed or used in a legal proceeding. Storage requires strict access controls to protect the items from accidental or deliberate modification, as well as appropriate environment controls.

3.4 Differences between Cloud Forensics and Traditional Forensics

Although cloud forensics and traditional forensic practices share a common foundation, cloud forensics has unique barriers, challenges, and techniques. While many methods and techniques will transfer transparently into the cloud environment, there are unique practices as well. The following section will focus on the IaaS deployment model, but as noted previously, additional challenges will appear when moving to the PaaS or SaaS model (see *2.0 Forensic Requirements for CSPs*).

The first challenge in cloud forensics is the identification of potential digital evidence. With direct attached storage, it is easy to determine which storage device belongs to a given server. With the advent of storage networking and virtualization, mapping storage devices has become much more complex and this complexity increases in the cloud environment. For example, in the CSA Trusted Cloud Reference Architecture [8] under "Infrastructure Services" (**Error! Reference source not found.**), torage is highly virtualized. A group of

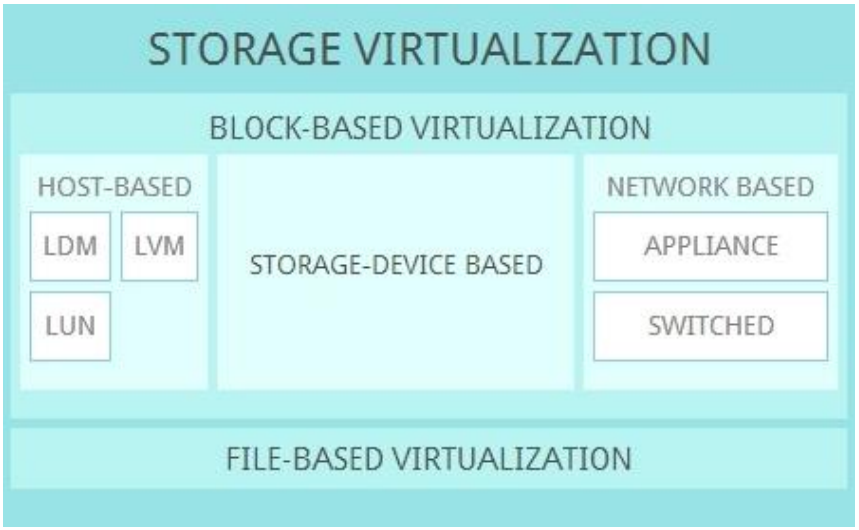

Figure 4: Storage Virtualization in the Trusted Cloud Reference Architecture

physical disk devices may be virtualized as a set of logical units presented to a cloud user (or a server supporting a cloud user) with RAID level, cache settings, etc., to match the specific cost, reliability and performance profile required. These logical units may even be transparently moved from place to place based on global performance and availability issues (perhaps storage instance "A" needs to be taken down for preventive maintenance so its logical units would be migrated to instance "B"). The identification process would have to be cognizant of the mapping and frequent migration to assure that the correct logical units were acquired.

Since past instances of storage objects (e.g., previous versions of digital documents, deleted files, the remains of temporary objects in free space, etc.) are often included in the corpus of relevant potential evidence, it is possible that the previous instance of the logical units on "A" would be within the scope of the investigation.

4.0 Mapping ISO 27037 to the Cloud

Although ISO 27037 is a relatively new standard (issued in October 2012) and only addresses the initial stages of a digital investigation (identifying, collecting/acquiring, and preserving potential digital evidence), it represents an international public and private sector consensus of how potential digital evidence should be handled in the critical initial steps of an investigation. There are many complex challenges of digital forensics in a cloud environment and this section will map and reinterpret the ISO 27037 guidance for a cloud context.

4.1 General Requirements

4.1.1 Requirements for Identification, Collection, Acquisition and Preservation of Digital Evidence - ISO 27037

ISO 27037	CLOUD
5.3.2 Auditable It should be possible for an independent assessor or other authorized interested parties to evaluate the activities performed by a DEFR[3] and DES[4]. This requires appropriate documentation regarding actions taken, why and how.	While this high-level requirement itself remains the same for cloud environments, execution becomes more difficult as investigations will likely be conducted on dynamic, distributed, and complex systems that can neither be frozen nor easily

[3] Digital Evidence First Responder – "individual who is authorized, trained and qualified to act first at an incident scene in performing digital evidence collection and acquisition with the responsibility for handling that evidence" [14]
[4] Digital Evidence Specialist – "individual who can carry out the tasks of a DEFR and has specialized knowledge , skills and abilities to handle a wide range of technical issues" [14]

<table>
<tr><td></td><td>identified. Thus, the necessity for appropriate qualifications and documentation becomes even more important.</td></tr>
<tr><td>5.3.3 Repeatable

Repeatability is established when the same test results are produced under the following conditions:

✓ Using the same measurement procedure and method

✓ Using the same instruments and under the same conditions

✓ Can be repeated at any time after the original test</td><td rowspan="2">While standard procedures and methods seem to be achievable even in a SaaS environment, conducting tests “under the same conditions” and “at any time after the original test” becomes more challenging (but not always impossible) within a dynamic, distributed, and complex cloud environment.

For acquisition in current forensic practice regarding imaging memory, an active log file, or other dynamic process, the concept of “snapshot forensics” is used. The analogy is that no two successive snapshots of a running child will capture exactly the same image (since the child is moving) but the snapshot accurately captures the appearance of the child and her background at a moment in time. Assurance of reliability for the snapshot then becomes assurance of its provenance and that it has not been modified since acquisition. Documentation can assure the identity, place and time of the snapshot while traditional techniques such as cryptographic hashes and chain-of-custody processes can provide integrity assurances.

The snapshot process is repeatable as it can be demonstrated that the camera will take an “accurate” photograph each time the shutter release is pressed. Reproducibility can similarly be shown by using a camera of similar capability from a different manufacturer. Thus, the process can be repeatable and reproducible even though no two successive snapshots of the running child will ever produce exactly the same results.</td></tr>
<tr><td>5.3.4 Reproducible

Reproducibility is established when the same test results are produced under the following conditions:

✓ Using the same measurement method

✓ Using different instruments and under different conditions

✓ Can be reproduced at any time after the original test</td></tr>
<tr><td>5.3.5 Justifiability</td><td></td></tr>
</table>

The DEFR should be able to justify all actions and methods used.	No changes for cloud environments.

4.2 Digital Evidence Handling - ISO 27037

ISO 27037	CLOUD
5.4.1 General Potential digital evidence should be treated according to the following principles: ✓ Minimize handling ✓ Account for any changes and document actions taken ✓ Comply with local rules of evidence ✓ Do not take actions beyond your competence	No changes for cloud environments.

ISO 27037	CLOUD
5.4.2 Identification The search for, recognition and documentation of potential digital evidence should be undertaken according to the following principles: ✓ Prioritize the evidence collection based on volatility ✓ Minimize the damage to the potential digital evidence ✓ Identify hidden digital evidence	It is recommended that customers identify the additional data sources unique to the cloud service model. Specifically: ✓ SaaS - application level logs like authorization errors, accounting (who did what, when), performance issues, data volumes, … ✓ PaaS - application specific logs available ideally via an API, patch status, authentication errors, operating system

✓ Recognize that identification may be difficult (cloud)	exceptions and warnings, anti-malware software warnings, … ✓ IaaS - system level logs, Infrastructure: hypervisor events and logs, raw virtual machine files suspend files capturing unencrypted RAM snapshots , Intrusion detection and firewall events, network events and packet capture, storage logs, backups, …
5.4.3 Collection *"Collection is a process in the digital evidence handling process where devices that may contain digital evidence are removed from their original location to a laboratory or another controlled environment for later acquisition and analysis."*	Due to the multi-tenant nature of cloud infrastructures, acquisition should usually be preferred over collection to avoid impacts to parties not involved in the matter and the gathering of irrelevant information that must be excluded during analysis. However, the specifics of the legal mandate in a particular situation must be followed. It must be emphasized that collection of digital evidence can often only be performed by the CSP and not by the tenant.
5.4.4 Acquisition The process of creating a copy of an item of potential digital evidence.	Because of the virtual nature of the cloud infrastructure, items normally thought of as physical (hard drives, server memory, etc.) will be logical items (a virtual hard disk file, a file that contains the contents of server memory for a suspended virtual machine, etc.) and acquisition must focus on these logical items rather than the physical containers where they reside.
5.4.5 Preservation Preservation is the protection of the integrity of potential digital evidence. Potential digital evidence and digital devices must be safeguarded from tampering or spoliation.	No changes for cloud environments. However, the chain of custody must be preserved as well, which is challenging in multi-geographical and multi-jurisdictional environments.

4.3 Key Components of Identification, Collection, Acquisition and Preservation of Digital Evidence – ISO 27037

Chapter 6 of the ISO standard refers to rather non-technical information including the chain of custody, roles and responsibilities, competencies and briefings. As there is little or no change for cloud environments, it will not be mapped in this document.

5.0 Instances of Identification, Collection, Acquisition and Preservation - ISO 27037

5.1 Computers, Peripheral Devices and Digital Storage Media - ISO 27037

Parts of ISO 27037 referring to stand-alone computer systems will not be mapped, as they are not applicable for cloud computing with the exception of mobile devices.

ISO 27037	CLOUD
7.1.1 Identification **7.1.1.1 Physical incident scene search and documentation** This refers mainly to stand-alone systems and is therefore only partially applicable for cloud	As cloud environments consist of multiple distributed, networked systems which are used/consumed over a network, it is likely infeasible or even impossible for the DEFR to access the physical incident scene. However, access to the client side (mobile devices) might be possible.
7.1.1.2 Non-digital evidence collection Additional, non-digital information should be collected e.g. by interviewing individuals to obtain passwords.	No change for cloud environments.
7.1.1.3 Decision-making process for collection or acquisition	

A determination must be made to collect or acquire potential evidence.	Within a cloud service, potential digital evidence is likely fragmented and distributed across the underlying storage infrastructure. As a result, physically collecting evidence (e.g. hard disk drives) might be impossible. In cloud environments, acquisition is likely the appropriate way to obtain a digital evidence copy. However, mobile devices (e.g. mobile clients that access a cloud application) might be collected.
7.1.2 Collection	Not applicable for cloud environments (see 5.4.3 Collection).
7.1.3 Acquisition **7.1.3.1 Powered on digital devices** **7.1.3.1.1 Overview** Scenarios exist in which acquisition may need to be conducted when the digital devices are powered on.	Within a cloud environment, the physical computing, network and storage systems will most likely be powered on while particular virtual systems might be offline or only available as snapshots and backups
The DEFR should make an accurate digital evidence copy of the digital device's storage media.	The virtualization layer of cloud systems can aid this process in providing the capability of creating a snapshot of a live system in a "non-intrusive" way. This snapshot will also contain all data within the system's memory.
Acquisition of volatile live data is important.	Snapshots (see above) or even Virtual Machine Introspection [16] could be helpful here (e.g. in case of encrypted VM hard disks) Example: Today's regulatory and industry vertical compliance requirements for multi-tenant cloud environments have lead into the increased usage of encryption within cloud environments. The scope of an investigation may include VMs and data that utilize encryption mechanisms at different levels (within the VM, at the hypervisor, at the storage network layer, on the NAS or Storage device …).

	Depending on where encryption is applied, "plugging" into a hypervisor API may help to gain access to data that would be unavailable otherwise. In addition, an encryption key or password might be found in a snapshot or vMotion file.
Suspect systems' programs or tools should never be used. Only use verified external (statically linked) tools.	The ability in cloud environments to use the virtual hardware layer to freeze a live system without spoiling it helps to avoid the usage of acquisition tools within the suspect systems' operating system software, while still being able to obtain evidence (e.g.. a full RAM dump for examination).
Store volatile data on prepared/sanitized storage media in file container and conduct appropriate hashing.	No change. If the target system is a VM, the VM files could be used as a container.
Use validated imaging tools for non-volatile data.	No change for cloud environments.
7.1.3.1.1 Additional Activities Try to detect data encryption on volatile data	No change for cloud environments.
Use a reliable time source.	While there is no fundamental difference to the non-cloud environment in facing the challenge to identify a reliable time source, the virtualization layer itself might sometimes add a time drift and/or complexity.
It may be appropriate to associate the DEFR with the acquired potential digital evidence.	No change for cloud environments.
7.1.3.2 Powered off digital devices **7.1.3.2.2 Acquisition** Easier to conduct as no volatile data has to be acquired. Conduct a proper imaging.	In a cloud environment, "powered off" digital devices are physically invisible as they are not represented by a physical workstation, laptop, docking station, charging or network cable or server. They only exist as files and database entries on the virtualization platform. Also, these offline VMs files contain data

	that was considered to be “volatile” on a physical workstation, such as data within the VM’s RAM. While the imaging process is not difficult after the VMs have been identified, knowing that they exist might be more difficult compared to physical systems.
7.1.3.3 Mission-critical digital devices In some cases, digital devices cannot be powered off.	This applies to cloud environments in general since multi-tenant, distributed infrastructures cannot be powered off due to investigation requests from only one tenant. On the other hand it demonstrates how the introduction of virtualization technology used for cloud can be an opportunity. The ability to snapshot and move a VM to a “lab” or “collection” platform and power up a new instance to take its place allows greater agility in keeping operations intact while conducting an investigation.
7.1.3.4 Partial acquisition Partial acquisition may be performed when: ✓ System storage is too large ✓ A system is too critical ✓ Only selected data must to be acquired ✓ A search warrant limits scope	Within cloud environments, acquisition will likely always be a partial acquisition, as all of the reasons listed in the standard requirements will likely apply.
7.1.3.5 Digital storage media Various types of storage media will exist.	Within cloud environments, data will typically be stored on large storage arrays. Several copies and backups may exist. Data and files might be fragmented across several physical locations. This will

	make data difficult to re-assemble without the storage system itself.
The location should be documented and checked.	Identifying or visiting the physical location of a physical storage array device might be difficult if not impossible. Additionally, a document might exist as data fragments stored in multiple physical locations.
Collection or on-site acquisition.	As mentioned earlier, in most cases only acquisition will be applicable in a cloud environment.
Consider different data retention capabilities of different storage media.	No change for cloud environments.
7.1.4 Preservation Seal acquired data with verification function and sign it.	No change for cloud environments.

5.2 Networked Devices - ISO 27037

As the NIST definition of cloud computing [9] implies a model of providing dynamic computing resources over a network, cloud environments are a special use case of section 7.2 of the ISO 27037 standard, which covers a broad scope of networked devices, including technologies like Bluetooth devices and CCTV systems.

ISO 27037	CLOUD
7.2.1 Identification **7.2.1.1 Overview** In a networked environment, it is difficult to ascertain where potential digital evidence is stored.	This becomes even more difficult in cloud environments as the DEFR might face a globally distributed environment where physical access is impossible and devices might be located in different jurisdictions. **Example**: The suspect may have used a cloud storage SaaS application to store illegal content. This application, provided via a reseller channel, is using a third-party identity provider via a standard API. Content is stored on multiple, different object storage platforms, provided by some other third-party storage provider which uses a sub-provider itself. Access token to stored

	objects may resist on multiple management nodes. Finally, logs are collected at yet another central site. All of these instances may be implemented in different countries. **Recommendation**: Although the overall IT systems involved are distributed, it may well be possible to trace back or reverse lookup the overall topology and obtain essential information by contacting the system owner on the remote end. Therefore, the DEFR should try to obtain all information in order to understand the overall system architecture, topology and information flows. However, the cloud environment is much more dynamic than server infrastructure was in the past. So topology information (e.g. allocated IP address, storage file space, etc.) may change rapidly. In addition, as cloud environments are typically highly overprovisioned, unallocated disk space may be overwritten rapidly as well. Thus, faster response is required.
Identification by observing physical characteristics such as device design elements, power connector or device labels.	For cloud environments, this is only applicable for the client side of cloud computing. This includes cases involving small form-factor devices (i.e. smartphones) or instances of a private cloud located in an accessible data center only.
Reverse Lookup (Example): Use a mobile phone number to lookup the network operator.	Within cloud environments, reverse lookups might sometimes be the only available method to identify digital evidence because physical access might be impossible. **Examples**: Through the analysis of a “local” cyber incident, a suspicious remote host had been identified that might still hold valuable data. The only information available at that time is a DNS request issued by a piece of malware caught on a smartphone. The DNS request maps into the IP address space of a major CSP. With the help of this information, the CSP could identify a particular VM and provide a snapshot. However, whether or not this would be successful would depend on various technical and legal circumstances. As stated before, time is critical. Useful information examples for a reverse lookup: - DNS, IP, VLAN, MAC

DEFR needs to take special care with mobile devices. The DEFR must identify all relevant devices.	Although there might be no distinct or physically searchable crime scene in cloud environments, it might be possible to identify relevant client devices including mobile phones. These devices may hold critical information (like cryptographic keys or passwords) that are required to access digital evidence within, for instance, a cloud storage service.
7.2.1.1 Physical incident scene search and documentation Before any acquisition or collection, the incident scene should be recorded in a visual manner by either photographing, video-graphing or sketching the scene as it looked upon entry. This documentation should be balanced with circumstances, cost, time, available resources and priorities.	The physical incident scene might become less relevant for cloud environments because physical access might be impossible, prohibitively expensive or limited to smartphones or smaller private or enterprise cloud deployments. However, documenting the "logical scene" (federated systems, APIs, third-party sub-contractor systems, external storage services, etc.) might become crucial.
Document the type, brand, model and serial numbers of any digital devices. For mobile devices this might include memory cards, cradles, original packaging (which might include PIN or PUK).	Applies unchanged for the client side of cloud environments (smartphones, etc.). It may not be applicable for the cloud environment itself (see above).
Critically evaluate a device (services provided and dependency for other services, how to best protect evidence). Make decisions on disconnecting or taking down a device.	For cloud environments, the understanding of the dependencies becomes a basic requirement to conduct meaningful identification of possible digital evidence participating in a distributed environment. While understanding all dependencies might become complex even in an IaaS deployment, the DEFR will have to limit the evaluation of the overall system dependencies when investigating a SaaS environment and focus on the incident relevant information and system parts (particular software modules and their storage, etc.). Taking down a device might not be required, as virtualization offers new way to "freeze" a system without shutting it down (like snapshotting a VM). On the other hand, taking down a physical host or storage in a provider environment that holds

	hundreds of virtual systems from multiple tenants would be highly disruptive. **Recommendation**: The DEFR should explore and utilize the new options for non-disruptive acquisition, introduced by the virtualization technology.
Preserve the status of digital device (don't switch on/off) unless transport is required and it cannot be done while the device is operating.	Stays the same for mobile devices. As mentioned above, virtualization offers a new way of handling digital evidence and virtual systems. A snapshot might be appropriate to freeze a system for transport and/or investigation without impacting the evidence while keeping the system up and running.
Use wireless signal detector to identify possibly hidden systems.	Applicable for the client side of cloud environments (i.e. mobile devices) only.
Not part of the standard yet Utilize new sources of information introduced by the virtualization technology and their management used for cloud environments	While the cloud environment introduces some challenges for forensic investigations, the introduction of a new abstraction layer between the physical hardware and the computer systems operating system software introduces powerful new options. For example, it creates the capacity to create non-intrusive system snapshots on live systems or plugging into the hypervisor to record data on the OS kernel level. Additionally, virtualization software typically comes along with powerful, central management systems capable of managing an entire virtual data center. Thus, the DEFR should consider the virtualization management systems and APIs as a new source of digital evidence and input (i.e. for creating timelines out of central log information, information on network events regarding virtual switches and firewalls but also VLAN mapping, locating target file systems or systems in scope within virtual containers or on storage systems, and tracking moving VMs, access logs and system configuration details).

Not part of the standard yet If the platform in scope uses encryption provided by the platform, the DEFR should understand the encryption service's topology to evaluate options to access relevant key material.	As mentioned before, encryption becomes increasingly relevant for cloud computing. Thus, more and more CSPs include this functionality into their service wrap. Sometimes it's a free, transparent (cloud platform applies encryption for data stored and the CSP holds the encryption keys) service. In other cases, the CSP offers APIs for encryption at different cloud layers and the customer can either use his own key server (at the CSP or customer premises) or even a third-party service provider. Key management is complex and typically includes different types of keys but also backup and recovery options. The key management infrastructure used within the cloud (topology, processes, technologies) may create the option to make the key accessible for the DEFR.
7.2.2 Collection, acquisition and preservation **7.2.2.1 Overview** Make a decision on whether to collect or acquire potential evidence.	As within a cloud service, potential digital evidence (like a file) is likely fragmented and distributed across the underlying storage infrastructure. As a result, physically collecting the evidence might be impossible. Thus within cloud environments, acquisition is likely the appropriate way in obtaining a digital evidence copy. Mobile devices might be collected however.
In case of acquisition, networked devices should be kept running for further analysis. The DEFR should consider sabotage through other active network connections. The DEFR may either disconnect or monitor the system.	If the DEFR has to consider sabotage, she/he should monitor not only the suspected systems but also the management system(s) and APIs (for the compute, storage and network layer) of the virtualization environment as it could be used to change or remove evidence "below" the virtual machine layer. If the system must be kept running in its original environment, create periodical snapshots of the system to be able to roll back in case of tampering. Alternatively, create a snapshot and bring up the suspected system in a lab environment that simulates the original cloud environment. This allows the DEFR to analyse the system's behaviour and to have more control over inbound and outbound connections.

7.2.2.2 Guidelines for networked device collection Skipped as not feasible or applicable for cloud environments	Collection may only apply to mobile client devices of the cloud environment.
7.2.2.3 Guidelines for networked device acquisition Devices with one physical network connection might be connected to several logical and/or virtual networks. Thus, before disconnecting, the DEFR should conduct a logical acquisition of data related to logical connections	Within cloud environments, virtualized networks and converged network adapters/fabric switches providing the physical connectivity for a whole chassis with multiple computing blades, running multiple VMs, are a common deployment model for IP and storage networks. Depending on the scope of the investigation, the DEFR will have to use the central cloud management system to track down the relevant network setup for the target system(s). When disconnecting a system (VM) for isolation and protection purpose is required, the configuration of the relevant virtual switch should be changed and connected via the virtual network management system to an isolated, forensic VLAN to keep the systems network adapter state up and unchanged.
Blocking wireless connections.	Applicable for mobile devices only.
7.2.2.4 Guidelines for networked device preservation *The DEFR should seal the acquired data using verification functions or digital signatures to determine that the digital evidence copies are equivalent to the originals*	No change for cloud environments.

6.0 Analysis and Interpretation

Once potential digital evidence is acquired and preserved, the processes of analysis, interpretation and reporting can begin as shown in Figure 5.

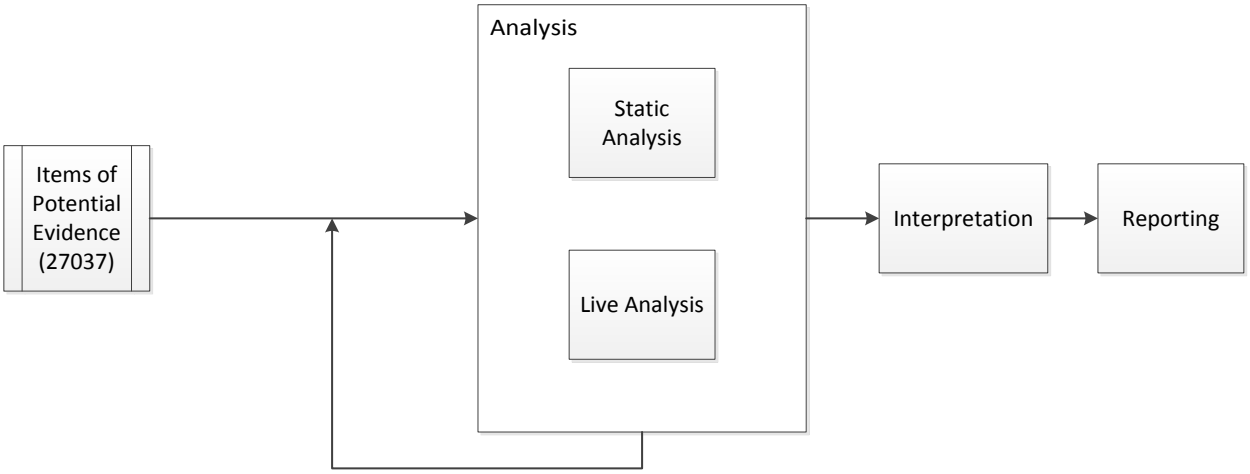

Figure 5: Processes of Analysis, Interpretation and Reporting

These processes are concerned with reconstructing a defensible narrative regarding a series of events or questions of fact arising in the real world based upon the gathered digital evidence. The responsibility for these activities may reside solely with the cloud consumer, or it may be shared with the CSP or other third party. The analysis and interpretation may be provided as an additional service by the CSP.

Analysis includes *"identification and evaluation of items of evidence from a source of potential digital evidence"* [14]. Analysis may be either static (by inspection only) or live (either in situ or by executing a sacrificial copy of an imaged system to observe its behavior). As noted above, analysis may be an iterative process where questions arise during analysis that suggest additional analytical tasks.

Interpretation is the process of *"synthesis of an explanation, within agreed limits, for the factual information about evidence resulting from the set of examinations and analyses making up the investigation"* [14]. In other words, interpretation assesses the meaning of the evidence regarding the real-world questions of fact that gave rise to the investigation.

Reporting covers the presentation of the results of the analysis and interpretation in either written (e.g., a forensic report) or verbal form (e.g., testimony in a legal forum) or both. Reporting is critical in determining the probative value of the evidence in the eyes of the triers of fact.

7.0 Current Status

To date, not all CSPs provide complete forensic support to their clients as a standard offering or via standard APIs. In addition, there are many ongoing challenges regarding forensics in cloud environments. The table below lists a few of these challenges [16]:

Organizational Challenges

- ⊕ Split of control
- ⊕ Segregation of duties
- ⊕ Chain of dependencies
- ⊕ Lack of transparency

Legal Challenges

- ⊕ Multi jurisdiction
- ⊕ Multi tenancy
- ⊕ Data ownership
- ⊕ Privacy
- ⊕ Service level agreement

Technical Challenges

- ⊕ Forensic acquisition
- ⊕ Live forensics
- ⊕ Evidence segregation
- ⊕ Virtualized environment
- ⊕ Data location
- ⊕ Time synchronization
- ⊕ Log management
- ⊕ Identity and anonymity
- ⊕ Data recovery
- ⊕ Proliferation of endpoints
- ⊕ Encryption
- ⊕ Interoperability

However, as a few items from the previous mapping have shown, virtualization and cloud environments sometimes make forensics easier (i.e. VM snapshots). In these cases, cloud forensics provide opportunities including:

- ⊕ Scalability and elasticity
- ⊕ Cost effectiveness
- ⊕ Data abundance
- ⊕ Overall robustness
- ⊕ Forensics as a Service
- ⊕ Security and forensics integration
- ⊕ Standard acceleration

8.0 Conclusion and Future Work

In this paper, we surveyed the issues related to forensic investigation in cloud environments, described in detail international standards for cloud forensics, and summarized the current integration of cloud forensic requirements into service level agreements (SLAs).

As with any new technology, there are challenges in supporting digital investigations in the context of cloud environments. While digital investigations, on the surface, seem to have little to do with the competitive position or profit-and-loss of CSPs, forensic readiness cannot be ignored.

In the short term, the cloud consumer bears the responsibility to ensure that CSPs selected for a particular purpose can respond appropriately to a forensic investigation. This is especially true because consumers ultimately suffer the loss from crimes in the cloud environment.

When contracting for services with a CSP, the customer should ensure that explicit language and SLOs are incorporated into the contract (as shown in the CSA Trusted Cloud Reference Architecture under the "Service Delivery" domain) to ensure they can respond appropriately when the need to perform a digital investigation arises.

For CSPs, integrating forensic capabilities into cloud offerings would increase transparency for the consumer and likely lead to greater revenue streams. As more organizations become reliant on cloud computing for critical operations, we foresee that forensics will become a key motivator on choice of CSP. Additionally, as the cloud market matures, we foresee legal and regulatory changes that may shift duties to include, collaboratively, CSPs.

9.0 References

[1] Orton, I., Alva, A., Endicott-Popovsky, B. Legal Process and Requirements for Cloud Forensic Investigations. In *Cybercrime and Cloud Forensics*, K. Ruan, Ed., IGI Global, 2012.

[2] Zhang, Y., Juels, A., Reiter, M., Ristenpart, T. Cross-VM Side Channels and Their Use to Extract Private Keys, In *Proceedings of the 2012 ACM conference on Computer and communications security* (CCS '12). ACM, New York, NY, USA, 305-316.

[3] Risenpart, T., Tromer, E., Shacham, H., Savage, S. Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds, In *Proceedings of the 16th ACM conference on Computer and communications security* (CCS '09). ACM, New York, NY.

[4] Garfinkel, S. Digital forensics research: The next 10 years, In *The Proceedings of the Tenth Annual DFRWS Conference* (August 2010), vol. 7, pp. S64–73.

[5] Birk, D. Technical Issues of Forensic Investigations Cloud Computing Environments, In *Proceedings of the 6th International Workshop on Systematic Approaches to Digital Forensic Engineering* (SADFE), Oakland, CA, 2011.

[6] Ruan, K., James, J., Carthy, J., Kachadi, T. Key Terms for Service Level Agreement to Support Cloud Forensics', In *Advances in Digital Forensics VIII*, Springer, 2012.

[7] Grobauer, B., Schreck, T. Towards Incident Handling in the Cloud: Challenges and Approaches, In *Proceedings of the 2010 ACM workshop on Cloud computing security workshop* (CCSW '10). ACM, New York, NY, USA, 77-86.

[8] Cloud Security Alliance. TCI Reference Architecture v1.1. Available at https://cloudsecurityalliance.org/wp-content/uploads/2011/11/TCI-Reference-Architecture-1.1.pdf [accessed 22 April 2013].

[9] National Institute of Standards and Technology. The NIST Definition of Cloud Computing. Available at http://csrc.nist.gov/publications/ nistpubs/800-145/SP800-145.pdf, 2011 [accessed January 8, 2012].

[10] Microsoft, Collecting Logging Data by Using Windows Azure Diagnostics. Available at http://msdn.microsoft.com/en-us/library/windowsazure/gg433048.aspx, 2011 [accessed 22 April 2013].

[11] Möller, S., Perlov, C., Jackson, W., Taussig, C., Forrest, S. A polymer/semiconductor write-once read-many-times memory, *Nature,* vol. 426, pp. 166-169, 13 November 2003.

[12] Cloud Security Alliance. Cloud Controls Matrix v1.4. Available at https://cloudsecurityalliance.org/research/ccm/, 2013 [accessed 22 April 2013].

[13] Hay, B., Nance, K. Forensics examination of volatile system data using virtual introspection, *SIGOPS Oper. Syst. Rev.* 42, 3 (April 2008), 74-82.

[14] ISO 27037, Guidelines for identification, collection, acquisition and preservation of digital evidence, Available at http://www.iso.org/iso/catalogue_detail?csnumber=44381, 2012 [accessed 22 April 2013].

[15] Dykstra, J. Riehl, D. Forensic Collection of Electronic Evidence from Infrastructure-As-A-Service Cloud Computing, *Richmond Journal of Law and Technology* 19, Available at http://jolt.richmond.edu/?p=463, 2012 [accessed 12 June 2013].

[16] Ruan, Keyun. "Cloud Forensics: Assessing Cloud Computing's Impact on Digital Investigation". Presentation. The Inaugural Cloud Security Alliance Congress EMEA, 25-26 September 2012, Amsterdam, Netherlands.

10.0 Acronyms

IaaS	Infrastructure as a Service
PaaS	Platform as a Service
SaaS	Software as a Service
SLO	Service Level Objective
SLA	Service Level Agreement

CSP	Cloud Service Provider
LEA	Law Enforcement Agency
NIST	National Institute of Standards and Technology
AAFS	American Academy of Forensic Sciences
BOSS	Business Operation Support Services
ITOS	Information Technology Operation and Support
CCM	Cloud Control Matrix
WORM	Write-Once-Read-Many
SAN	Storage Area Network
NAS	Network Attached Storage