Cloud computing describes a computing concept where software services, and the resources they use, operate as (and on) a virtualised platform across many different host machines, connected by the Internet or an organisation’s internal network. From a business or system user’s point of view, the cloud provides, via virtualisation, a single platform or service collection in which it can operate.

Exposure to threats

Cloud computing involves potentially greater exposure to security threats and privacy breaches, especially when the cloud is based on the Internet rather than an organisation’s own internal network. For example, it might be unclear as to where data is processed within a cloud computing system, and such processing can occur in differing jurisdictions. Current commercial cloud service providers include Microsoft Azure Services Platform, Amazon Web Services and Google, as well as open source cloud systems such as Sun Open Cloud Platform. There are three typical cloud service delivery models:

• The Software as a Service (SaaS) model where the customer rents software for use on a subscription or pay-per-use model.
• The Platform as a Service (PaaS) model where the customer rents a development environment for application developers.
• The Infrastructure as a Service (IaaS) model where the customer rents the hardware infrastructure on a subscription or pay-per-use model, and the service can be scaled depending upon demand.

The nature of cloud computing services means that, in practice, an organisation may not know where data for which it is responsible is geographically located at any particular time. Some public cloud service providers do provide regional data storage options – for example, US and EU data storage options. However, data may be replicated by some cloud providers, so multiple sources may exist. In addition the cloud provider should know where virtual machine instances are running and have a historical record.

Fighting crime

Computer forensics has emerged in recent years as an important tool in the fight against crime. It is defined as the application of computer investigation and analysis techniques to determine potential evidence. Traditionally, computer forensics has classified crime involving computers and associated technologies in three ways – the computer is: the target of the crime; a repository of information used or generated during the commission of a crime; or a tool used in committing a crime. These definitions focus on a more traditional view of digital forensics involving data, software, operating systems and analysis of seized physical computing devices. However, this analytical paradigm is less applicable within the cloud computing environment as it is distributed in all senses – from users to applications, from servers to data.

Standards and protocols relating to the identification, preservation, recovery and analysis of computer-based evidence are becoming important as organisations
increasingly require computer forensic services or investigations. Organisations need to understand the computer forensic process and their legal obligations in order to ensure the integrity and robustness of such investigations. It is important that any computer forensic investigation demonstrates integrity, in terms of the evidence collected (that it is correct and has not been altered in any way) and in terms of the methodologies used in the collection, preservation and analysis of digital data. Unless the integrity of digital evidence is maintained, it is unlikely to be admissible in court, if criminal activity is identified by the investigation.

“The issues relating to preservation of data in a cloud computing environment would concern access to the data prior to it being seized, and the preservation of the data being done correctly.”

Evidence acquisition

Digital evidence can be more ethereal and dynamic in the virtual environments provided in a cloud computing system. If a software application is accessed via a cloud computing system, data traditionally written to the operating system, such as registry entries or temporary Internet files, will reside or be stored within the virtual environment and so lost when the user exits the cloud. Virtualisation sanitises resources so the traditional analysis of leftover artefacts could be limited. This can make digital evidence traditionally stored on hard drives potentially unrecoverable.

Also the legal process required to gain authority for access to data held in a public cloud computing system, especially one that might utilise computing devices in different jurisdictions, can be potentially time consuming and could delay investigations where the recovery of digital evidence is typically time critical. Although the legal framework for forensic investigations is intrinsically local, there are jurisdictional implications with regard to globally distributed cloud services. Although similar jurisdictional considerations may have existed to some extent since the advent of the Internet, cloud services have already involved, and will further involve more complex jurisdictional concerns for forensic investigators. For example, the Interpol investigation of Mr Swirl, or Swirl Face (an individual who made worldwide broadcasts of indecent images) involved co-operation between the police forces of many different countries.

The identification of evidence within the cloud computing environment can be very complex. The first consideration that the analyst must make is whether it is a private or public cloud (or, indeed, a hybrid of the two). As Grossman notes, a private cloud is for a single organisation’s internal use and it can be run by the organisation itself or outsourced to a third party. A public cloud is managed by another organisation that provides cloud services. These clouds offer publicly accessible remote interfaces for creating and managing data. This general architecture will have serious ramifications
for the identification of evidence. If, for example, the analyst is examining evidence within a private cloud, data will reside within the organisation or within its outsourced supplier. The key sources of potential evidence will be identifiable, such as servers, applications, and data repositories residing within the organisational IT infrastructure. In addition, the investigating team may also have access to key personnel identified by the investigation, such as the suspect or system administrators.

“Even the existence of data will be quite complex to identify as data is pushed further back into the network rather than purely being delivered to the user’s physical computing device”

However, if the evidence resides within a public cloud, it will be much more difficult to identify. As Treacy suggests, the cloud computing environment aims to be dynamic and customisable.6 This is achieved through the seamless interaction of a variety of applications being delivered to the user as if they were accessing just a single site or logical location. This seamless delivery from distributed sources will make the identification of sources of potential evidence, or the evidence itself, much more complex. Moreover, even the existence of data will be quite complex to identify as data is pushed further back into the network rather than purely being delivered to the user’s physical computing device.

There are different computer forensic challenges related to the different cloud computing models, PaaS, IaaS and SaaS. These models present subtly different challenges to the forensic investigator:

- **Remote web access**: key evidence may reside across several client browsers’ web history/caches, presenting difficulty in collection, collation and verification.
- **Virtualised platform or resource**: evidence collection and/or seizure may have to occur via the virtualisation software, with the potential to render the digital information forensically unsound.

### Third-party location

Related to this architectural issue is that of the third party’s location. Potentially, this can have a major impact on the investigatory jurisdiction and responsibility. Currently, the procedure for identifying and extracting evidence is the seizure of the computing device itself. The investigation team has the tangible evidence of the device and they are able to analyse the data held on it. However, the cloud computing paradigm aims to push data and services back into the network. Therefore, there could be some doubt as to the actual location, either logically or physically, of the source of potential evidence.

In addition, a computer forensics analyst views information found in a variety of sources within the computer in order to produce a case. Without access to the physical device itself, the analyst may only have a few files on which to base his or her analysis. This might not be enough evidence to provide a robust case for a successful conviction. Within the UK legal system, prosecutions are based upon evidence that a specific person, at a specific time, date and place, undertook unlawful action (actus reus), and that they knew or should have known the action was unlawful (mens rea). A few computer evidence files might be sufficient for a prosecution, but it might be difficult to prove that a specific individual undertook a specific action at a particular time, date and place.

So, although cloud computing is more complex, as there is OS and application independence, the focus of a forensic investigation is still on events. Therefore the mechanisms to control access, and the location where information is to be delivered could be more important from a prosecution point of view as they provide evidence of a specific person undertaking a specific action at a specified time, date and place. To some extent the location of the artefact that is being used that is within the cloud is not that important. What is important for the prosecution is evidence concerning logs of access, evidence regarding the individuals who control access, what kind of access permissions are involved, and how and where access occurs. If this kind of information is present then the methodology for investigation in a cloud environment might be no different from any other computing environment – the cloud-based issues mainly concern preserving the data.

“Identification of evidence in cloud computing is more complex as data could be spread out across these applications and servers”

In traditional computer forensics investigations, evidence is distributed across a small number of sources such as hard drives, network servers, cell phones, etc.7,8 The analysis of these allows the examiner to retrieve a large amount of information regarding the suspect’s activities from this small number of sources. An advantage of cloud computing is that multiple applications and servers across geopolitical locations are able to interact seamlessly to provide the services and applications that a user requires. Moreover, as Pendyala and Shim suggest, a major advantage of cloud computing is independence from hardware and operating system profiles.9 In addition, the end-user experience requires data to be processed by multiple applications or computers but delivered as if it comes from a single source computer. Therefore, identification of evidence in cloud computing is more complex as data could be spread out across these applications and servers.

### Physical seizure

Cloud computing impacts upon the ability of law enforcement agencies to physically seize computing assets in order to pursue an investigation. If a law enforcement agency could find the appropriate server systems in a timely manner, it would be unlikely that they could get to information in a timely manner, and in particular, the agency could have difficulty in seizing such systems let alone ‘seize’ a datacentre.

Cloud computing can be broken down into different categories. These include virtualisation technology, remote hosting
and software as a service (that is software applications). Each of these areas can have different challenges from a computer forensics perspective.

With regard to virtualisation technologies, the challenge is whether the computer forensic analyst needs to be concerned with the client computer operating system or the host computer operating system. If the investigation is mainly concerned with the client, then forensic analysis may be performed from the host operating system, which can be advantageous for the forensic analyst.

With regard to remote hosting, the main computer forensic challenge is that there could be difficulty in gaining physical access to the relevant server computer for the purpose of making a disk image. If the hosting service provides forensic analysis as an option, then this difficulty may be reduced. Copying VMware machines can provide a simple means of obtaining a forensically sound copy of virtual servers that is self contained.

With regard to software as a service, the challenge is that computer forensics will be different from traditional computer forensics. Although this aspect of computer forensics is relatively new, forensic investigations have been conducted on Hotmail, Gmail and other web-based email services for some years.

Suspects in the cloud

Finally, there is the issue of identifying the actual suspects within the cloud environment. In traditional forensics, a computer will be seized that has physical links to the suspect. For example, a computer might be seized from the suspect’s home or work environment, thus physically tying the suspect to the machine and the evidence. Within the network environment, computers interact without a suspect’s knowledge. As there is no physical interaction, identities within this networked environment often rely on usernames and passwords. For many years these have been seen by security experts as quite insecure. For example, a malicious person might intercept data in transit and gain a username and password to use in later nefarious activities. Therefore, it is more problematic to link a suspect to malicious activities within this virtual environment.

"Within a cloud environment, larger-scale encryption may be used by organisations concerned with the security of their data within the cloud"

Extraction of data and evidence can be complicated within the cloud computing environment. In addition to the issues identified above, a number of problems will need to be addressed by computer forensics analysts in extracting evidence in the distributed environment. Due to the requirement of placing large amounts of sensitive data within a distributed environment, proponents of cloud computing have suggested that security should be a major concern. Researchers have proposed encryption as one means of ensuring end-user security. As with more traditional computer forensics investigations, any form of encryption places a large burden on the forensics investigation and increases the complexity of the investigation. Evidence has to be identified, which is complex with encrypted data as it defeats widely-used techniques such as keyword searches, but it also must be decrypted (if it has been identified in the first place). Although data may be encrypted in common Internet-based activities, such as e-commerce, and some organisations might even encrypt data processed on in-house hardware, within a cloud environment, larger-scale encryption may be used by organisations concerned with the security of their data within the cloud.

Related to encrypted data is the proposal for more secure authentication within the cloud computing environment. Authentication will ensure that restricted access is enabled to all but those that need to access that data. Therefore, investigators may have to complicate their interaction with the law – for example, their invocation of clauses of the UK Regulation of Investigatory Powers Act (RIPA) 2000 – to enable access to this additional data. However, more secure authentication could also be an advantage for the computer forensics analyst. For example, if a user interacts with the Public Key Infrastructure (PKI), particularly using public/private key pairings and certification, evidence can be undeniably associated with a particular suspect.

Evidence analysis

It could be difficult to analyse the sequence of events in a particular transaction in a cloud computing system since a variety of different machines might have been involved in the transaction – for example, a transaction involving updates to a large number of database tables in a distributed database. Another challenge facing the forensics community when analysing evidence in cloud computing environments is the degree of software standardisation, as there are several virtualisation platforms. As with the forensics analysis of mobile platforms, this may cause complexity in the investigation due to the need to extract evidence from a wide range of applications or servers.

For example, in the computer market, there are very few operating systems compared to the mobile phone market. Tools such as EnCase and FTK can be updated to account for any new developments in this market, and will apply to the majority of computers. In the mobile phone market, there are many operating system variants and legacy devices that must be accounted for by the examiner. Cloud computing often uses newly written applications to provide the environment that users expect. Users are able to access relevant data or services through a variety of devices, from a desktop PC to a mobile phone as well as a wide range of applications. This lack of standardisation and cross-platform development makes it more problematic to write programs for the forensic extraction of data.

Meta data embedded within documents that had subsequently entered the cloud storage could provide important clues to how the data has been used and manipulated beforehand (such as change tracking in MS Word documents). Financial services organisations and some other types of organisations might typically have audit trails built into their cloud-based application systems that can be used to provide...
digital evidence. If an investigation of a cloud computing environment involved the analysis of email evidence, then typically logs of sent and received emails from the user's computer could be used as evidence, unless the tampering of emails was being investigated, in which case evidence from the computing devices within the cloud might be required.

Unless a cloud computing application provides some form of audit trail, it may be difficult to extract digital evidence in an admissible manner from such applications, and in some cases, there may be little evidence available to extract. In such instances, digital evidence potentially available from the user's computer (assuming that this can be identified) may, in some cases, have to be the major source of digital evidence relating to cloud-based activity. As an example, if a computer forensic investigation involved analysis of a Google document transaction, then with regard to user data stored on the user’s personal computer after such a Google document transaction there would be cookies for user login and documents and also Google gears may have created a SQLite database on the user's machine to allow the user to work offline. All these artefacts stored on the user's personal computer could provide potential evidence, even if further digital evidence from computers in the Google cloud could not be easily obtained.

Audit trail

Cloud computing systems could potentially be made easier to forensically investigate whether some form of audit trail was maintained by the cloud system. For example, with regard to SaaS cloud environments, the software applications themselves could maintain appropriate audit trails of changes made to application data. With IaaS and PaaS cloud environments, logs could be maintained of users’ activity within the cloud environment. However, given the typically large scale of cloud environments, such approaches would involve the storage of large volumes of data, which would not be welcomed by cloud service providers. Compression techniques could potentially be used to reduce the volume of audit trail and log data stored: however, even with compression techniques, the volumes of data would be considerable. Equally, in remote-access service platforms, internal software-specific audit trails may maintain login/action histories of interest to investigators, which may or may not be made available. However, externally-kept logs may be of equal interest, which may be stored or fall into the scope or jurisdiction of the investigator. In the case of web-based SaaS platforms, a useful source of external logs may be the investigated user’s web browser history or cache. However, obtaining the complete user cache may be impossible or may necessitate seizing several machines – home, work and – potentially – mobile devices.

“If data is crossing geopolitical borders rather than logical infrastructures, more international co-operation will be required in conducting investigations”

Finally, as discussed above, the cloud computing environment abstracts users away from the physical computers that they are using. This leads to the problem of extracting evidence from the relevant source(s). This is less problematic in the case of private clouds where relevant evidence may reside predominantly within the organisational boundary and, as such, can be more readily extracted. However, public clouds make the extraction of evidence more complicated. For example, a suspect may access an application on a server based in the US from a mobile phone in the UK. In this interaction, no code may be executed on the client machine. Therefore, if a UK-based investigator wishes to extract evidence, no evidence would be available as it resides on a different machine and in another jurisdiction. In addition, the offence that is being investigated may be a crime in the UK but not in the US. Therefore, these third parties may refuse access to the evidence. Related to the issue of jurisdiction is that of agency responsibility for the investigation. If data is crossing geopolitical borders rather than logical infrastructures, more international co-operation will be required in conducting investigations. In addition, there are issues of data crossing these geopolitical borders. Evidence may be missed as it only exists in the jurisdiction area for a very short time. As such, investigatory authorities may not know of the existence of evidence or may be unable to collect this highly transient data.

Management of computer forensic investigations

The first stage of a computer forensic investigation involving a cloud computing system would be to clearly identify the purpose of the investigation. For example, the investigation might concern unauthorised access to application systems based in the cloud, suspected fraud or money laundering or the accessing or dissemination of offensive material. It is important that the purpose of a computer forensic investigation is clearly defined so that the full scope of the investigatory process can be decided. Management support in the investigation of a private or hybrid private-public cloud system would be required to ensure that the disruption to the organisation’s operations is minimised where possible during the investigation.

A cloud computing forensic investigation may appear daunting due to the potential size of the system to be investigated. However, if the cloud computing system just involves remote service, for example with IaaS systems, logging and authentication may provide much of the digital evidence required. Virtualised cloud computing environments can be handled with logging and forensics in monitors. In a cloud computing environment, actions taken from the moment that misuse or criminal activity is suspected can have a profound impact on both the amount of digital evidence available and the extent to which it will be acceptable in future legal proceedings. It is important that organisations have defined procedures and practices that are adhered to in any computer forensic activities that are undertaken within the organisation, in order that digital evidence is not contaminated in any way, especially if such digital evidence might later be required.
by a police force (or other agency) for a criminal prosecution.

Currently there do not appear to be any published guidelines that specifically address the conduct of the computer forensic investigation of cloud computing systems. For example, the UK Association of Chief Police Officers’ (ACPO) good practice guide for computer-based electronic evidence does not specifically cover cloud-based computing systems. The ACPO guide does, however, cover crime scenes on the Internet, in terms of the forensic examination of physical devices connected to the Internet, and the necessity in some instances of capturing evidence directly from the Internet possibly during ‘live’ interaction with a suspect or by capturing live website content. The use of checkpoints might be useful in order to make live systems easier to analyse. The ACPO guide also provides advice regarding the difficulty in capturing evidence where target machine(s) may be sited outside UK jurisdiction.

Forensic agent

A private cloud computing system is for a single organisation’s internal use and it may be run by the organisation itself or outsourced to a third party. Even in such a cloud computing environment within an organisation it may be difficult to shut down servers. In the public or hybrid private-public cloud computing environment this would be even more complex. A public cloud is managed by another organisation that provides cloud services. The more dispersed architecture of a public cloud system can pose difficulties for the identification of digital evidence. In such instances where server shut down would not be possible, it is common practice within UK police forces that a network-enabled ‘forensic software’ agent be installed, which can give the ability to image data across the network on-the-fly. However, other forensic software is available that does not entail the installation of an agent. In internal corporate cloud computing environments, devices such as routers and firewalls could be used to give an insight into network configuration through Access Control Lists (ACLs) or security rule sets. This may be achieved by viewing the configuration screens as an administrator of the device.

“When digital evidence is required from a public cloud service provider there may also be the issue of continuity of service for other users of the cloud services”

If a computer forensic investigation involves a private or a hybrid private-public cloud, then at least some digital data will reside within the organisation or within its outsourced supplier. Some of the sources of potential evidence will be identifiable, such as servers, applications and data repositories residing within the organisation’s IT infrastructure. However, if the digital evidence resides within a public cloud, it will be much more difficult to identify. Moreover, even the existence of digital evidence might be complex to identify as data is pushed further back into the network rather than purely being delivered to the user’s physical computing device and may only exist within tight temporal constraints. When digital evidence is required from a public cloud service provider there may also be the issue of continuity of service for other users of the cloud services. Ideally a computer forensic investigation should not impact upon other cloud service users who are not the target of the investigation.

Any computer forensic investigation should keep within the appropriate guidelines for computer-based electronic evidence within the jurisdiction concerned. That is to show a court, if required, that the digital evidence produced is no more and no less than when it was first taken into the possession of the forensic examiner. At present, commonly adopted computer-based electronic evidence guidelines do not specifically appear to address cloud computing investigations. However, their principles should still be maintained in any cloud computing forensic investigation. In this respect, cloud computing systems investigations are no different from other forms of computer forensic investigations. Legislation such as RIPA makes it unlawful to intercept any communication in the course of transmission without the consent of one of the parties or without lawful authority. Appropriate internal corporate authorisation would be required for any investigation of an organisation’s internal private cloud system. Investigation of a public cloud computing system involving Internet-based computing resources would require the cloud computing services provider to provide the police (or other agency) investigation with required digital data. However, due to the nature of cloud computing systems, some of the digital data required may not be practicable to obtain from the cloud.

In a cloud computing environment, due to the potentially greater effort required to identify and examine computing devices that had stored or processed digital data of interest to the investigation, there might be limited time and resources available to identify digital material of wider relevance than that which specifically concerns the investigation. If a computer forensic investigation needed to consider the potential effects of malware – such as spyware, computer viruses, worms and trojan software within a cloud computing environment – this might be complex. If a defence case relied upon the actions of malicious software within the cloud without the knowledge of the accused, it might be difficult to obtain digital evidence to support such a defence.

Conclusions

Cloud computing is likely to make the acquisition and analysis of digital evidence more complex. Computer forensic investigations of cloud computing systems are likely to require more time and effort to undertake, due to the number of computing devices within the cloud that may need to be forensically examined. In addition, in legal terms, cloud computing systems will make it potentially more difficult for the computer forensic analyst to acquire and analyse digital evidence to the same standards as that currently expected for traditional server-based systems, due to the difficulty in establishing what data was stored or processed by what software on what specific computing device. This might lead to legislation requiring public
A secure data protection strategy

Russell Sanderson, Iron Mountain UK

Information is the lifeblood of your business. Do you know what information you are holding, who’s holding it and whether it’s safe?

Let’s start by examining a true story. An IT administrator in a UK-based company steals some backup tapes, one of which contains important data about the company’s European operations. In an attempt to extort money, he alerts the company anonymously and demands a £275,000 ransom fee. The company works closely...