2011

Organisational preparedness for hosted virtual desktops in the context of digital forensics

Nirbhay Jawale
Digital Forensic Research Laboratories, Auckland

Ajit Narayanan
Digital Forensic Research Laboratories, Auckland

DOI: 10.4225/75/57b2be5e40ced
Originally published in the Proceedings of the 9th Australian Digital Forensics Conference, Edith Cowan University, Perth Western Australia, 5th -7th December 2011

ORGANISATIONAL PREPAREDNESS FOR HOSTED VIRTUAL DESTOPS IN THE CONTEXT OF DIGITAL FORENSICS

Nirbhay Jawale, Ajit Narayanan
AUT University
Digital Forensic Research Laboratories
Auckland, New Zealand
j.nirbhay@gmail.com

Abstract

Virtualization in computing has progressed to an extent where desktops can be virtualized and accessed from anywhere. The server hosted model has already surpassed 1% market share of the worldwide professional PC market, with estimates indicating that this is a rapidly growing area. This paper investigates the adequacy of current digital forensic procedures on hosted virtual desktops (HVDs) as there does not appear to be specific methods of locating and extracting evidences from this infrastructure. A hosted virtual desktop deployed in private clouds was simulated to reflect two different computer crime scenarios. It was found that current digital forensic procedures may not be adequate for locating and extracting evidence, since the infrastructure introduces complications such as persistent/non-persisted disk modes and segregating data in a multi-tenant environment.

Keywords

Hosted virtual desktops; digital forensics; server hosted models.

INTRODUCTION

The development of virtualization started in 1960, when VMware introduced partitioning of large mainframes for better hardware utilization. (Virtualization History, 2010) Since then virtualization has matured and been adopted to a wide extent in the industry. Recent developments include branching into areas of server virtualization, storage and application virtualization and, very recently, desktop virtualization. Desktop virtualization has so far been through two models: the Client hosted model, which is typically operated from the user's workstation using Windows Virtual PC; and the VMware workstation or Java Virtual Machine (VM). However, recently a third model has emerged, called the server hosted model or Hosted Virtual Desktop (HVD), which is a virtualized form of desktop (VM) delivered to users from the cloud infrastructure. In other words, users see a desktop that looks identical to a standard, local desktop, but the desktop is virtualised and operated from a remote server. Enterprises are increasingly considering HVD solutions to control hardware costs as well as to reduce maintenance and upgrading costs: a HVD is automatically updated remotely. All that is required is an Internet connection and a local screen, with enough memory to run the HVD. According to (Gartner, 2009) HVD has surpassed 1% market share of the worldwide professional PC market already, estimating that "15 percent of current worldwide traditional professional desktop PCs will migrate to HVD by 2014, equal to about 66 million connected devices" as the demand to access one's desktop from anywhere continues to grow. Likewise the use of this model for in-house cloud (private cloud) computing is expected to grow as organizations choose lower cost options involving vendors offering products that will allow organisations to build their own private cloud. The question then arises as to how HVDs will impact on digital forensics. Forensic investigators Diane Barrett and Gregory Kipper mention in their book (Barrett & Kipper, 2010) that forensic experts can perform procedures on physical machines in full confidence because of the physical presence of the hardware in the forensic laboratory. If a system is fully virtualized, the confidence may not be at the same level. This is due to various characteristics of the environment, such as the data storage in the cloud environment being in a state of flux, custody of data not being guaranteed(Taylor, Haggerty, Gresty, & Hegarty, 2010), and the data not being segregated as a consequence of resource sharing by users in a multi-tenant environment (Ruan & Carthy, 2011).

Most importantly there does not appear to be a specific method or set of methods for extracting evidence from cloud systems. Although in past several techniques to investigate the presence of a VM or using a VM as a forensic tool have been proposed, not many exist on investigating the cloud as an infrastructure. It is possible that current digital forensic techniques are sufficient to investigate crimes in the cloud, but to date there has been no systematic investigation of the adequacy of existing digital forensic techniques and methods when applied to

possible crimes committed in the cloud. The aim of this study is to investigate the question of whether new cloud-oriented digital forensic techniques are required or whether existing techniques can continue to be used for investigating security breaches and crimes committed using HVDs. .

METHODS AND EXPERIMENTS

The experiments are conducted through simulations of two case scenarios. The aim of the experiments is to locate and extract evidence within the HVD infrastructure. The case studies used to generate the experiments are based on plausible scenarios typically found in standard forensic IT investigations but modified to be applicable to a HVD infrastructure and environment. The methods adopted are taken from standard digital forensic procedures with the aim of identifying their appropriateness and adequacy for analysing the data and results of the experiments.

With regard to the simulation infrastructure. the simulated model of a hosted virtual desktop in a private cloud is constructed on a workstation consisting VMware Workstation 7 and VMware 4.5 evaluation version. To deploy HVD successfully, five major components are required: Active Directory, ESX Server, Connection Server, Infrastructure Tool and Event Database. To accommodate the essential components of the infrastructure, five VM's were created with relevant operating systems, as shown in Figure 1. Every component of the infrastructure is on the same LAN segment because every component needs to be part of the same domain with a static IP address.

Virtual Machine	Guest Operating System	LAN 1	LAN 2
Active Directory	Windows Server 2008		✔
Connection Server	Windows Server 2008		✔
ESX Server	ESX Server 4.0		✔
Manage HVD	Windows XP Professional		✔
Test User	Windows XP Professional		✔
EVENTDB	Windows Server 2008		✔

Figure 1. HVD Infrastructure In VMware Workstation 7

The purpose of Active directory in this infrastructure is to manage users, groups and apply control using group policy while the virtual desktops are in operation. ESX Server is a hypervisor which stores every VM created. The Connection Server is the key component that is responsible for maintaining the connection between the VM stored on the hypervisor and clients to deliver desktops virtually. It can also be used to group VM's in pools and apply permissions. VMware vSphere is used to implement and maintain VM's (Manage HVD). It can also be used to maintain hardware allocations for every VM. The event database (EVENTDB) maintains records related to VM's and changes according to the infrastructure. Data such as time and date of VM creation, shutdown, and restart by which user is held in this database. Users (Test User) within the infrastructure are generated as required by creating number of VM configured with Windows XP. Every VM needs to be implemented with View Agent software and also configured on the infrastructures domain to be visible in the infrastructure. Similarly, Client Agent needs to be installed on every client machine trying to access their VM.

The method consists of two case scenarios being designed and developed to produce simulations that, in turn, lead to data and results for evaluation against currently used forensic IT tools, methods and techniques.

SCENARIO 1

In a company, the vice president (VP) is suspected of surfing pornographic websites during office hours. An anonymous complaint has been lodged with the company's director of HR that the VP's desktop screen had been briefly observed displaying pornographic material. Due to the high standing of the suspect, the company's HR department wants to gather solid evidence of the VP's activities before confronting the VP. The director of HR approaches a computer forensic team to investigate whether evidence exists concerning the VP's alleged activities. In the company every employee has either a designated workstation or a laptop connected to the company's in-house cloud infrastructure. The VP was using his company laptop while surfing and could access external sites through the company's in-house cloud.

SCENARIO 2

A school's ISP has detected and blocked the attempt of a user in the school to surf an offshore network containing objectionable material. The ISP wants to warn the school's head teacher for everyone's safety.

The school's administrator has limited knowledge about the recently deployed virtual desktops, which are linked to an in-house cloud. The head teacher doesn't want to risk the school's reputation, so a private computer forensic team is hired to investigate this matter and report their findings.

RESULTS

Both scenarios were investigated using widely accepted digital forensic procedure and best practice was followed as far as possible. So, in the case of the first scenario, the following are assumed: photographing the crime scene and involved hardware; measuring systems power states; collecting live data if possible; Producing connectivity diagrams of the system; labelling evidence; documenting equipment with serial numbers; and maintaining a chain of custody (Henry, 2009). The progress of the investigation was marked using the core elements of the investigation process i.e. Preparation, Investigation and Presentation. The person under investigation is the suspect.

(a) Scenario 1

Aim: Perform forensically sound data acquisition and find evidence useful for proving that the suspect was surfing pornography during office hours while using the company's IT resources.

GENERAL FINDINGS

- The company has recently deployed their private cloud and has hosted virtual desktops in place.
- The infrastructure consists of 100 VM's organized in three pools, i.e. task workers, power users and kiosk users.
- The suspect is part of the power user pool, which consists of stateful VM's.
- The reported time and day of the alleged incident was on 22/07/11 around 2:50pm, according to the anonymous report.

FORENSIC PROCESS

Search & Recognition: The suspected location of potential evidence is the VP's laptop; there is no other workstation or mobile device in the office, as far as can be ascertained. To avoid evidence tampering and alterations, network and local activities were paused.

Collection: The VP's laptop was seized while it was powered on; hence a live acquisition on the laptop's local disk could be and was performed. The virtual desktop was still active; hence a live acquisition could be and was performed on the virtual desktop also. However, performing live acquisition on the virtual desktop was slightly more complex than performing it on a local machine. Acquiring a live VM hosted on ESX server was possible by two techniques.

Technique one involved the use the e-fense Helix disk to acquire live image of the system. This requires an independent collection machine connected to the domain. To prepare the collection, the following autoplay was performed (also possible with Helix.exe):

Select Incident Response - Start NetCat Listener - Define Listening Ports - Define Image name/Location - Wait for the "Notice" Prompt.

A Notice dialogue will prompt for a confirmation of the NetCat command you are about to run and, upon confirmation, a command prompt will display the port number and the user being listened to (Figure 2).

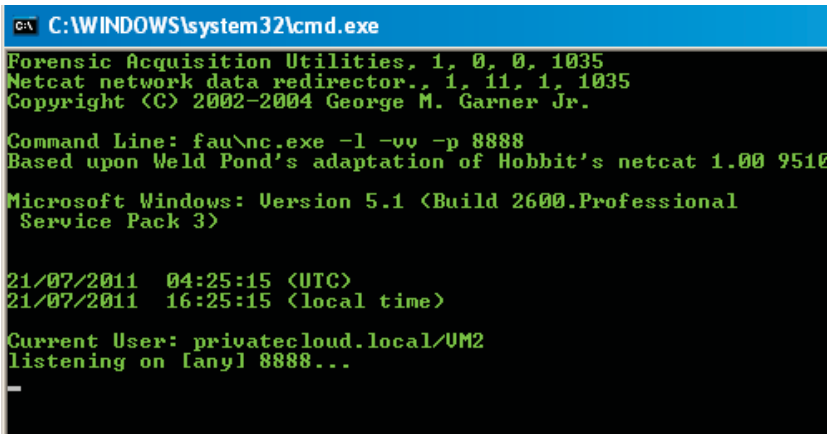

Figure 2. NetCat Listening

The suspect's machine needs to be prepared for live acquisition via the Virtual Infrastructure tool based on the management console machine i.e. VMware vSphere. In order to capture the current state of the VM, Helix CD or Image needs to be emulated via the management console. Upon the launch of Helix.exe or Autoplay, a window opens, such as in Figure 3.

Figure 3: Prepare Suspects VM

Select Acquisition - Source: HDD or Memory - Location: NetCat - Destination: collection machine IP - Image name: default - Block size: default - Conv: noerror - Acquire!

This will acquire HDD/Memory of the VM stored on the ESX server and write to an image on the collection machine using NetCat. It is considered good practice to save the activity log after the prompt, as part of the final documentation.

The second technique involves the use of vCenter server or vCenter Standalone converter to capture the live copy of the running VM. This can be achieved by suspending/Pausing the suspect's VM via an infrastructure tool like vSphere. Suspending will enable the live copy to be frozen in the current state, enabling the forensic team to identify actual running processes and potential information in the volatile memory during the analysis stage. Once suspended, a vCenter Standalone converter can be pointed at the suspect's VM using its IP address and the copy process will begin (Figure 4). vCenter converter can be run on the collection machine as long as it is in the same domain. If the private cloud is managed using the vCenter Server the process is very straightforward: Select the suspect's VM, Right Click and Copy.

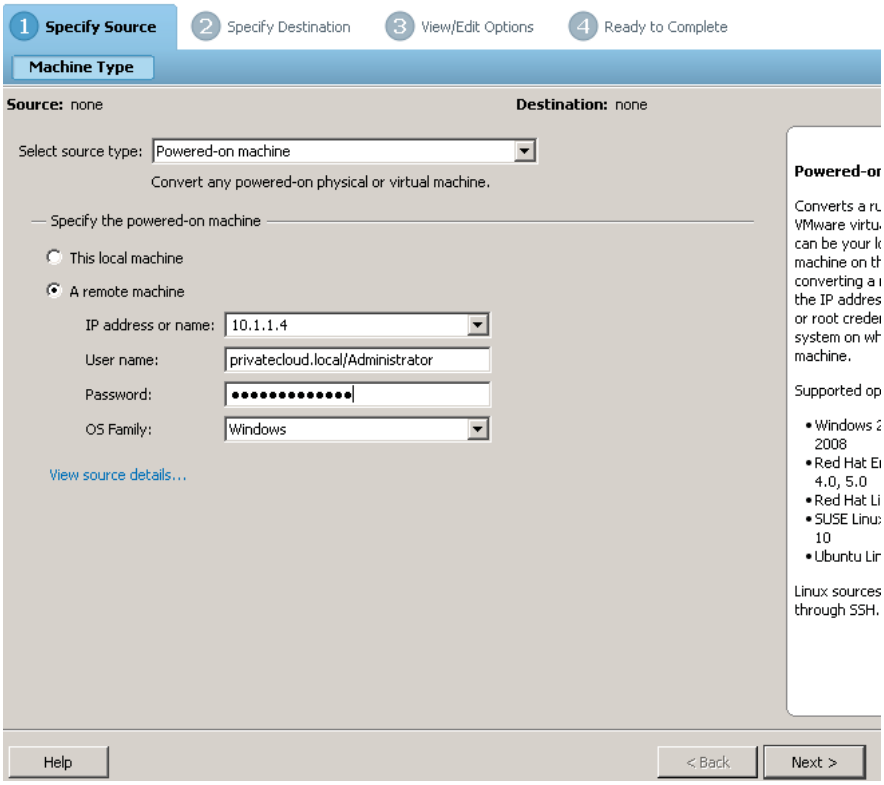

Figure 4. vCenter Standalone Converter

As part of comprehensive collection phase along with acquiring the user environment, system logs were also collected from the following **Location of VMware View log files:**

Event Data base server

View Admin Console \ Monitoring\Events - Filter by suspects VM - Export.

Connection Server and Security Server logs

<DriveLetter>:ProgramData\Application Data\VMware\VDM\logs

View Agent & Client Logs

Suspects VM\<DriveLetter>:\Documents and Settings\All Users`Application`* Data\VMware\VDM\logs.*

Suspects Laptop\C:\Documents and Settings\%username%\Local Settings\Application Data\VMware\VDM\Logs

In this case scenario, both techniques were used to acquire the evidence to maximise the measurements made during the investigation analysis phase of the captured artifacts. However both techniques have technical constraints making them less forensically sound. The first technique involves running the e-fence Helix CD in the live state of the VM and relies on NetCat to transfer the output to a collection machine. In networks that restrict the use of NetCat this technique may not be successful. The second technique relies on a management tool to hold the live state of the VM while a conversion tool copies it to a collection machine. But the success of this technique relies on the standalone converter agent on VM's. The agent allows the standalone converter to perform remote cloning during any state of the VM. The agent may or may not be installed by default on the suspect's VM, depending on the template used to create the VM in the first place. Although including the agent in the template is not part of practice guide (Dodge, 2006), it is highly recommended for administrators and forensic readiness. The agent however can be installed later but may alter system files at the same time.

Examination & Analysis:

During the examination process the following evidence was collected:-

- Image of suspect's physical machine
- Image of suspect's VM located on the ESX server in a suspended state
- System logs

The image of the suspect's physical machine was analysed using forensic software and the results show no traces of access to objectionable images. However, objectionable images were found in temporary internet files upon analysing the suspect's VM. Other attempts to visit pornographic websites during office hours were also found logged in the internet history.

Although it is possible that the suspect's VM may have been logged into by other user in the network, the View Client logs found on the suspect's physical machine and event logs on the server show the time stamps of user login/log off times. These were found to match the timestamps of internet history and objectionable images written to disk, as illustrated in Figure 5 and 6. This strongly suggests that the user who engaged in objectionable activity on the targeted VM was logged in via the VP's laptop.

User	Severity	Time 1 ▾	Module	Message
PRIVATECLOUD\vice.president	Audit success	7/22/11 3:10:39	Connection Server	User PRIVATECLOUD\vice.president has logged out
PRIVATECLOUD\vice.president	Info	7/22/11 2:52:43	Agent	User PRIVATECLOUD\vice.president has logged in to a new session on
PRIVATECLOUD\vice.president	Info	7/22/11 2:51:46	Agent	The agent running on machine XPVM has accepted an allocated session for user
PRIVATECLOUD\vice.president	Info	7/22/11 2:51:29	Connection Server	User PRIVATECLOUD\vice.president requested Pool Scenario_1,
PRIVATECLOUD\vice.president	Info	7/22/11 2:51:29	Connection Server	User PRIVATECLOUD\vice.president requested Pool Scenario_1
PRIVATECLOUD\vice.president	Audit success	7/22/11 2:51:08	Connection Server	User PRIVATECLOUD\vice.president has logged in

Figure 5. Event Monitor Log

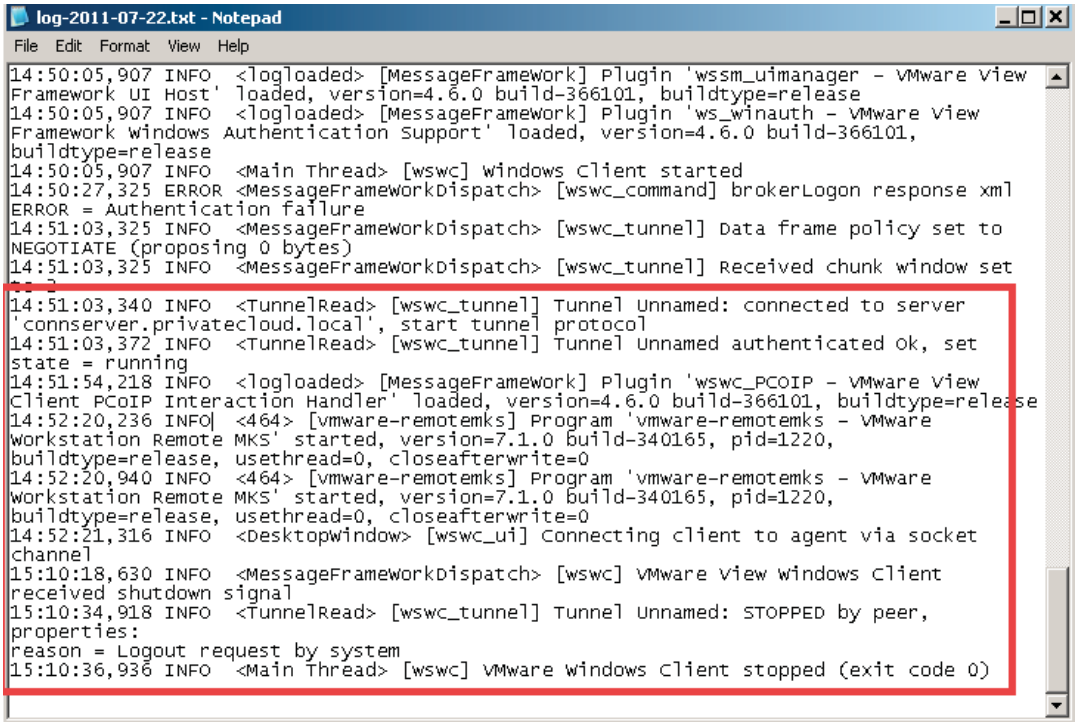

Figure 6. View Client Log

Conclusion to case scenario 1

This scenario was successfully investigated using current digital forensic procedures . The forensic team was successfully able to investigate because the user and his environment were known. Secondly, the suspect's VM was configured with persistent disk mode, which left the internet temporary files behind. Figure 7 shows that all standard investigation techniques can be followed.

Investigation Process	
Preparation	✓
Search	✓
Recognition	✓
Collection	✓
Investigation	✓
Examination	✓
Analysis	✓
Presentation	✓
Documentation	✓

Figure 7: Summary Of Case Scenario 1

(b) Scenario 2

Aim: Locate the virtual desktop which attempted to browse objectionable images.

GENERAL FINDINGS

- The ISP has leased 3 IP addresses to the school which have been further distributed, so that every computer can been uniquely identified on the school's network.
- The school has a hosted virtual desktop infrastructure in place.
- All the VM's are distributed among three virtual servers i.e. Teachers, Students, Staff. Each virtual server is using 1 of 3 ISP provided IPs.
- All virtual desktops are part of a dedicated pool and every VM is configured for persistent disk mode.
- The school's internet filtering system is very basic and did not detect attempts made to connect to the external network containing objectionable material.
- The school had no WIFI to restrict students using smart devices during school hours

FORENSIC PROCESS

The potential locations of the evidence can be either in the virtual desktop or in backups which are performed every day at 12:00 midnight. Possible traces that may show a user's attempt to connect to the offshore network can be found in the browser cache, internet history or DNS cache. But in this scenario, the forensic team only has reports provided by the ISP including address of the offshore network with timestamps confirming the activity occurred via the IP range belonging to the school. With this information, it is hard to narrow down to a user, as this could potentially be any user or group of users within the school since the IP range is using NAT to accommodate every computer in the campus. It is also possible that a rogue user may be accessing the school's network to gain access to the external network containing objectionable material.

Possible ways to start search and seizure are:-

1. Create a list of keywords from information provided by the ISP and perform a search using the keywords on every virtual server i.e. Teachers, Staff, Student, assuming the search will go through the file system of every VM present on the server. Possible results can be internet temp files or visited URL's. This may mean taking the servers offline but the biggest challenge is to perform a search on the entire server containing numerous VM's. Not only is this time consuming but also the ability to perform keyword search in such a manner is very restricted unless supported by third party tools. Currently none of the forensic suites is able to read, analyze or perform such search on ESX file system (VMFS) (Haletky, 2011). This technique also voids best practice as the search will be performed on a live system. The advantage of using this technique is that it avoids intrusion into a user's personal space as the search will only show results based on the information provided in response to the key words.
2. Attempt to identify the user rather than looking for traces. i.e. examine the record of previously occurred incidents involving the school's computers and internet. This depends on the school's records, and if such campus incidents are logged it will help to narrow down to the number of users who may be involved.
3. Audit firewall logs. Depending on the type of firewall, it is possible the logs show visited URLs with timestamps and user/s.
4. Seize and image every virtual server in order to import them into a forensic software for analysis. The time taken to image every virtual server will result in downtime of school work as all the academic application that are served to users via their virtual desktops. More importantly, due to the nature of the analysis stage the likelihood of browsing through other user's personal space is very high as multiple tenants are hosted on the same server.

This scenario is set in a school context. In reality, a school is likely to have an internet traffic monitoring system that will detect such activity. But if a similar incident occurs in an organization, the ability to identify a single user can be very complex. An Internet user does not have a unique identity, in contrast to an internal network. The primary concern of this scenario is respecting privacy of other users' personal space during the investigation. In cases where techniques 1, 2, 3 are not practical, investigators are left with option four, where data of all users is forcefully exposed rather than specific users' data. Trawling through the resulting data can be very difficult without knowing what to look for. This clearly shows typical forensic procedures fail on multi-tenant environments because the industry lacks tools and procedures to segregate forensic data among multiple tenants (Ruan & Carthy, 2011). Figure 8 summarises the investigation process.

Investigation Process	
Preparation	
Search	X
Recognition	X
Collection	X
Investigation	
Examination	X
Analysis	X
Presentation	
Documentation	X

Figure 8. Summary Of Case Scenario 2

CONCLUSION

By simulating a working HVD model various case scenarios were designed, developed and implemented. This permitted us to observe the adequacy of current digital forensic procedures in such cases. The findings show that locating and extracting evidence in scenario 1 was possible but not straightforward. Scenario 2 proved to be too complex for current digital procedures to handle easily, with neither location nor extraction of evidence being possible. This problem is not caused by the actual virtualization suite used but is inherent in the simulated cloud architecture.

In other words, the differences between the outcomes of the two case scenarios are the result of the two independently configured cloud computing models. In the case of scenario 1, the cloud computing model consists of VM's configured with persistent virtual HDD, whereas in the case of scenario 2 the cloud computing model is multi-tenanted. While these results are preliminary and more work is required to investigate the implications of multi-tenanted architectures and non-persistent environment for forensic IT procedures, it is clear that using a multi-tenanted architecture could have digital forensic consequences for protecting the privacy of individuals not involved in any objectionable behaviour. In other words, it is possible that companies and organisations adopting multi-tenanted architectures may need to get the written agreement of employees and users to their private data being in the scope of digital forensic investigations even if they are not suspected of doing anything wrong. Likewise, the non-persistent (stateless) environments may have advantages, yet it affects availability of potential evidence.

In conclusion, these experiments indicate that organisations considering the deployment of HVDs will need to review their digital forensic preparedness to ensure that their auditing and investigative procedures are placed on as sound a footing as in non-HVD environments.

REFERENCES

Avison, D. E., Lau, F., Myers, M. D., & Nielsen, P. A. (1999). Action research. Communications of the ACM, 42(1), 94-97. doi:10.1145/291469.291479

Barrett, D., & Kipper, G. (2010). Virtualization and Forensics

Dodge, J. (2006). VirtualCenter 2: Template Usage and Best Practices: Foedus. Retrieved from

http://www.vmware.com/pdf/vc_2_templates_usage_best_practices_wp.pdf

Gartner. (2009). Gartner says Worldwide Hosted Virtual Desktop Market to Surpass $65 Billion in 2013. Retrieved from http://www.gartner.com/it/page.jsp?id=920814

Haletky, E. L. (07/09/2011). Retrieved from http://communities.vmware.com/thread/152476

Henry, P. (2009). Best Practices In Digital Evidence Collection. Retrieved 18/09/2011, 2011, from http://computer-forensics.sans.org/blog/2009/09/12/best-practices-in-digital-evidence-collection/

Ruan, K., & Carthy, J. (2011). Cloud forensics: An overview.

Taylor, M., Haggerty, J., Gresty, D., & Hegarty, R. (2010). Digital evidence in cloud computing systems. Computer Law & Security Review, 26(3), 304-308. doi:10.1016/j.clsr.2010.03.002