Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2015, Article ID 689870, 10 pages
http://dx.doi.org/10.1155/2015/689870

Research Article

Investigation Methodology of a Virtual Desktop Infrastructure for IoT

Doowon Jeong,[1] Jungheum Park,[1] Sangjin Lee,[1] and Chulhoon Kang[2]

[1] *Center for Information Security Technologies (CIST), Korea University, Anam-dong, Seongbuk-gu, Seoul 136-713, Republic of Korea*
[2] *Supreme Prosecutors' Office, Seocho-dong, Seocho-gu, Seoul 137-730, Republic of Korea*

Correspondence should be addressed to Sangjin Lee; sangjin@korea.ac.kr

Received 13 March 2014; Accepted 31 July 2014

Academic Editor: Young-Sik Jeong

Cloud computing for IoT (Internet of Things) has exhibited the greatest growth in the IT market in the recent past and this trend is expected to continue. Many companies are adopting a virtual desktop infrastructure (VDI) for private cloud computing to reduce costs and enhance the efficiency of their servers. As a VDI is widely used, threats of cyber terror and invasion are also increasing. To minimize the damage, response procedure for cyber intrusion on a VDI should be systematized. Therefore, we propose an investigation methodology for VDI solutions in this paper. Here we focus on a virtual desktop infrastructure and introduce various desktop virtualization solutions that are widely used, such as VMware, Citrix, and Microsoft. In addition, we verify the integrity of the data acquired in order that the result of our proposed methodology is acceptable as evidence in a court of law. During the experiment, we observed an error: one of the commonly used digital forensic tools failed to mount a dynamically allocated virtual disk properly.

1. Introduction

In the recent past, cloud computing has experienced phenomenal growth for IoT (Internet of Things). To offer IoT services, many companies have managed to reduce costs and enhance the efficiency of their servers by adopting a virtual desktop infrastructure (VDI) which is classified into private cloud computing. Private cloud computing involves the use of virtualization technology of cloud servers. Resources such as CPU, RAM, and server storage are shared. Unlike a public cloud, the servers are only used by internal users. The use of private cloud computing is continually increasing owing to its efficiency.

However, as VDI is widely used, threats of cyber terror and invasion are also increasing. In VDI, all resources are shared; it would be critical to whole users. To minimize the damage, response procedure such as investigating causal relationship and identifying a criminal on a VDI should be systematized either scientifically or technically. However, investigation methodology for private clouds are not keeping pace with this growth in private cloud computing, despite much research into investigation and digital forensics for cloud computing. Taylor et al. outlined challenges and considerations relevant to examiners when investigating general cloud computing environments [1]. Chung et al. proposed a procedure for investigating and analyzing artifacts for users of cloud storage services [2]. Dykstra and Sherman researched a forensic collection method for infrastructure-as-a-service cloud computing [3]. However, to the best of our knowledge, research on digital forensic investigation (DFI) for a complete VDI has yet to be accomplished. Other research into digital forensics for cloud computing tends to focus on concepts or processes for general investigation and evidence collection. Therefore, more research into DFI for VDI is necessary.

In cloud-hosted virtual desktop environments, user data may not be stored on the local system but in distributed storage linked by a hypervisor, unlike noncloud-hosted virtual desktop environments [4–6]. An investigation of a computer requires an image of the entire target device [7]. However, this is becoming increasingly impractical because storage

Table 1: Hypervisor and desktop virtualization solutions.

Component	Citrix	VMware	Microsoft
Hypervisor	XenServer 6.0	ESXi Server 5.0	Hyper-V (Windows Server 2008 R2)
Hypervisor management system	XenCenter 5.6	View 5.0	Hyper-V (Windows Server 2008 R2)

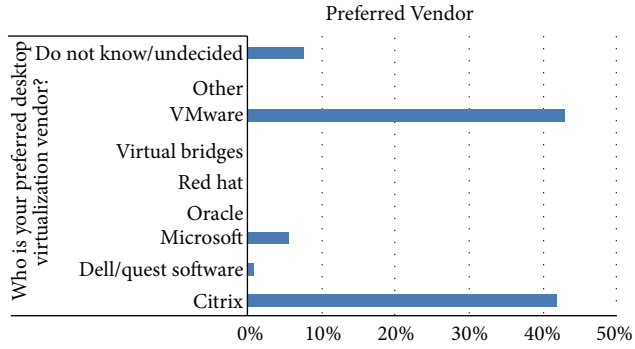

Figure 1: Desktop virtualization trends at the Gartner Data Center [12].

can contain many terabytes of data. Partial or selective file copying such as a virtual hard disk for a specific user may be considered for DFI in a cloud computing environment [8–11]. Therefore, we believe that this new approach will be very useful for investigating crimes and causal relationship related to VDI invasion accident.

The remainder of the paper is organized as follows. In Section 2, we present VDI for IoT and briefly introduce popular desktop virtualization solutions, such as VMware, Citrix, and Microsoft. In Section 3, we propose a DFI method that searches for user traces, assigns information between a user and a virtual machine, and collects data using structural features and functions of each desktop virtualization solution. In Section 4, we verify the integrity of VDI data acquisition in an experiment. In Section 5, we report an error identified during this experiment: Encase, a widely known digital forensic tool, failed to mount a dynamically allocated virtual disk properly. Section 6 concludes.

2. Virtual Desktop Infrastructure

2.1. Desktop Virtualization Solutions. In computing, virtualization is a technique for sharing resources such as hardware platforms, operating systems, storage, and network devices [13, 14]. Desktop virtualization involves separating the logical desktop from the physical server, which is realized by a hypervisor. A hypervisor is a logical platform for simultaneous operation of multiple operating systems on a host server. VDI is a desktop-centered service that hosts user desktop environments on remote servers and/or blade PCs; the hosts can access VDI over a network using a remote display protocol. Desktop virtualization solutions are software packages consisting of several programs, and these solutions are based on the hypervisor. There are various desktop virtualization solutions; Citrix, VMware, and Microsoft are the most popular (Figure 1). Therefore, we focused on these three solutions here. Each solution has its own hypervisor: Citrix uses XenServer, VMware uses ESX/ESXi Server, and Microsoft uses Hyper-V. Here, we construct a VDI that consists of a hypervisor and a desktop virtualization solution. Table 1 lists the hypervisor versions and desktop virtualization solutions we used in the study.

2.2. VDI Structure. Although the hypervisor and desktop virtualization solution comprising each VDI differ, a survey revealed that the configuration methods are very similar [15–17] (Table 2). As shown in Figure 2, a hypervisor and hypervisor management system are required to create and manage virtual machines. A local storage device such as the hard disk of a hypervisor system can be used as a storage unit for the virtual machine. However, in the majority of cases, shared storage devices are used because companies require many virtual machines to offer private cloud computing services to their members. An authentication management system and a connection management system are also essential for user authentication and delivery of a virtual machine to the user. A user can access the virtual machine using a specific program or web once the VDI is constructed. The access process for the virtual machine is as follows (Figure 2).

(1) A connect request (login) is sent to the connection management system.

(2) The connection management system sends the user login information to the authentication management system.

(3) On successful user authentication, the connection management system asks the hypervisor to assign a virtual machine, which is stored in the shared storage.

(4) The connection management system delivers that virtual machine to the user.

(5) Then, the virtual machine can be used as a personal desktop.

Table 2: VDI components.

Component	Citrix	VMware	Microsoft	Role
Hypervisor	XenServer	ESXi server	Hyper-V	Create and manage virtual machines
Hypervisor management system	XenCenter	vCenter server	SCVMM (system center virtual machine manager)	Manage the hypervisor
Connection management system	DDC (desktop delivery controller)	View Manager	RDCM (remote desktop connection manager)	Connect and assign a virtual machine to a user
Authentication management system	Active Directory	Active Directory	Active Directory	Register (create/delete) and authenticate the user
Virtual machine access program	Web browser (Citrix receiver should be installed)	View client or web browser	Web browser	Access to virtual machine

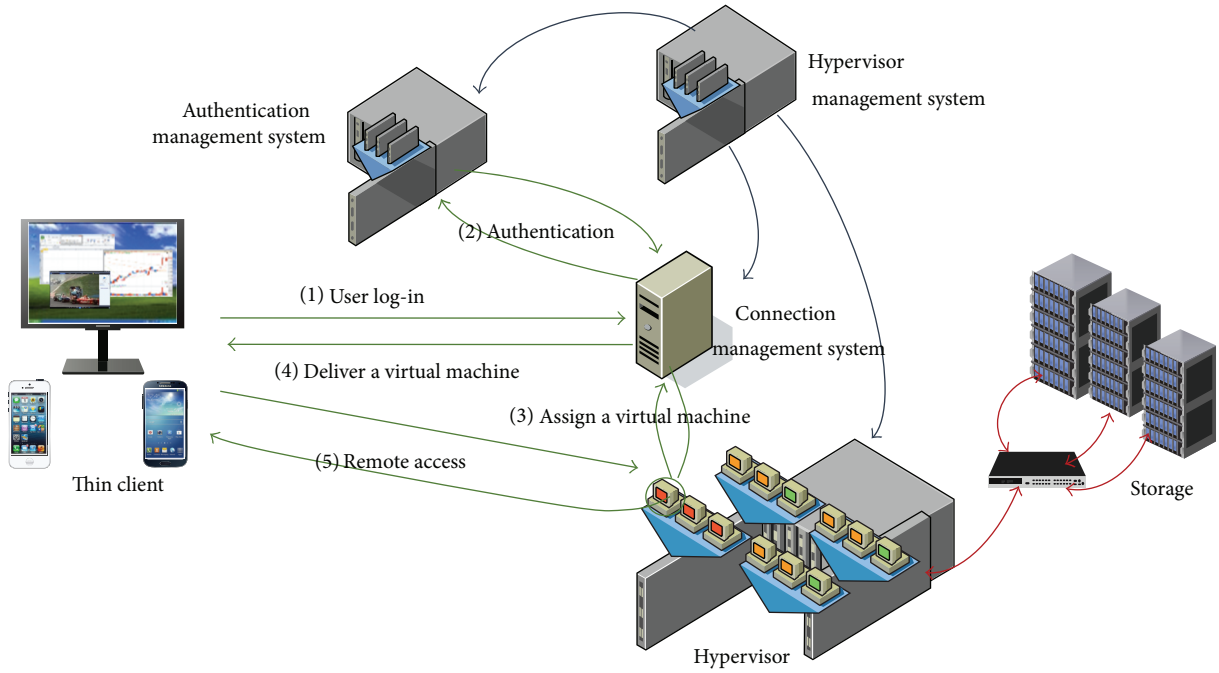

Figure 2: General VDI structure.

3. DFI Method for VDI

In VDI, user data are stored in the central storage for virtual machines. There are two methods for gathering a user's data: one is to investigate the entire central storage, and the other is to remotely extract the virtual machine allocated to that user. The first method is inefficient because the central storage capacity is huge and so investigation is very time consuming. Therefore, the second method is preferable because it is similar to disk imaging for investigation of the hard disk of a local desktop. Hence, extraction of a virtual machine is the main point for investigating a VDI. To achieve this, an investigator must determine whether or not the suspect uses a particular virtual machine.

DFI for VDI targets systems that carry user traces. The trace recorded by a system is used to access the virtual machine. To find the trace, the first step is to investigate the thin client for a user using the virtual desktop as in Figure 3. When a user accesses a virtual machine, access information such as registry data, log files, or web history is recorded in the thin client and can be discovered via a signature search, depending on the solution. However, if this information cannot be uncovered (e.g., the records have been deleted and the programs have been removed), it is difficult to obtain virtual machine access information from the thin client. In this case, the investigator only needs to check the user access information and virtual machine assignment information in the connection management system and the authentication management system.

After inspecting the relevant virtual machine access information, the investigator should collect data for the virtual machine used by the suspect. For this, the investigator requires administrator authority for the hypervisor

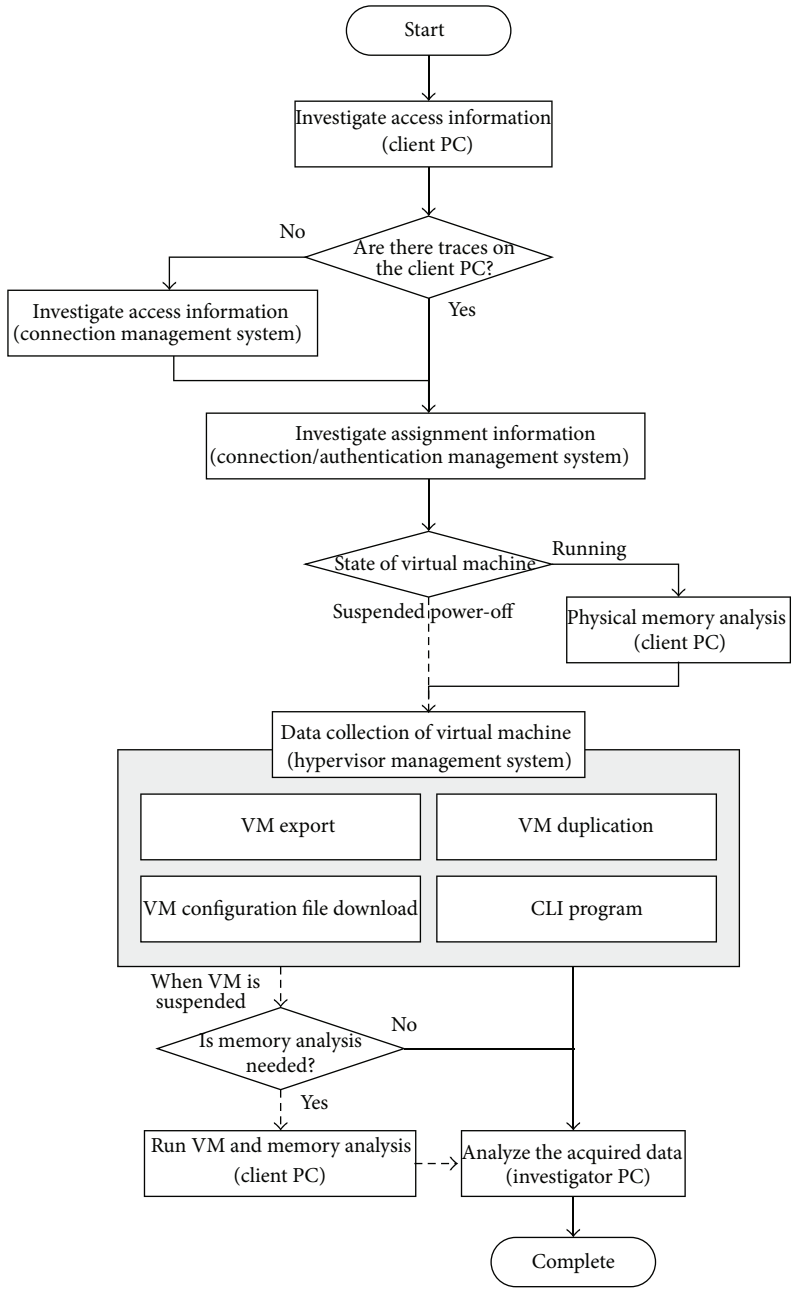

Figure 3: Digital forensics procedure for VDI in private cloud computing.

or its management system or user authority for the virtual machine. If access authorities are obtained, then the data can be collected via the hypervisor management system, shell connection, or virtual machine access. Data collection via the hypervisor management system or shell connection requires a dedicated program for each solution. If the virtual machine is already running, the investigator can analyze live memory and perform a memory dump by executing memory forensics tools in the virtual machine. Detailed information is presented in Section 3.3. The collected data can then be analyzed using general DFI methods and tools.

Here, we make two assumptions: (i) the investigator already knows the suspects, because private cloud computing services are provided to restricted users who have access authority; and (ii) the investigator has administrator or user authority with assistance from the organization.

3.1. User Access Information. As mentioned above, the VDI structure of Citrix, VMware, and Microsoft is very similar. Therefore, the DFI method is similar to these solutions. Evidence of use of a virtual machine is logged in the user's computer, hypervisor management system, connection management system, and authentication management system. Here, a DFI method for a general VDI using Citrix, VMware, and Microsoft and local computers operating on Windows 7, Ubuntu 12.04, and Mac OS 10.8.2 is studied.

Table 3: Access information for a virtual machine logged in the local Windows system.

Solution	Registry	Log/web browser signature
Citrix	*KEY_CURRENT_USER\Software\Citrix\XenDesktop\DesktopViewer\[VM name]* ⇒ VM name, IP address of connection management system (DDC)	*%UserProfile%\AppData\Roaming\ICAClient* ⇒ VM name, connection/disconnection time Signature: DesktopWeb ⇒ connection time, IP address or name of connection management system (DDC)
VMware	*HKEY_CURRENT_USER\Software\VMware, Inc.\VMware VDM\Client* ⇒ VM name, IP address or URL of connection management system (View Manager), domain name, user computer name	*%UserProfile%\AppData\Local\VMware\VDM\logs* ⇒ URL of connection management system (View Manager), connection/disconnection time, domain name, user computer name ※ log-[yyyy]-[mm]-[dd].txt
Microsoft	*KEY_CURRENT_USER\Software\Microsoft\Terminal Server Client\Default* ⇒ VM name or IP address	*Signature: RDWeb* ⇒ connection time, Hyper-V server name, domain name

3.1.1. Traces on the Client PC. In Windows 7, registry and log entries are created when VMware is used. When Citrix is used, registry, log entries, and web history traces are created. When a Microsoft VDI is used, registry entries related to the remote desktop are created, but log entries are not created. However, Microsoft uses a specific signature, RDWeb, when a connection using the web is made to a virtual desktop environment. Therefore, access information can be determined from the web history. Table 3 shows the access information for a virtual desktop environment logged in the local Windows system.

In Ubuntu 12.04 and Mac OS 10.8.2, access information for VMware can be found from the log created as in Windows OS. However, for Citrix, when a connection is made via a web browser, an investigator should check the history of the web browser. Thus, we studied Firefox, the default web browser in Ubuntu, and Safari, the default web browser in Mac OS. Further, for Microsoft, unlike in Windows, access information cannot be found via web history analysis since an RDWeb connection is impossible using web browser in Ubuntu and Mac OS. Instead, the access information can be found from the information retained when a remote desktop connection from each OS to the Microsoft virtual machine is made. Table 4 shows the access information logged in local Ubuntu and Mac systems.

3.1.2. Traces on the Connection Management System. If there are no connection traces on the user's local computer, the investigator should focus on the connection management system, which assigns virtual machines to users, manages the machines, and connects or disconnects virtual machines according to user requests. Therefore, all information pertaining to connections to virtual machines is managed and logged here. An investigator can find information on the exact time at which a user connected to or disconnected from the virtual machine by analyzing these log files. Table 5 shows the access information logged in the connection management system.

3.2. Virtual Machine Assignment Information. To connect to a virtual machine, a user must be assigned a virtual machine through the connection management system. A virtual machine assigned to a specific user cannot be accessed by others and will be used only by that user. The assignment information is stored in the connection management system and authentication management system. It is useful to prove the relationship between a suspect and a virtual machine. The assignment information in the connection management system should be investigated to establish connection information between the virtual machine and its user. Table 6 shows how to find this assignment information in the connection management system. Assignment information is also stored in the database of the connection management system or authentication management system. Table 7 summarizes the method for finding assignment information between a user and a virtual machine from the database.

3.3. Data Collection for a Virtual Machine. In a virtual desktop environment, data for a virtual machine are stored in the storage area for the server and not on the local computer. Therefore, an investigator should investigate the central storage area. However, when a cloud environment is constructed, the central storage area is typically made up of multiple independent storage devices [18]. It is not feasible to collect all the data from these devices. Thus, it is most efficient to acquire a virtual hard disk for the virtual machine. However, it is difficult to acquire data for a virtual machine because the virtual hard disk can be allocated in various ways: as single or multiple files and via static or dynamic allocation. The data could be stored on one physical disk or distributed over multiple disks. Therefore, we use the hypervisor management system and shell connection program for each solution to acquire a virtual hard disk for the suspect because data extraction is possible without reference to the type of allocation. If a user is connected to a virtual machine, the investigator can collect data such as a memory dump, specific files, or the entire virtual hard disk of the virtual machine.

Table 4: Access information logged in local Ubuntu and Mac systems.

Solution	Ubuntu 12.04	Mac OS 10.8.2
Citrix	Cache: *home**[user name]*\.mozilla\firefox\6lhwv183.default\Cache_CACHE_[numbers]_ History: *home**[user name]*\.mozilla\firefox\6lhwv183.default\places.sqlite Cookie: *home**[user name]*\.mozilla\firefox\6lhwv183.default\cookies.sqlite Session: *home**[user name]*\.mozilla\firefox\6lhwv183.default\sessionstore.js ⇒ IP address or URL of connection management system (DDC)	Cache: *Users**[user name]*\Library\Caches\com.apple.Safari\Cache.db History: *Users**[user name]*\Library\Safari\History.plist Cookie: *Users**[user name]*\Library\Safari\Cookies.plist Session: *Users**[user name]*\Library\Safari\LastSession.plist ⇒ IP address or URL of connection management system (DDC)
VMware	*tmp*\vmware-*[user name]*\vmware-view-*[numbers]*.logs ⇒ IP address or URL of connection management system (View Manager), connection/disconnection time, user ID, VM name, domain name	*Users**[user name]*\Library\Logs\VMware View Client\vmware-view.logs ⇒ IP address or URL of connection management system (View Manager), connection/disconnection time, VM IP address, domain name
Microsoft	*home**[user name]*\.bash_history ⇒ VM name or IP address, user ID (option), user password (option), domain name (option)	*Users**[user name]*\Documents\RDC Connections\Default.rdp ⇒ VM name, user ID, domain name

Table 5: Access information logged in the connection management system.

Solution	Log
Citrix	*%SystemDrive%\inetpub\logs\LogFiles\[folder name]* ⇒ connection/disconnection time, connection management system (DDC) and user IP address ※ [yymmdd].log
VMware	*%SystemDrive%\ProgramData\VMware\VDM\logs* ⇒ VM name and IP address, connection/disconnection/reconnection/logoff time, domain name, user computer name ※ log-[yyyy]-[mm]-[dd].txt
Microsoft	*%SystemDrive%\inetpub\logs\LogFiles* ⇒ connection/disconnection time, user ID ※ [yymmdd].log

Table 6: Method for finding assignment information in the connection management system.

Solution	Method
Citrix	DDC (1) Start Citrix Desktop Studio on DDC (2) Select Desktop Studio-Assignments (3) Select VM or Group
VMware	View Manager (1) Start View Administrator Console on View Manager (2) Select Inventory-Desktops
Microsoft	Active Directory (1) Start Active Directory user and computer on Active Directory (2) Select user-properties—personnel virtual desktop

3.3.1. Hypervisor Management System. A target virtual machine can be exported or duplicated and the component files can be downloaded using the hypervisor management system provided by each solution. Table 8 summarizes methods for collecting virtual machine data using the hypervisor management system.

When using VM export, the virtual machine data are converted to the solution format (e.g., xva file format for Citrix). VM duplication means that the raw data for the virtual machine can be obtained. In the case of VMware, we can select and download some configuration files using the VM configuration file download method.

3.3.2. Shell Connection Program. Each solution provides a command-line interface (CLI) with various administrative and management-oriented utilities. One such utility provided by each solution allows acquisition of a copy of the state of the virtual machine. VMware and Microsoft can collect the raw data duplicated from the original virtual disk. Citrix, however, can only collect compressed data. Thus, XenCenter is required to recover and analyze virtual machine data hosted and acquired via Citrix. Table 9 summarizes the method for

Table 7: Method of finding assignment information in the database of the connection or authentication management system.

Solution	Method
Citrix	DDC (1) Connect to DB by using MS SQL Server Management Studio (2) [*DDC PC name*]-[*Databases*]-[*CitrixXenDesktopDB*]-[*Tables*]-[*chb_Stat e.AccountNames*]: user name and Uid (3) [*DDC PC name*]-[*Databases*]-[*CitrixXenDesktopDB*]-[*Tables*]-[*chb_Stat e.WorkerDiags*]: VM assigned user (Uid)
VMware	View Manger (ADAM DB) (1) Connect to ADAM DB by using Active Directory Explorer (2) [*DC=vdi,DC=vmware,DC=int*]-[*OU=Servers*]: specific VM CN (Common Name) value and other information (a) Description: VM name (b) Member: user CN (3) [*DC=vdi,DC=vmware,DC=int*]-[*CN=ForeignSecurityPrinciple*]: user CN value and other information (a) Description: user and domain name
Microsoft	Active Directory (ADAM DB) (1) Connect to ADAM DB by using Active Directory Explorer (2) [*DC=domain name*]-[*OU=Hyper-V*]: user name and other information (a) msTSPrimaryDesktop: assigned VM name

Table 8: Data acquisition method using the hypervisor management system.

Solution	VM export	VM duplication	VM configuration file download
Citrix (XenCenter)	Select VM-Menu-VM-Export ⇒.xva or.ovf file Export	Select VM-click mouse right button-Copy VM-Full copy	
VMware (vCenter)	Select VM-Menu-File-Export-OVF Template Export ⇒.ovf file Export	Select VM-duplication	Select Hypervisor or VM-Summary-Resource-Storage-select Datastore-Browse Datastore-select folder or file-download
Microsoft (Hyper-V Manager and SCVMM)	Hyper-V Manager-select VM-click mouse right button-Export ⇒.vhd file Export	SCVMM-select VM-duplication-deploy VM on host	

collecting virtual machine data using the shell connection program.

3.3.3. Consideration of the State of a Virtual Machine. In a virtual desktop environment, a virtual machine can be running, suspended, or in a power-off state. An investigator should check the state of a virtual machine before acquiring data, because the acquisition method that is applicable varies, depending on the state. Table 10 lists applicable acquisition methods. It is evident that when the virtual machine is running, it is impossible to acquire the virtual disk using the Citrix and Microsoft solutions. For the Microsoft solution, the investigator should turn off the virtual machine. If analysis of the memory is essential, the investigator should analyze the memory before turning off or suspending the virtual machine. For analysis of the memory when the virtual machine is in a suspended state, the investigator should first acquire the virtual disk and then resume the virtual machine for memory forensics.

4. Verification of Acquisition Data Integrity

The integrity of the acquired data should be demonstrated for admissibility of evidence in a court of law. Hence, in this section, we verify the integrity of the virtual hard disk drive (HDD) acquired according to our method.

4.1. Experiment #1: Comparison of Hash Values for the Original Virtual HDD and the Acquisition Data. Several methods can be used to acquire a virtual hard disk. In VMware, acquisition is via a shell connection program and VM export, duplication, and file download through the hypervisor management system. As Microsoft and Citrix do not provide VM file download, we acquire data via a shell connection program and VM export or duplication through the hypervisor management system. After acquiring the data, we compared hash values for the original virtual HDD of the virtual machine and the acquisition data. Table 11 lists the results.

For VMware and Microsoft, the hash values match perfectly, regardless of the acquisition method used. The sizes of the original virtual HDD and acquisition data are also the same. Therefore, investigation using VMware or Microsoft according to the proposed acquisition method yields that data are admissible as evidence in a court of law.

However, for Citrix, the hash values are different. First, there is a difference between the format of the original virtual HDD data and the acquisition data. The format of the original

Table 9: Acquisition of virtual machine data using a hypervisor CLI with default utilities for each solution.

Solution	Shell connection program
Citrix (XenCenter console Tab)	Connect to shell or select "Console" tab on XenCenter Virtual disk collection: xe vm-export vm=[*VM name*] filename=[*file mane*].xva
VMware (vSphere PowerCLI)	Connect to shell using vSphere PowerCLI Virtual disk collection command: copy-datastoreitem [*datastore drive*]:\[*Src. path*] [*Dst. path*] ∗vSphere PowerCLI should be installed
Microsoft (Windows PowerShell)	Connect to shell using Windows PowerShell Virtual disk collection command: export-vm-vm "[*VM name*]"-server [*Hyper-V Server name*]-path [*Dst. path*] ∗PowerShell Management Library for Hyper-V should be installed

Table 10: Applicable acquisition method depending on the solution and state of the virtual machine.

Solution	Acquisition method	State		
		Running	Suspended	Power-off
Citrix	VM export	No	Yes	Yes
	VM duplication	No	Yes	Yes
	VM configuration file download	No	No	No
	CLI program	No	Yes	Yes
VMware	VM export	No	No	Yes
	VM duplication	Yes	Yes	Yes
	VM configuration file download	No	Yes	Yes
	CLI program	No	Yes	Yes
Microsoft	VM Export	No	No	Yes
	VM duplication	No	No	Yes
	VM configuration file download	No	No	No
	CLI program	No	No	Yes

Table 11: Results for experiment #1 on integrity verification.

Solution	Acquisition method	Hash value		Result
		Original virtual HDD	Acquisition data	
VMware	VM export	0440B1A068A0A9D116B2184E824196D7	0440B1A068A0A9D116B2184E824196D7	Match
	VM duplication		0440B1A068A0A9D116B2184E824196D7	Match
	VM file download		0440B1A068A0A9D116B2184E824196D7	Match
	CLI program		0440B1A068A0A9D116B2184E824196D7	Match
Citrix	VM export	CEDB64BD9510566BD3A7A516CADF6444 (Disk size: 5,309,903,360 bytes)	06D6A00AD0A51EFE1E31B04B0D473BE2 (Disk size: 5,200,160,256 bytes)	**Mismatch**
	VM duplication		06D6A00AD0A51EFE1E31B04B0D473BE2 (Disk size: 5,200,160,256 bytes)	**Mismatch**
	CLI program		06D6A00AD0A51EFE1E31B04B0D473BE2 (Disk size: 5,200,160,256 bytes)	**Mismatch**
Microsoft	VM export	328D07681CD90C98BB71F625F47B3F07	328D07681CD90C98BB71F625F47B3F07	Match
	VM duplication		328D07681CD90C98BB71F625F47B3F07	Match
	CLI program		328D07681CD90C98BB71F625F47B3F07	Match

data is VHD, but that of the acquisition data is XVA or OVF and the data are compressed. Decompression of an acquisition file leads to a smaller size than of the original. This is because Citrix rearranges the original data when the data are acquired via XenCenter. Figure 4 shows that the offset of a specific file is changed from 0x10CFFF to 0x10C800.

Repetition of the experiment revealed that when data are acquired or duplicated using XenCenter, they are transmitted via blocks and the transmitted data are rearranged. It is impossible to verify the integrity of the original virtual HDD and the acquisition data by comparing hash values because the data order is inverted when the original HDD is acquired.

Table 12: Results for experiment #2 on integrity verification for logical drives for Citrix.

Area	Original HDD	Acquisition data	Result
Boot	092D9487556456C6881F16BEA9FABCDA	092D9487556456C6881F16BEA9FABCDA	Match
Data	27A83C3709DEE6F042AA064C56B7DE29	27A83C3709DEE6F042AA064C56B7DE29	Match

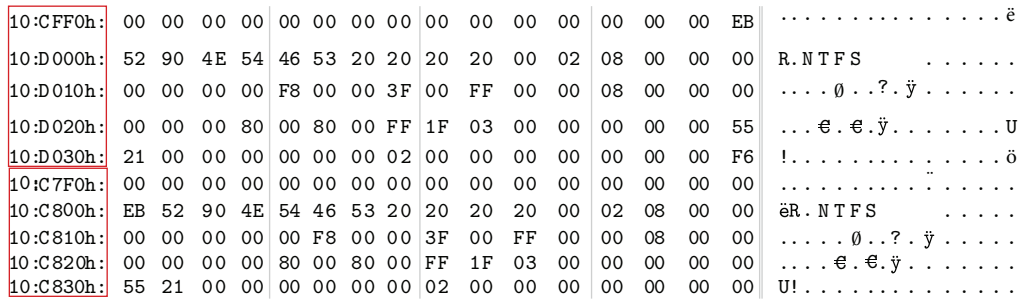

Figure 4: Comparison of offsets for the same file: top, original HDD; bottom, acquisition data.

4.2. Experiment #2: Comparison of Hash Values for Logical Drives. The integrity of Citrix acquisition data was verified in a different manner. We mounted the original HDD and the acquisition data on a local computer to verify the integrity. The hash value for each logical drive was then calculated. Various tools were used to enhance the reliability of the experimental results. The tools Mount Image Pro, FTK Imager, and X-Way Forensics were used for mounting the disk image, and Encase, FTK Imager, and X-Way Forensics were used for calculating hash values. The reason why Encase was not used for mounting is explained in Section 5. Table 12 lists the results for these experiments.

Table 12 reveals that the size and hash values match for the original HDD and the acquisition data. We also verified the integrity of the acquisition data by comparison of hash values for each mounted logical drive. The results for experiments #1 and #2 prove that the proposed acquisition method ensures data integrity.

5. Reliability Verification of Forensic Tools for Virtual Machine Data

During experiment #2, we found that Encase 6 and Encase 7 could not parse the acquired data in their entirety when mounting the virtual HDD, which is dynamically allocated in VHD format. This problem was observed both for data acquired through Citrix and for the Microsoft solution. To explore this problem further, we compared various tools. Table 13 shows the ability of each tool to correctly parse the acquired dynamic VHD formats.

Encase failed to properly mount the original virtual HDD as well as the copy. To understand the reason behind this problem, we calculated the hash values for all the entries for virtual drives mounted by Encase, FTK Imager, and X-Way Forensics. There were 59,127 entries and the hash values for 13 of these entries were mismatched.

To analyze this issue in detail, we compared the mismatched files using a hex editor. As observed in Figure 5, the hex values are different even though they are at the same offset in the same file (pagefile.sys). We found that unknown values were repeatedly written at a specific offset for some files, but the reason why these are written when Encase mounts a dynamic VHD format remains unknown.

This finding indicates that an investigator should avoid Encase when mounting acquired data in a dynamic VHD format. However, Encase may be used to analyze the data after mounting via some other tool.

6. Conclusion

Adoption of a VDI for IoT can save costs and is a convenient alternative for users. However, investigation methods for VDI invasion accidents have not kept pace with the VDI market, which is rapidly growing and experiencing wide development.

Here, we explained VDI and popular VMware, Citrix, and Microsoft desktop virtualization solutions. The infrastructure of the three solutions is very similar, so we were able to establish a framework for VDI investigation. Since VDI is different from general PC environments, we focused on acquiring the data for a virtual machine using user access information from the PC thin client, the connection management system, and the authentication management system. By applying the proposed method to VDI, an investigator can obtain a virtual disk image and analyze this as for general disk forensics. We verified the integrity of data acquired via our method through experiments for admissibility of evidence in a court of law. Moreover, we discovered that a widely used tool has an error and failed to properly mount acquired data in a dynamic VHD format.

This paper will be useful for investigation of cases in which VDI plays an essential role. We hope that it will inspire further research on DFI methods in response to the rapidly growing cloud computing environment.

Table 13: Comparison of hash values for various tools.

Index	Original virtual hard disk	Copy of virtual hard disk	Result
EnCase	**C69289228xxxxxx**	**64C4D1298xxxxxx**	**Mismatch**
FTK	C5F64F49Cxxxxxx	C5F64F49Cxxxxxx	Match
X-Way Forensics	C5F64F49Cxxxxxx	C5F64F49Cxxxxxx	Match

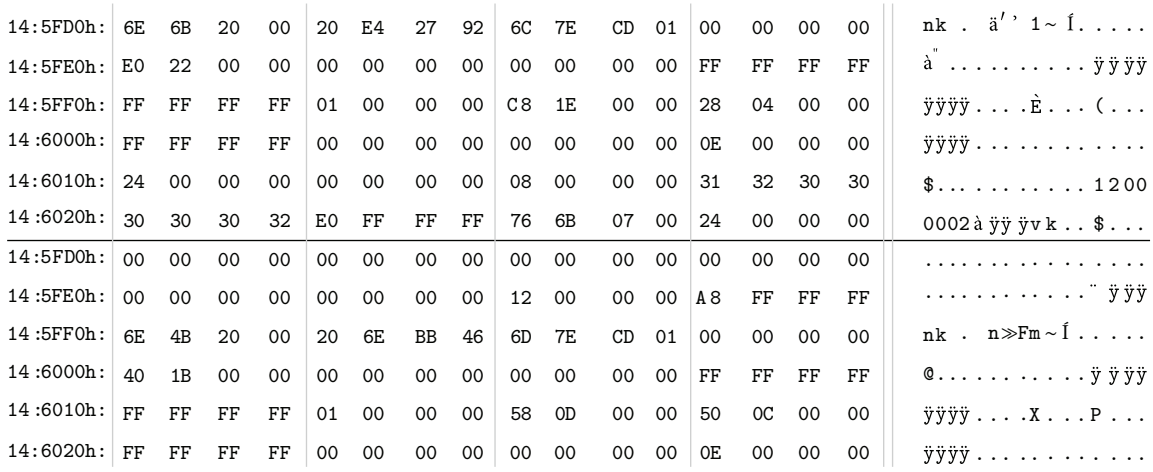

Figure 5: Image of pagefile.sys hex values while mounting a virtual disk using Encase (top) and FTK Imager (bottom).

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This research was supported by the Public Welfare & Safety Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (2012M3A2A1051106).

References

[1] M. Taylor, J. Haggerty, D. Gresty, and D. Lamb, "Forensic investigation of cloud computing systems," *Network Security*, vol. 2011, no. 3, pp. 4–10, 2011.

[2] H. Chung, J. Park, S. Lee, and C. Kang, "Digital forensic investigation of cloud storage services," *Digital Investigation*, vol. 9, no. 2, pp. 81–95, 2012.

[3] J. Dykstra and A. T. Sherman, "Design and implementation of FROST: digital forensic tools for the OpenStack cloud computing platform," *Digital Investigation*, vol. 10, pp. S87–S95, 2013.

[4] A. Huth and J. Cebula, *The Basics of Cloud Computing*, Burlington, 2011.

[5] P. Mell and T. Grance, "The NIST definition of cloud computing," NIST Special Publication 800–145, 2011.

[6] Y. Pan and J. Zhang, "Parallel programming on cloud computing platforms," *Journal of Convergence*, vol. 3, pp. 23–28, 2012.

[7] S. Biggs and S. Vidalis, "Cloud computing: the impact on digital forensic investigations," *Internet Technology and Secured Transactions*, pp. 1–6, 2009.

[8] B. Martini and K.-K. R. Choo, "An integrated conceptual digital forensic framework for cloud computing," *Digital Investigation*, vol. 9, no. 2, pp. 71–80, 2012.

[9] M. Taylor, J. Haggerty, D. Gresty, and R. Hegarty, "Digital evidence in cloud computing systems," *Computer Law & Security Review*, vol. 26, no. 3, pp. 304–308, 2010.

[10] T. Teraoka, "Organization and exploration of heterogeneous personal data collected in daily life," *Human-Centric Computing and Information Sciences*, vol. 2, article 1, 2012.

[11] S. Silas, K. Ezra, and E. B. Rajsingh, "A novel fault tolerant service selection framework for pervasive computing," *Human-Centric Computing and Information Sciences*, vol. 2, pp. 1–14, 2012.

[12] T. J. Bittman, Top five private cloud computing trends, 2012, http://blogs.gartner.com/thomas_bittman/2012/03/22/top-five-private-cloud-computing-trends-2012/.

[13] S. Thorpe, "Virtual machine history model framework for a data cloud digital investigation," *Journal of Convergence*, vol. 3, 2012.

[14] X. Xie, H. Jiang, H. Jin, W. Cao, P. Yuan, and L. T. Yang, "Metis: a profiling toolkit based on the virtualization of hardware performance counters," *Human-Centric Computing and Information Sciences*, vol. 2, pp. 1–15, 2012.

[15] EMC White Paper, "Sizing EMC VNX Series for VDI workload," EMC, 2012.

[16] Citrix, "XenServer Citrix eDocs," 2012, http://support.citrix.com/proddocs/topic/xenserver/xs-wrapper.html.

[17] Vmware, "VMware View architecture planning," 2012, http://pubs.vmware.com/view-50/topic/com.vmware.ICbase/PDF/view-50-architecture-planning.pdf.

[18] J. Dykstra and D. Riehl, "Forensic collection of electronic evidence from infrastructure-as-a-service cloud computing," *Richmond Journal of Law and Technology*, vol. 19, no. 1, pp. 1–47, 2012.