Disposable Virtual Machines and Challenges to Digital Forensics Investigation

Mohammed Yousuf Uddin,[1], Sultan Ahmad*[2], Mohammad Mazhar Afzal[3]
Department of Computer Science and Engineering, Glocal University,
Saharanpur, Uttar Pradesh, India[1,3]
Department of Computer Science, College of Computer Engineering and Sciences,
Prince Sattam Bin Abdulaziz University,
Al-Kharj 11942, Saudi Arabia[2]

***Abstract*—Digital forensics field faces new challenges with emerging technologies. Virtualization is one of the significant challenges in the field of digital forensics. Virtual Machines (VM) have many advantages either it be an optimum utilization of hardware resources or cost saving for organizations. Traditional forensics' tools are not competent enough to analyze the virtual machines as they only support for physical machines, to overcome this challenge Virtual Machine Introspection technologies were developed to perform forensic investigation of virtual machines. Until now, we were dealing with persistent virtual machines; these are created once and used many times. We have extreme version of virtual machine and that is disposable virtual machine. However, the disposable virtual machine once created and are used one time, it vanish from the system without leaving behind any significant traces or artifacts for digital investigator. The purpose of this paper is to discuss various disposable virtualization technologies available and challenges posed by them on the digital forensics investigation process and provided some future directions to overcome these challenges.**

***Keywords*—Digital forensics; digital investigation; disposable virtual machines; light weight virtual machine; Microsoft sandbox; QEMU; qubes**

I. Introduction

Digital forensics is the process with four basic phases: collection, examination, analysis and reporting. During collection phase, data related to a specific event is identified, collected, and its integrity is maintained. Examination phase uses forensic tools and techniques as well as manual processes to identify and extract the relevant evidences from the collected data. Analysis phase deal with analyzing the results of the examination phase to generate useful information related to the case. Final phase generates reports of evidence from the results of the analysis [1]. A virtual machine (VM) is a tightly isolated software container with an operating system and applications inside. VM is self-contained and independent. Multiple VMs on a single physical machine with different operating systems and applications to run on just one physical server, or host. Hypervisor is the software layer, which decouples the virtual machines from the host and dynamically allocates and manages the computing resources to each virtual machine as per requirement [2]. Forensic investigation of virtual machines is challenging task if in a case virtual machine is subject of crime investigation, obtaining the image of the physical drive will not result in significant evidence since the virtual hard drive holds the evidence and more over vulnerabilities and attacks that affect the physical drive will have same effect on virtual environment. Analyzing multiple virtual machines using traditional tools of forensics is not possible. Virtual Machine introspection is the technique to monitor a virtual machine through hypervisor or a privileged VM, where the evidence collected without affecting the target VM [3]. Virtual machines created using oracle virtual box can be recovered using autopsy and other tools but VMs which were deleted using destroy command cannot be recovered [4]. The goal of this paper is to explore the disposable virtual machines and challenges posed to the digital forensics practitioners. Next section will discuss the virtualization technologies. Section 3 explores the disposable virtual machine technologies. Section 4 explores the challenges and roadblocks introduced by disposable virtualization to digital forensics. Section 5 will discuss current solutions to the issues related to disposable virtualization. We conclude with possible research directions to overcome these challenges.

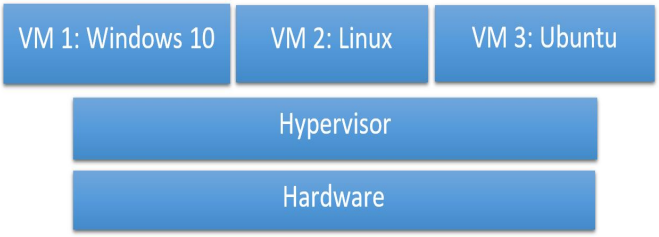

Fig. 1. Type-1 Hypervisor.

II. Virtual Machines

Virtualization technology enables utilization of resources in an effective way, reduces maintenance and security cost for the end-users. Virtual machine runs up on hypervisor. Hypervisors are of two types, one, which directly operates on physical hardware and does not require operating system, is called type-1 hypervisor, often called as "bare metal" hypervisors, examples include Citrix, Xen Server, ESXi from VMware, and Microsoft's Hyper-V. Layerd architecure of type-1 hypervisor illustrated in Fig. 1.

Second type of hypervisor rests upon operating system known as type-2 hypervisor. Most popular type-2 hypervisors are VMware, Virtual Box, and Parallel Desktop for MAC OS.

* Corresponding Author : Sultan Ahmad

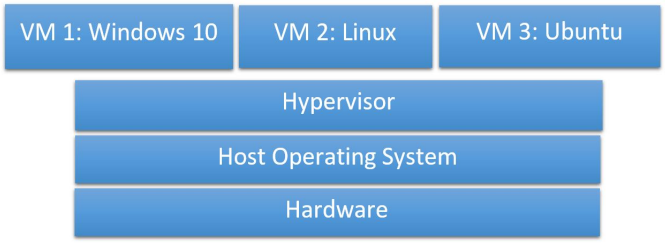

Fig. 2. Type-2 Hypervisor.

Type1 hypervisors provide greater performance and security and there is no overhead task for hypervisor to interact with host operating system. Type-2 hypervisor runs as an application on top of the host operating system (OS), it gives convenience to the individual users who intend to emulate a different operating system other than their OS, example: windows users can install Linux on virtual machine [5]. Fig. 2 shows the type-2 hypervisor's architecture. VMware files like vmdk file is virtual hard disk and vmem file is paging file act as primary memory RAM[6]. Oracle Virtual Box hypervisor also maintains such files, for each virtual machine there is a machine folder, inside machine folder vmname.vbox file and vmname.vdi , vdi format file for disk image, and Log files folder and a snapshot folder. These specific files of virtual machine collected from the host machine, to conducted investigation on virtual machine [7].

A. Virtual Machine Forensics

VM Forensics is similar to traditional digital forensics in many ways but at the same time, it introduces new pitfalls. Forensic approaches for virtual machines are many. Simplest form of forensics investigation of virtual machine starts with acquiring disk image of host computer on which virtual machines are running, after acquiring disk image files are extracted for the respective Virtual machine manger. Along with VM's files network logs and host operating system's registry also extracted. Disk image acquisition has to be done with utmost care, to preserve the integrity to ensure the legal admissibility of the evidence. There are standard procedures and guidelines for digital evidence acquisition approved by the Association of Chief Police Officers of the UK (ACPO), ISO Standard 27037, U. S. Department of Justice Office, and the EU publication Guidelines on Digital Forensic. First, the machine is powered off by disconnecting power supply. Then the hard disk drives or solid-state drives disassembled from the suspect machine. Extracted disk drive is write protected with write blocker kit. Disk drive then connected to forensic machine to create a duplicate image of disk drive using specialized tools such as dd, FTK imager and "encase", etc. Disk image acquired from previous step is used for analysis. In case of VM disk image there are two approaches, first is resuming the suspended virtual machine on corresponding virtual machine manager. Second approach is to create the snapshots of virtual machine.In case of resuming the suspended virtual machine VM disk files vmdk, or vdk or vhd files and other files related to virtual machine are restored, down side of this approach is during resuming process VM files may change and integrity of the evidence is compromised. While snapshot of VM used for forensic analysis, there will be no changes on state of the HDD. Forensic analysis tools; Encase, FTK supports the conversion of virtual disk image files (.vmdk, .vdi) to raw dd format files [8]. Virtual machine introspection technique uses virtual machine manager to view inside virtual machine, to track and view virtual machine state. VMI can inspect and view VM-memory, processor, installed Operating systems, applications and services. Evidence Search through injected code. This strategy is inspired by code injection attacks. Which uses vulnerabilities to inject malicious code in to applications and kernel to control and corrupt the system [9].

TABLE I. Disposable Virtual Machines.

Disposable VM	Hypervisor	Type
Microsoft Sadbox	Microsoft Hypervisor	Type 2
Qubes Disposable VM	KVM, Xen	Type 1
Virtual box Nested VM	Virtual Box	Type 2
Shade SandBox	Microsoft Hypervisor	Type 2
QEMU	Xen, KVM, Hax	Type 2
Bitbox	Virtual Box	Type 2

III. Disposable Virtual Machines

Disposable virtual machine is the lightweight virtual machine, created instantly and it will be disposed when it is closed. Disposable VMs commonly used to host single application, such as web browser, viewer, editor and suspicious applications. This concept of single use virtual machines also adopted by various operating systems. In Table I, the few popular disposable Virtual machine managers are listed.

A. Microsoft Windows Sandbox (WSB)

Microsoft Windows sandbox runs applications in isolation. Secure execution of application in sandbox environment does not affect the host operating system. New instance of sandbox created each time and disposed as soon as it is closed. Preinstalled applications in host operating system are not accessible in sandbox environment instead explicit installation of application is required. Sandbox uses hardware virtualization for kernel isolation. Windows Sandbox is a new lightweight disposable desktop environment. Which runs application in isolation. Windows 10 pro and enterprise editions include sandbox environment. As soon as sandbox is closed, applications and residual files, and data related to that particular sandbox deleted permanently. Every time you start a Windows Sandbox, it is as clean as a brand-new installation of Windows. Windows 10 operating systems has all required files pre-loaded to run the sandbox. It is disposable nothing persists on the host device as soon as you close the sandbox.Windows Sandbox (WSB) gets the dynamically generated base image with its own directory structure as host operating system, except the mutable files are copied in to WBS directory structure. Immutable files of host operating system can be accessed through links. Efficiency of the windows sandbox achieved by following: process scheduling integrated with kernel scheduler. Smart memory management where memory pages are allocated to WSB and Host operating system on demand, there is no fixed chunk of memory for WSB, it gives more flexibility and improves efficiency overall. virtual GPU enables dynamic utilization of graphics processing. Windows sandbox architecture

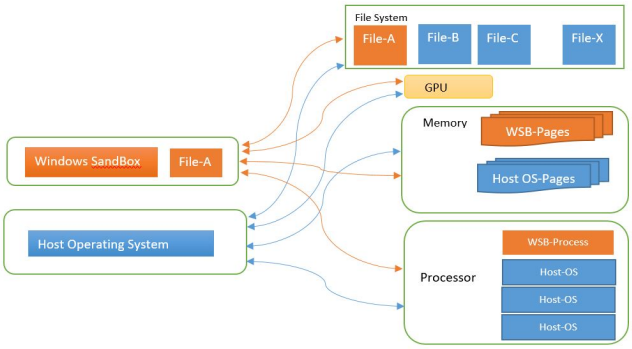

Fig. 3. Windows SandBox Architecture.

is illustrated in Fig. 3. To use windows sandbox you must start the sandbox first and copy the executable file you wish to run from the host file system and paste the executable file in sandbox. Once the file copied, you can run it as a normal application. Windows Sandbox gives two options; one is to run a full desktop in sandbox. Second option is just the application in sandbox and as known as rails. Sandbox has many advantages over tradition virtual machines creates a virtual machine and installs complete operating systems where resources are shared among host operating system and virtual machines. In case of windows, only few files used for sandbox from host file system it is dynamically generated image. Memory management is dynamic based on payload system allocates memory to the sandbox. Process scheduling is integrated where sandbox and host systems are managed together . Windows sandbox is secure as it runs on a separate kernel that provided by Microsoft's hypervisor keeping it isolated from the host kernel. Virtualization in case of sandbox is hardware-based illustrated in Fig. 1. Thus to implement type-1 hypervisor host system must support virtualization, which can be enabled or disabled from BIOS of the host system. Any malicious code will not affect the host kernel and will not persist as soon as sandbox is closed. WSB can be accessed remotely from server where Sandbox is created in two modes 1.WSB with full desktop 2.WSB Rails in Rails a specific application is launched on sandbox it is similar as Application VM. Remote clients can access and launch the WSB from server, once it is closed no files or changes are saved in host server[10].

B. Qubes Disposable VM

Qubes OS developed with focus on Security through isolation approach. Virtualization is based on Xen hypervisor. Domains created with different security levels, which runs on virtual machine. Work domain is more secure than Shopping domain. Dom0 is the administrative domain it can access all the hardware directly, such as graphics devices, input output devices like keyboard and mouse. This administrative domain manages the virtual disks of the other VMs, it stores these virtual disk images on its file system. Disk space saved by storing virtual disk on same file systems and accessed in read only mode. Qubes allows users to launch disposable VM directly from dom0's start menu or from an AppVM you have to choose open with disposable VM. In disposable VM you can work with untrusted files without compromising other Virtual machines. Disposable VMs created using Disposable VM Template. Disposable VMs created with these templates has its own user file system, one for each disposable VM. Qubes R4.0 has multiple Templates and default template for disposable VM is fedro-xx—dvm(xx here refers to version number[11].

C. VirtualBox

Oracle VirtualBox is an open source type 2 hypervisor for virtualization of window and Linux operating systems from Oracle Corporation. Creation and management of guest virtual machines is very much user friendly. Intel VT-x and AMD-V hardware-assisted virtualization is supported on VirtualBox. It supports nested virtualization that is one of the challenges for digital forensics experts [7]. Nested virtual machines runs on hypervisor which is on top of other virtual machine, this stacking of hypervisor recursively increases overhead but at the same time provides extra layer of security and decouples the VM from physical host [12]. Eventually it comes with extra overhead for digital forensic investigation.

D. Shade Sandboxie

Shade Sandboxie is an application based sandboxing. It creates isolated environment to execute suspicious code. Such an environment is used to track and notice code behavior and output activity, it creates functional layer of network security against ATPs and other cyber threats. Applications run inside simulated virtual environment without hardware virtualization support. Running malicious code and browsing websites with potential threats will not affect the host Operating System [13].

E. QEMU (Quick Emulator)

QEMU is the hosted virtual machine monitor it operates in different modes. System emulation mode where it emulates hardware including processor, peripheral devices. In user mode, it runs programs using different instruction set rather than its instruction set by cross-compilation and cross debugging. KVM hosting mode, QEMU emulates hardware but guest operating system runs on KVM. XEN hosting mode, here also QEMU emulates hardware and XEN run the guest operating systems [14].

F. BitBox

BitBox is secure firefox encased in virtual machine with linux OS on oracle virual box. Only drawback of this is the setup, which takes 2GB of disk space. Developed by German cyber Security Company Rohde and Schwarz to prevent cyber-attacks such as APTs, Zero-day exploits and Ransomwares[15].

IV. Challenges Posed by Disposable Virtual Machines in Digital Forensics

That, in essence, attackers can start a disposable VM to carry out their act and close the disposable VM, which leaves no traces for forensics expert. Existing Virtual machine forensic techniques are not going to yield significant results. The disposable virtual machines not designed with digital forensics

and evidence integrity in mind, instead the objective was to completely isolate applications from host operating system and leave a pristine system without leaving any traces behind. However, not any significant work has been done in disposable virtual machine forensic. We could not find any substantial information about disposable virtual machine or lightweight VM forensics. Forensic investigation begins with identifying the system, which contains potential evidence or involved in suspicious activity. First step is to identify the incident and next is to acquire evidence to prove the incident. When it comes to disposable virtual machines, no traces are left. The very nature of disposable virtual machines architecture is the main challenge in data identification and subsequent collection of evidence. Mostly no artifacts left after closing disposable virtual machines. Possible solution could be capturing the sandbox or the disposable virtual machine instances while they are active other possible solution is to perform data carving from memory dumps log files of hypervisor. In presence of hypervisor, it is difficult to take, the memory dump of the physical memory it is difficult to extract the data from memory reserved fur virtual machine monitors. One possible way to use memory acquisition tools like volatility, Rekall and Layout Expert [16].It might be able to analyze virtual machine processes running on the machine even after capturing memory dumps it is difficult to analyze the memory dump for virtual machine data. Here we use the standard forensic investigation steps to discuss the challenges posed by disposable VMs at each stage. Stage of forensic investigation are as follows: 1. Forensic Image creation 2. Identification and Recovery 3. Analysis 4. Presentation and Documentation[17].

A. Forensic Image Creation

Disk image of suspected system is created from physical machine. At this stage, integrity of the image created must be preserved. This is performed using tools like DD, DDRescue, Encase and Photorec etc.[18][19]. Investigator never uses the original disk to conduct investigation; instead, image of the disk used to conduct analysis and further investigation. This image used to collect the information about virtual machine and hypervisor used. Information included execution time logs, temporary files, snapshots and Internet activity log files etc. Therefore, investigator must collect the image carefully without tampering its integrity to extract vital information. Write blockers are used to prevent accidental writes on to the original disk. MD5 hashing is one of the method to ensure the integrity. Forensic tools allow us to complete this task by mounting disk image for further analysis of the Virtual machines and Hypervisor. Graphical user interface such as Dymanage and AIR are developed for DD find DD rescue. In case of disposable virtual machines, data is not persistent so it is not possible to create disk image of disposable virtual machines.

B. Identifcation and Recovery

At first, host machine is analyzed to find the traces of virtual machine in hypervisor. Host operating system maintains log files, which lead to extract traces of virtual machine. Windows operating system maintains registry entries, prefetched files, shared DLL, log files, thumbnails, icons, temporary files, and system event logs etc. that can prove the virtual machine existence in host computer. Even after files are deleted most of the time operating system do not completely delete the files instead removes the file reference from master file table. Specifically for large size, files such as virtual machine files still exist in the disk. Each of these files can be extracted from the unallocated space of the secondary disk. Data recovery tools like best disc, handy recovery and R-studio etc. commonly used to recover data from the disk image.

TABLE II. DISPOSABLE VM CHALLENGES.

Investigation Stage	Challenge
Image creation	No persistent files of disposable VM exist on disk drive
Information identification	Host OS or the Hypervisor do not maintain activity logs of disposable VM
Analysis	Snapshots or .vdi files are not available
Presentation	No specific format of reporting is available

C. Analysis

Virtual machine analysis: regular virtual machines can be Analyzed by mounting it as disk drive or by accessing it through a hypervisor. In case of disposable virtual machines, it is not possible; files related to disposable VM are deleted. Virtual machine files could be recovered by identifying its format based on hypervisor. Files with extensions like .VDI, .VMDK is used by popular hypervisors [20]. Other options to investigate the virtual machine is by picking a snapshot of VM and further analyzing snapshot to extract the vital information that can be presented as evidence. This provision of snapshot is not available for disposable virtual machines. Only option left for disposable VM is to analyze host operating system log files, registry entries, etc.

D. Presentation and Documentation

Documenting and presenting the evidence found during investigation is the final stage of forensic investigation. Evidence includes time stamps, who accessed and when accessed data or performed an activity. Forensics tools have their proprietary format of reports. There are no specific forensic tools for disposable VM forensics, so there exists no specific reporting formats for disposable VM. It is preferable to use the same reports as virtual machine. In Table II we have presented the challenges posed by disposable VM at every stage of forensic investigation.

V. CONCLUSION

In this paper, the investigators have explored challenges posed by lightweight VM to the digital forensics experts at every stage of digital forensics investigation. We discovered that there is not much research done in disposable VM forensics. These challenges needs to be addressed by conducting experiments on disposable VM. One of the possible thing is to compare the complete system image before and after running disposable virtual machine on various platforms and in this way we find possible traces or changes in the system.

ACKNOWLEDGMENT

The authors would like to acknowledge the support of the Deanship of Scientific Research at Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia.

REFERENCES

[1] K. Kent, S. Chevalier, T. Grance, and H. Dang, "Guide to integrating forensic techniques into incident response," *NIST Special Publication*, vol. 10, no. 14, pp. 800–86, 2006.

[2] VMWare.com, *VMware*, 2020 (accessed October 20, 2020). [Online]. Available: https://www.vmware.com/solutions/virtualization.html

[3] J. Poore, J. C. Flores, and T. Atkison, "Evolution of digital forensics in virtualization by using virtual machine introspection," in *Proceedings of the 51st ACM Southeast Conference*, 2013, pp. 1–6.

[4] E. Wahyudi, I. Riadi, and Y. Prayudi, "Virtual machine forensic analysis and recovery method for recovery and analysis digital evidence," *International Journal of Computer Science and Information Security*, vol. 16, 2018.

[5] P. Tobin and T. Kechadi, "Virtual machine forensics by means of introspection and kernel code injection," in *Proceedings of the 9th International Conference on Cyber Warfare & Security: ICCWS*, 2014, p. 294.

[6] S. Lim, B. Yoo, J. Park, K. Byun, and S. Lee, "A research on the investigation method of digital forensics for a vmware workstation's virtual machine," *Mathematical and computer modelling*, vol. 55, no. 1-2, pp. 151–160, 2012.

[7] Virtualbox.org, *VirtualBox*, 2020 (accessed October 20, 2020). [Online]. Available: https://www.virtualbox.org/manual/ch10.html

[8] M. Hirwani, Y. Pan, B. Stackpole, and D. Johnson, "Forensic acquisition and analysis of vmware virtual hard disks," 2012.

[9] P. Tobin, N.-A. Le-Khac, and T. Kechadi, "Forensic analysis of virtual hard drives," *Journal of Digital Forensics, Security and Law*, vol. 12, no. 1, p. 10, 2017.

[10] Microsoft, *Windows Sandbox*, 2020 (accessed October 15, 2020). [Online]. Available: https://docs.microsoft.com/en-us/windows/security/threat-protection/windows-sandbox/windows-sandbox-overview

[11] Q. OS, *DisposableVMs*, 2020 (accessed October 15, 2020). [Online]. Available: https://www.qubes-os.org/doc/disposablevm/

[12] B. Kauer, P. Verissimo, and A. Bessani, "Recursive virtual machines for advanced security mechanisms," in *2011 IEEE/IFIP 41st International Conference on Dependable Systems and Networks Workshops (DSN-W)*. IEEE, 2011, pp. 117–122.

[13] shadesandbox.com, *Shade Sandbox*, 2020 (accessed November 22, 2020). [Online]. Available: https://shadesandbox.com/blog

[14] qemu.org, *Quick Emulator*, 2020 (accessed November 22, 2020). [Online]. Available: https://www.qemu.org/documentation/

[15] https://www.rohde schwarz.com, *Browser In The Box*, 2020 (accessed November 22, 2020). [Online]. Available: https://www.rohde-schwarz.com

[16] T. Wu, F. Breitinger, and S. O'Shaughnessy, "Digital forensic tools: Recent advances and enhancing the status quo," *Forensic Science International: Digital Investigation*, vol. 34, p. 300999, 2020.

[17] S. R. Selamat, R. Yusof, and S. Sahib, "Mapping process of digital forensic investigation framework," *International Journal of Computer Science and Network Security*, vol. 8, no. 10, pp. 163–169, 2008.

[18] N. Reddy, "Linux forensics," in *Practical Cyber Forensics*. Springer, 2019, pp. 69–100.

[19] S. Widup, *Computer forensics and digital investigation with EnCase Forensic v7*. McGraw-Hill Education Group, 2014.

[20] H. Riaz and M. A. Tahir, "Analysis of vmware virtual machine in forensics and anti-forensics paradigm," in *2018 6th International Symposium on Digital Forensic and Security (ISDFS)*. IEEE, 2018, pp. 1–6.