D30 Virtual Machines in Computer Forensics Research

John Tebbutt, and Douglas White, MS, National Institute of Standards and Technology, 100 Bureau Drive Stop 8970, Gaithersburg, MD 208998970*

After attending this presentation, attendees will learn about some of the ways in which virtual machines can be used in computer forensics research to expedite the research process while providing more detailed information on the machines under observation.

This presentation will impact the forensic science community by showing how virtual machine technology is a significant addition to the toolkit of computer forensics researchers, facilitating adherence to the scientific model by enabling high-resolution capturing of machine states for investigation.

Virtual machines (VMs) are increasingly used by computer forensics investigators because they offer numerous well-documented advantages over physical machines. Computer forensics researchers also have found virtual machine technology to be of great use, for somewhat different reasons. Principal among these are the potential to create small (4GB or less) virtual drives; the ability to freeze or snapshot these drives while the machine is in a particular state; the ability easily to store such snapshots indefinitely (e.g. on a DVD); and the ability to investigate several such VMs simultaneously and over time on a single physical workstation.

VM technology has been helpful in the observation of changes in system characteristics as a result of specific actions. A snapshot of a machine in a known state is taken, an action performed, a second snapshot, taken and the two snapshots are compared in order to determine the consequences of the action. It is then a simple matter to return to the known state, perform different actions, and identify any patterns which may exist. The NSRL has used this approach primarily to examine changes occurring as a result of software installation: which existing files are changed, and how; which files are added; and what is the effect on the Windows™ registry?

VM technology is used to obtain coverage statistics for the NSRL data set. The NSRL derives its reference data directly from installation media, which raises the question as to the utility of the reference data when compared with files installed on computers. Using VMs, it is possible rapidly to quantify the coverage of the NSRL data set with regard to real systems. Beginning with a “bare metal” install in a VM, it was possible to investigate which operating system files are not found in the NSRL data set, progressing to known applications, and so on. Previously it had been necessary to attempt coverage estimates using arbitrary machines operating in everyday environments and with incompletely known installation histories.

Finally, VMs are occasionally used in the production of the NSRL reference data. For example, a new operating system is installed into a VM with a 4GB virtual hard drive, the VM is shut down, and the virtual hard drive is written to a DVD. The virtual hard drive can then be processed in the same way as an installation disc to obtain file hashes for inclusion into the NSRL data set. It can then be stored together with the installation media, rendering the process repeatable and verifiable. While this approach is labor intensive, there are circumstances in which it can be useful, for example, when files are stored in unknown proprietary archive formats on the installation media and would thus otherwise be unavailable for processing.

Virtual Machine, Machine State, Computer Forensics

** Presenting Author*