

SHADOWBOX: A Low-Artifact Framework for Analyzing Evasive Malicious Code

Javad Zandi

Florida International University

Lalchandra Rampersaud

Florida International University

Amin Kharraz

Florida International University

Abstract—Over the years, cyber crime attacks against critical services have become more sophisticated in launching low-profile operations. However, an even more alarming trend is the increasing difficulty of collecting relevant evidence about modern cyber crimes and the involved threat actors in the early stages before significant damage is done. This issue puts defenders at a significant disadvantage, as it becomes exceedingly difficult to understand the attack details and formulate an appropriate response.

Developing a robust analysis framework to collect evidence about modern threats has never been easy. One main challenge is to provide a robust trade-off between achieving sufficient visibility while leaving minimal detectable artifacts. This paper introduces SHADOWBOX, an open-source, low-artifact and portable analysis framework that can provide system-wide monitoring capabilities while satisfying contemporary checks that are used by modern malicious code. We designed multiple deployment scenarios, showing SHADOWBOX's potential in evidence gathering and threat reasoning in a real-world setting. By making SHADOWBOX and its execution trace data available to the broader research community, this work encourages further exploration in the field by reducing the engineering costs for threat analysis and building a longitudinal behavioral analysis catalog for diverse security domains.

I. INTRODUCTION

Modern cyber crimes are continuously getting cheaper, faster, and more consequential. For instance, in December 2021, just a few days after the Log4j vulnerability was disclosed [22], adversaries started to adapt their code to locate exposed log4j services on the Internet and actively exploited this vulnerability. Very recently, the Clop ransomware group claimed responsibility for attacking the MOVEit transfer service on several known financial, education, U.S. federal, and state governments over the Memorial Day holiday. The attack was so consequential that the U.S. State Department announced a \$10 million bounty for information on the actors and involved threat campaign [70]. These incidents, on the one hand, suggest that adversaries are opportunistic and have asymmetric power in repurposing their tools to target new vulnerabilities in exposed services. However, a more concerning issue is that collecting necessary forensic evidence about these incidents at the early stages of the attacks before they cause consequential damage to the target systems is often challenging.

Over the years, the security community has invested significant effort to close this gap by developing solutions to extract

insights from large volumes of raw network [7], [34], [42] or default logging traces [37], [42], [47], [50], [64], [76], [134] that primarily reflect temporal usage of systems resources (e.g., CPU, RAM, Network, Disk). While these efforts have been instrumental in detecting specific forms of attacks with predictable patterns (e.g., Ransomware [61], crypto mining [123]), they often fail to manifest fine-grained information on attack techniques that implement not evident or less known evasive techniques to deliver their payloads. Consider an attack where the offending process injects a payload into another process and uses that as a proxy to establish a backdoor to a remote server. This incident, as a sequence of system-wide temporal data, is not collected effectively by engines that rely on default logging mechanisms. This insufficient visibility over the dynamic behavior of malicious operations can put defenders in a highly disadvantaged position to understand the attack tactics and lateral movements, perform root cause analysis, and formulate a proper response.

The core **insight** in this paper is today's threat intelligence against modern evasive cyber crimes has to provide fine-grained system-wide visibility, leave minimal detectable artifacts, and be portable. These design goals have been historically difficult to attain at the same time. For instance, having fine-grained temporal visibility requires looking into very low-level components of the systems, which could make the solution less portable. For instance, hypervisor-based approaches can produce fine-grained behavioral artifacts; however, their effectiveness is highly dependent on the deployment scenario and the underlying virtualization environment. That is, the effectiveness of these techniques can vary with the hypervisor vendors due to differences in the implementation details of exposed APIs, isolation guarantees, and monitoring capabilities [68], [128]. Furthermore, hypervisor-based monitoring might be restricted or disabled due to multi-tenancy or performance trade-offs when deployed in cloud infrastructure. On the other hand, approaches such as API hooking offer great portability, but there are several ways to bypass them [8], [59]. This paper aims to move towards addressing this need by introducing SHADOWBOX— a portable and low-artifact analysis framework for malicious code analysis to strike a balance between offering *fine-grained visibility* while *minimizing detectable* artifacts. These two design requirements are critical to staying effective over time in a landscape where adversaries actively try to identify defense solutions. To this end, SHADOWBOX operates as a protected-service

979-8-3315-8969-1/25/\$31.00 © 2025 IEEE

in high-privileged mode, making it almost impossible for a large number of applications to fingerprint, tamper, or kill the engine. At the same time, it offers system-wide visibility, temporal data on processes and threads and various forms of I/O requests which are essentials to record the behavior of various forms of evasive attacks.

We highlight the portable design of SHADOWBOX through two distinct case studies in the e-crime ecosystem: (1) a baremetal threat analysis environment: we deployed the sandbox on several live-running physical machines to perform large-scale malware analysis on baremetal Windows, machines(Section IV-A), (2) a distributed deception-based infrastructure: We deployed the SHADOWBOX-enabled images as in-cloud deception-based threat infrastructure for 46 days and collected attack artifacts about real-world attacks on intentionally vulnerable systems in the wild (Section IV-B).

We built one of the first scalable baremetal analysis environments. We created Windows 11 OS images enabled by SHADOWBOX and deployed them across eight physical machines, providing the capability to analyze approximately 1,000 samples in a baremetal environment. We analyzed 79,544 malicious binaries over 30 months, from June 2022 to December 2024. The analysis shows the adoption of hardware-based fingerprinting techniques across modern malware families in the wild. In particular, among all forms of fingerprinting techniques we observed in the dataset, 1,187 of the samples were calling Windows Management Instrumentation (WMI) to detect advanced virtualized analysis systems (e.g., Hypervisor-based sandbox), 9,254 samples from Sabsik, Wacatac, AgentTesla families checked for direct CPU clock access, and 7,147 samples checking the BIOS hardware, firmware version, manufacturer, and configurations loaded during the system boot process as shown in table III. While advanced sandboxes may attempt to obscure these artifacts by reporting altered driver versions or hardware information, such modifications can significantly compromise system robustness. Specifically, these changes may cause system crashes or interfere with legitimate drivers, as critical services often require precise hardware details for essential operations such as integrity checking, patching, updating, and driver management. This underscores the practical utility of solutions like SHADOWBOX, which can be readily deployed as a baremetal analysis environment.

The analysis of recorded activities also shows that a low-artifact analysis system similar to SHADOWBOX can potentially trigger different program paths in the malicious code compared to the behavior observed in contemporary virtualized environments. For instance, we observed that 3,981,555 (75.35%) of the dropped files in all the experiments were delivered only in SHADOWBOX, and 884,301(16.74%) were dropped in the virtualized environment – the underlying technology that enables a large number of contemporary analysis systems. We have shared 2% of the dataset (117GB) for this submission (randomly selected across 941 families), for baremetal and virtualized analysis environments. Finally, SHADOWBOX was deployed as a deception-based

threat intelligence environment for 46 days by intentionally exposing vulnerable Windows services on a set of exposed cloud hosts. During the course of the experiment, we identified 734 successful intrusions. Our analysis shows that it takes approximately 6 hours for adversaries to discover the exposed vulnerable service and 13 hours to compromise the service. The analysis of the process and filesystem traces shows the execution of 1,221 executables, a collection of 401 shell scripts, 235 installers, 14,706 source code files, 438 digital certificates, and 4,586 custom dynamic libraries.

The proposed analysis framework and discussed use cases aimed to highlight a critical gap in security defense for portable yet robust analysis tools for attack comprehension, bringing more behavioral visibility without introducing trivial detectable artifacts. Our hope is that this work serves to raise awareness about the importance of low-artifact analysis frameworks for global visibility and intelligence gathering in today’s attack landscape. We also hope that our approach will prove helpful to the security community in identifying emerging threats that go beyond what is routinely observed today and open the door to future research on threat intelligence and malicious code analysis.

Contributions. The paper makes the following contributions:

- We propose a portable and low-overhead threat analysis engine that can be deployed in all modern versions of Windows operating systems. The engine collects relevant information at the process and thread level as well as I/O activities for analyzing run-time behavior of malicious.
- We deployed the proposed analysis engine in two case studies, showing the benefits of a portable and low-artifact security engine in understanding the modern threat landscape. In addition to generating a large catalog of threat artifacts in a baremetal environment, we deployed a distributed deception environment collecting artifacts about 734 successful attacks, delivering malware, running proxies, and run monetization software.
- The source-code¹ as well as the artifacts used in the analysis of the paper (2% of the data catalog)² are made available. The distribution of malware families for the artifacts submitted can be found in Table IV.

II. BACKGROUND / MOTIVATION

A. Threat Model

In this paper, we assume that adversaries have significant freedom to develop evasive malicious code, as documented in prior work and community reports. In particular, our assumptions about attackers’ capabilities are as follows:

Environment Sensitive: There is no lack of evidence that modern evasive attacks are armed with various forms of anti-analysis techniques [4], [15], [17], [18], [29], [64], [65], [73], [95], [106], [115]. A long list of those artifacts (e.g., specific processes, registry keys, services, and network adapters) is publicly reported [27]. These techniques are often used by

¹<https://tinyurl.com/LASECode>

²<https://tinyurl.com/LASEArt>

malicious code developers to successfully intrude on a system where they plan to launch code.

Debugging Resistant: Adversaries have historically incorporated anti-debugging techniques to complicate forensic analysis. In particular, one common debugging approach relies on code injection techniques [5], [14], [69], [127] in the context of the target payload. That is, analysis agents (i.e., DLL) are injected into the context of a malicious process for behavioral monitoring. There have been several techniques to automatically detect such analysis techniques by simply listing active processes on target systems to identify debuggers or creating snapshots of their own memory structures, such as their heap and modules (DLLs loaded in the process's virtual address space), to identify if they are being analyzed. Adversaries can incorporate anti-debugging techniques at different layers, including CPU registers, in-memory data structure checks, or calling specific APIs or native functions [14], [35], [77], [82], [105], [127].

Launching Proxy Operations: Malicious processes often inject the actual malicious code into the context of legitimate processes to bypass reputation-based services [97], [98]. This is an important assumption as it makes most debugging techniques that target particular processes ineffective. Hence, our threat model in this project will consider these and other similar risks, and the solutions that will be developed should be robust and tested against these malicious activities [49], [54], [97], [98].

In this study, we also assume that user-initiated processes solely operate in user mode. Consequently, interactions with lower-level system resources are channeled through operating system API calls, which can be intercepted by a low-level forensic engine operating in the kernel space. In addition, we also assume that the trusted computing base includes the OS kernel and underlying software and hardware stack, and that normal user-based access control prevents attackers from running malicious code with superuser privileges.

B. The Need for a Low-Artifact Forensics Engine for Today's Modern e-Crimes

To respond effectively to highly evasive methods described in the threat model, the security community has innovated at various layers to define their analysis frameworks [9], [42], [47], [59], [63], [64], depending either on whether they found it the most effective way or where it was easier to incorporate the proposed method. In this section, we describe the limitations of those methods and make the case for how the analysis engine should be implemented to study the majority of threats.

User-mode Hooking Techniques. Classic hooking techniques [14], [49], [77], [82], [114] often rely on inline overwriting of APIs which requires injecting custom-formulated payloads into the context of the target process address space to identify the invoked functions, which inherently leaves several artifacts (e.g., checking the integrity of the API function, monitoring the list of loaded processes). Furthermore, this is not a scalable approach due to the engineering efforts required

to overwrite the target APIs. These modifications often are not without problems and could cause serious issues (e.g., system crashes) because of unexpected exceptions and unhandled cases. Lastly, function hooking or debugging mechanisms are inherently designed to be single process-centric and often lose their effectiveness rapidly when a more system-wide behavioral analysis is required. Furthermore, subverting these mechanisms is a common practice among adversaries by copying the desired malicious code into the address space of another process or using other common hooking evasion techniques such as customized code, stole code [20], [57], [58], [66], [118], and sliding calls [58].

Built-in Logging and Diagnosis Platforms. Modern operating systems are often equipped with integrated logging platforms designed for event monitoring and application performance tracking (e.g., syslog) [78]. In Windows, Event Tracing for Windows (ETW) [129] serves as a logging tool for system-wide monitoring and performance data collection. While ETW trace recording has been useful for collecting security events (e.g., authentication logs), it was not primarily designed to stay undetected. Adversaries can enumerate active ETW providers that use it for analysis/inferencing purposes to list defense tools on target machines. This technique is often used to detect anti-malware tools on the target system. Furthermore, a user in an active session can disrupt the logging process by temporarily modifying ETW environment variables to redirect logs, disable logging in a specific context, or modify behaviors in an ETW-reliant application. That said, ETW offers significant visibility over the overall dynamic behavior of systems. However, the approach is vulnerable to fingerprinting techniques and disruption [23]. In Table VI, we also discuss a set of open-source projects that are built on top of ETW service and compare with SHADOWBOX.

Hypervisor-based Approaches. Hypervisor-based methods [13], [33], [38] have been proposed to address the above shortcomings by putting the analysis module out of the operating system to achieve increased visibility and analysis capabilities (e.g., memory extraction, low-cost analysis, and debugging). While these techniques offer significant flexibility for code analysis, there are still low-cost techniques to fingerprint hypervisor-based solutions by, for instance, abusing critical Windows core APIs (e.g., Windows Management Instrumentation [119], [124]) to extract hardware information or direct access to the CPU for detecting instruction execution delay. Disabling these APIs is possible. However, these changes can be consequential because the normal operation of the many critical software components depends on accurately extracting hardware system information. An alternative solution is to overwrite these APIs and modify the returned values. For instance, HyperDbg [56] develops a custom hypervisor that utilizes hardware-based instruction-level emulation and OS-level API. However, such modifications are very likely to be distinguished by advanced decoy samples discussed in prior work [95], [132] that look for specific hardware versions and configuration details.

This paper. Table I, in Appendix, provides seven more popular

forms of anti-debugging and anti-VM checks in the wild. The list covers 60 specific methods tested on more known research prototypes. HyperDbg [56] is a hardware-based instruction-level emulation platform for binary analysis. In addition to HyperDbg, we analyzed Unveil [59], a research prototype that analyzed more than 2 million malware samples. As mentioned in Table I, SHADOWBOX, as a baremetal solution, does not need to modify the return values of a large number of checks such as VMI calls or API calls. These values often need to be modified in hypervisor-based solutions to look realistic, but their effectiveness is highly dependent on the precision of the modified value – which can be a significantly expensive process.

C. Design Requirements

In the following, we describe the security requirements we are seeking to achieve acceptable visibility without leaving detectable artifacts.

Low-Artifact Operations. Low-artifact behavioral monitoring is a critical property in analyzing modern threats (e.g., analyzing malicious code, recording techniques, tactics, and procedures) because, in almost all of these incidents, adversaries’ goal is to identify signs of an analysis environment. As mentioned earlier, approaches such as user-mode hooking techniques or hypervisor techniques introduce specific forms of artifacts that can be weaponized by adversaries to analyze the environment. New threat intelligence frameworks should make the target systems more robust against these fingerprinting efforts.

System-wide Behavioral Monitoring. System-wide behavioral monitoring refers to the property that enables run-time behavioral monitoring in a multi-process environment. This is a critical property in modern forensic analysis towards achieving complete mediation since malicious code can launch the actual malicious payload via proxy operations (as mentioned in the threat modeling). A system-wide view can provide sufficient visibility about the interaction of a process not only with the operating system resources but also with other processes [26]. From a threat intelligence standpoint, system-wide monitoring helps to establish a clear and comprehensive audit trail of all actions taken on a system by users or processes.

OS Support. It is critical that monitoring mechanisms operate in a high-privileged mode while introducing low integration costs. To this end, in-kernel analysis engine is a viable design choice to achieve these goals due to the significant OS support. An in-kernel engine inherently runs as a privileged operation in the kernel-level [45]. Therefore, termination or manipulation of core functionalities will not be possible by the malicious process that operates at the user level. In particular, Windows officially supports the concept of protected services, called Early Launch Anti-Malware (ELAM), to allow critical services to run as protected services [67]. After the service is launched, Windows incorporates a code integrity mechanism to only allow trusted code to load into a protected service. Windows also protects these services from code injection [89], [94]. This approach will guarantee that the engine is protected from

common forms of availability, memory corruption, tampering, and code injection attacks.

Considering all design options and the proposed threat model, an in-kernel behavioral recording method appears to be a suitable layer for enabling a portable and low-artifact analysis layer to study common threats in the wild. That is, the in-kernel module does not modify any Windows APIs or inject any code into running processes. Furthermore, the flexibility to deploy the engine in baremetal systems makes the entire analysis system robust against advanced fingerprinting techniques that target hardware specifications. *That said, this design should not be considered as an alternative solution to the current hypervisor-based solution that can offer significant reverse engineering flexibility. Rather, it can serve as the first line of defense in threat analysis by satisfying fingerprinting checks that are impossible or costly in other defense systems.*

III. MONITORING RUN-TIME BEHAVIOR

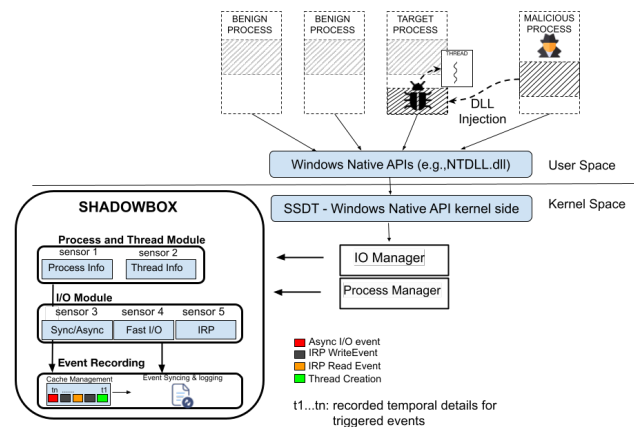


Fig. 1: The high-level view of SHADOWBOX’s architecture. SHADOWBOX provides a system-wide multi-process analysis engine by defining three in-kernel modules: (1) Process and thread module, (2) Filesystem and I/O module, (3) Event recording module. In this figure, SHADOWBOX records run-time behavior of the malicious process that uses DLL injection to launch a proxy attack.

In this section, we discuss the engineering decision and details of Low-Artifact Forensic Engine (SHADOWBOX). We elaborate on how we implemented the analysis engine. We provide examples of the collected artifacts and the deployment details of SHADOWBOX to better analyze modern threats.

A. Analysis Agents and Major Components

This section describes the main modules of SHADOWBOX that enables run-time trace recording.

I. Process and Thread Monitoring. Modern malicious code often employs code injection techniques such as DLL injection or process hollowing to inject code into legitimate processes [11], [12], [19], [58], [66], [96], [109]. This allows the malicious code to execute in the context of a legitimate process and run proxy attacks. Detecting such

TABLE I: Anti-Debugging and Anti-VM checks and their effectiveness in SHADOWBOX

Cat.	Method.	Example	HyperDbg [56]	Unveil [59]	SHADOWBOX
Anti-Debugging	API Call	GetCurrentProcessId() Process32Next() NtQueryInfor.Proc. FindWindow() NtSuspendThread() NtSetInf.Thread() CreateThread()	Modified via EPT-Hook	Detectable	Original Value
	HEAP Structure	HEAP.Flags HEAP.ForceFlags	Not Detectable	Not Detectable	Not Detectable
	Timing	GetTickCount QueryPerf.Counter GetLocalTime()	Modified via EPT-Hook	Detectable	Not Detectable
	Debugging	IsDebuggerPresentAPI IsDebuggerPresentPEB CheckRemoteDebuggerPresentAPI NSIT_ThreadHideFromDebugger NtGlobalFlag NQIP_ProcessDebugPort NQIP_ProcessDebugObject NQIP_ProcessDebugFlag CanOpenCsrss MemoryBreakpoints_PageGuard Interrupt_0x2d Interrupt_3 HardwareBreakpoints NQSI_KernelDebugger SharedUserData_KernelDebugger VirtualAlloc_WriteWatch_ NQO_ObjectTypeInfor. NQO_ObjectAllTypesInfor.	Not Detectable	Not Detectable	Not Detectable
Anti-VM/Hypervisor/Emulation	VM Checks	GetCurrentProcessId() reg_keys reg_key_value ldt_trick idt_trick gdt_trick str_trick vm_check_mac Firmware_RSMB Firmware ACPI Device Artifacts cpuid_hypervisor_vendor cpuid_is_hypervisor mouse_movement filesystem_artifacts setupdi_diskdrive	Not Detectable	Detectable	Not Detectable
	WMI Calls	manufacturer_computer_system_wmi model_computer_system_wmi vbox_mac_wmi process_id_processor_wmi disk_size_wmi serial_number_bios_wmi	Depends on Impl.	Detectable	Original Value
	Resource Profiling	Process_Enum Memory Space disk_size_getdiskfreespace disk_size_deviceiocontrol iSCSI_detection	Depends on Impl.	Detectable	Not Detectable

evasive operations has never been trivial since adversaries have significant freedom on what process to choose for injection, and making any assumption about the target process can impact the visibility factor. To this end, SHADOWBOX records system-wide monitoring for processes and threads by recording process-related events such as *initiation and termination*, thread-level events such as *creation and deletion*, and process image loading events such as *load*

and unload. This driver takes advantage of the Windows kernel APIs located in the `NtosKrnlib` [46], [53], [88] library and invokes embedded Windows Kernel functions: `PsSetCreateProcessNotifyRoutineEx()` [85], `PsSetCreateThreadNotifyRoutineEx()` [86] and `PsSetLoadImageNotifyRoutineEx()` [87] for process, thread and image activities respectively. Each of these functions allows

the SHADOWBOX driver to register a callback routine, which is registered in the kernel data structure within the process management subsystem and called by the Process Manager [133] whenever its corresponding event occurs. These functions are part of the process and thread management in Windows and provide mechanisms for kernel-mode components to monitor process lifecycle events.

II. I/O Request Packets. The execution of input and output operations in the Windows operating system, specifically related to driver operations, is facilitated by the utilization of packets. These packets, known as I/O packets or IRPs (I/O Request Packets), include encoded information that directs and controls the driver's actions [107], [110]. The self-contained structure of the Input/Output Request Packet (IRP) encompasses all the necessary information for a driver to manage an I/O request. To enhance operational efficiency and contextual relevance, IRPs are classified into major and minor operations. The major function code instructs the driver on what action to take to fulfill the I/O request. Minor codes refer to a specific category of major operations and supply further detailed information to the driver. A major function code specifies the specific type of I/O operation in conjunction with an optional minor function code [90], [107]. Recording the forms of IRPs shown in table IX in the Appendix is critical to understanding different classes of system interactions. That is, all filesystem activities, such as fetching files into the memory or interacting with cached files, are translated into a set of IRP requests that can be captured in the kernel. SHADOWBOX monitors all 103 IRPs, including 45 major and 48 minor I/O request packets (IRPs) based operations encompassing persistent and cached files.

III. Synchronous and Asynchronous I/O. In Windows, most of the I/O operations generated by processes are synchronous [92]. That is, a process waits until the OS returns a status code when the I/O is complete. For instance, ReadFile and WriteFile functions are executed synchronously. However, programs can modify the default behavior and issue asynchronous I/O requests by generating multiple I/O requests and continue executing while the OS performs the I/O operations. Our analysis shows that asynchronous behavior is largely ignored by almost all the current solutions [59], [63], [64], [117]. This is an important property because it allows processes to launch concurrent malicious operations, evading defense solutions that wait for the immediate return of the status code.

IV. Fast I/O. Programs can be designed to use Fast I/O requests for reading and writing to cached files, enabling rapid access synchronously [91]. In Fast I/O operations, the data is transferred directly between user buffers and the cache. This operation can bypass all current mechanisms that generate an IRP. For instance, if a process generates a Fast I/O operation requesting access to the content of an opened file in the memory, the requests will be processed immediately without generating any IRP events. This is an important observation since adversaries or malicious code can generate Fast I/O requests to access open files in the system buffer without

generating any IRP artifacts [116].

V. Event Recording. The event recording module operates as a producer-consumer architecture [16] where producers generate filesystem and process events, and consumers use the events for processing, which include storing on disk and sending high-priority events to the network. We developed a trace recording library that offers a set of Application Programming Interfaces (APIs) to store the events and manage I/O event monitoring. The library provides a mechanism to create data containers to maintain I/O traces in the memory for further processing. Generating the data containers requires close interaction with the OS cache manager to allocate the necessary memory space. One approach to implement the in-memory data cache is a write-back cache [40] where the data is updated only in the cache, and permanent storage is updated later. The write-back cache is a straightforward caching mechanism and is an excellent fit for general read and write I/O events. However, continuous I/O monitoring can significantly increase the required cache size on machines with limited memory and storage capacity. As an alternative approach, we implement a circular buffer mechanism using a queue, where I/O events are broken into smaller and more manageable data chunks for processing and consumption. The caching policy is implemented as a multi-threading model to deliver data from multiple producers to multiple consumers concurrently.

B. I/O Benchmarks

SHADOWBOX operates as an in-kernel module to record several forensics events. Consequently, it is important to evaluate the imposed overhead in a quantifiable fashion. To this end, we conducted an experiment to evaluate the overhead on filesystem operations as they are the more common and intense operations due to the customized instrumentation layer and log collection process. We generated workloads by creating 500 small (10Kb) and large (10Mb) files to test the throughput of block write, rewrite, and read operations. We used these operations as the major events that require significant filesystem interaction. The disk I/O performance was assessed using the popular Windows file system benchmarking tool, IOZone [2] by measuring the time to process the creation and access of generated files. Each experiment was repeated ten times for 4.42 minutes, and the average score was calculated to yield the final results presented in Table V in Appendix. The experiments show that SHADOWBOX performs well with both small and large files and imposes an overhead between 1.74% and 5.31%. We consider the 2.01% overhead for reading an existing small file for the first time. One reason for the relatively lower overhead for these critical operations is the design decision to monitor forensics events with minimal changes to the standard subsystems. Note that SHADOWBOX leverages an event-driven approach and does not require injecting code into the context of target processes. This approach has two main benefits: (1) it is a passive approach and avoids intrusive operations that may require any modification to the filesystem or standard OS

functions, introducing lower I/O overhead compared to inline API modifications, (2) it minimizes possible interference with the normal operation of processes and reduces the risk of crashes and memory corruption.

C. Resistance Against Common Dynamic Evasion Mechanisms

In the following, we briefly describe common fingerprinting checks used by adversaries in the wild and discuss the robustness of SHADOWBOX against each method.

VM Checks. Modern attacks are sensitive to virtualized environments and may not run correctly in those systems [9]. Checking registry keys (e.g., `reg_key`, `reg_key_value`) [1], firmware details (e.g., firmware ACPI, RSMB), CPU information (e.g., `cpuid`) [1] and hardware information (e.g., `model_computer_system_wmi`) are just a few examples of such fingerprinting techniques [95]. SHADOWBOX is portable and can be deployed on baremetal machines (see case study 1). Consequently, none of the VM checks would apply to the threat infrastructure enabled by SHADOWBOX. We should also highlight that SHADOWBOX-enabled infrastructure is robust against advanced fingerprinting techniques (e.g., device information retrieval through Windows Management Instrumentation) [93] used to fingerprint hypervisor-based environments given that the reference platform does not use any virtualization technologies vulnerable to such fingerprinting attempts.

Anti-Debugging Checks. Anti-debugging techniques are strategies to thwart efforts to understand and dissect malicious software. Anti-debugging aims to make debugging more difficult, time-consuming, or outright impossible. For instance, it is quite common in malicious payloads to check for debuggers in various levels (e.g., `SharedUserData_KernelDebugger`) by checking ports or debugging objects (e.g., `processDebugPort`, `processDebugObject`), interrupts, and hardware breakpoints (e.g., `Interrupt_0x2d`, `SystemKernelDebuggerInformation`) [36], [44], [74]. SHADOWBOX is an event-driven engine that relies on filter driver design to achieve its design goals. It does not require injecting the debugging agent into the context of target processes to record activities, leaving almost no traces for anti-debugging checks initiated by the malicious process.

Resource Profiling Checks. Resource profiling checks complement anti-debugging and VM checks [74]. For instance, checking the Windows data structure for processes (e.g., `process_enum`) [1], [95], or information on Disk (`disk_size_getdiskfreespace`, `disk_size_wmi`) [93] allows adversaries to identify sandboxes or virtual environments [1], [36]. Since SHADOWBOX is deployed in a baremetal environment, these checks will not likely assist adversaries.

D. Comparison with Other Open-Source Malicious Code Analysis Tools.

As the last comparison, we analyzed publicly available tools by comparing their architectural details, recorded artifacts, and

fingerprintable details that could impact the effectiveness of these engines.

Selection Metrics. We evaluated malicious code analysis engines using three key criteria for a head-to-head comparison: the engine (1) must be open-source, (2) must be compatible with Windows operating systems, and (3) must be installed and operate without compilation or fundamental deployment errors. Our initial analysis included 10 engines, from which we selected six for this evaluation following these metrics.

Analysis. Table VI illustrates the summary of the analysis. We observed that user-mode API hooking is still a common practice. For instance, BestEDR [130], OSquery [104], Wazuh [126], BlueSpawn [52] are all developed on top of API hooking or standard application logging mechanisms. Implementing this design strategy results in easier evasion and tampering since the forensic engine leaves detectable traces. Furthermore, tampering or terminating the engine is quite straightforward when the artifact recording engine operates at the same privilege level as other processes. Whids [108] approximates SHADOWBOX in terms of gaining visibility by augmenting Microsoft ETW. However, the log processing and the scheduling module reside in the user space, making the tool vulnerable to evasion and fingerprinting. As previously noted, ETW was initially developed for performance monitoring, and minimizing artifacts was not a primary design consideration. Furthermore, it is not designed to capture specific run-time security events, such as remote code injection, that can have significant security consequences. That said, SHADOWBOX offers system-wide visibility over processes, threads, and I/O activities while maintaining all functionalities (e.g., logging, monitoring tasks) within the kernel. This can potentially make SHADOWBOX a more generalizable solution for the same problem space.

E. Collecting Run-time Artifacts

In this section, we discuss the format and details of the recorded artifacts. To this end, we provide an example of a macro-based malware sample where the malicious code loads and compiles visual basic code. Table VII depicts the low-level temporal artifacts collected by SHADOWBOX in this attack. The sequence of commands provided illustrates a multi-faceted approach used by the macro-based malware to infiltrate a system, execute malicious payloads, and maintain persistence. In particular, the malicious code uses Dynamic Data Exchange (DDE) to pass commands to Excel, a method known to execute code within an Excel document. The subsequent command (line 2) indicates the use of embedded mode, likely to run a macro embedded within the Excel document.

The repeated use of ‘`eqnedt32.exe -Embedding`’ in the trace suggests the exploitation of vulnerabilities in the Equation Editor. Older versions of EQNEDT32.EXE have known security flaws that can be exploited to execute arbitrary code [32], [43]. Lines 14 to 25 show that the malware incorporates Windows command prompt and script hosts such as `wscript.exe` and `cscript.exe` to execute scripts. The filesystem activity shows that these scripts were used to download additional payloads

(i.e., xx.vbs, Podaliri4.exe), modify system settings, and establish persistence. Using commands such as ‘WmiPrvSE.exe -secured -Embedding’ invokes Windows Management Instrumentation (WMI) to perform system management tasks, including creating persistence and gathering system information. The artifact also shows that the malware uses Alternate Data Streams (ADS), a technique to hide malicious scripts within legitimate files, making them harder to detect. This multi-stage attack resulted in the launch of the final process, i.e., Podaliri4.exe, which delivered ransomware to the system. Figure 2 depicts the attack tree generated from the data collected.

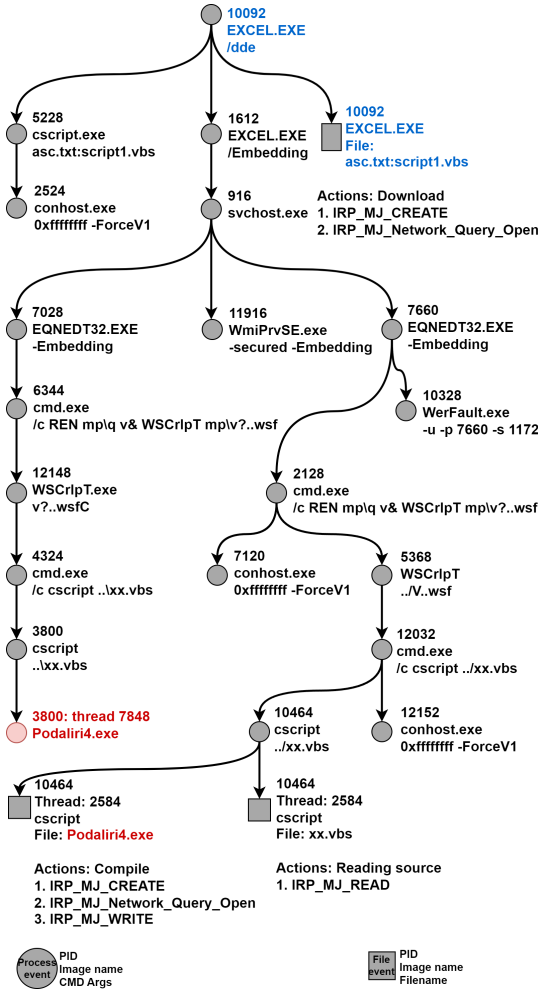


Fig. 2: The collected artifacts for a macro-based malware attack. The malicious process invokes 15 processes in the background to launch a successful attack.

IV. CASE STUDIES

SHADOWBOX is designed to be portable, low-artifact, and robust against common evasion mechanisms. In this section, we provide two case studies to illustrate how the engine can achieve these goals under different settings and deployment scenarios. In case study 1, we deployed SHADOWBOX in a baremetal malware analysis environment, aiming to address

common fingerprinting techniques in the wild. The analysis is based on 79,544 malware samples. In the second case study, we deploy SHADOWBOX as an in-cloud deception-based threat intelligence service with the intent to collect real-world artifacts on how adversaries exploit vulnerable machines and how they use those hijacked.

A. Case Study 1: A Baremetal-assisted Analysis Infrastructure for Malicious code

Virtualized environments have been used significantly in prior research [25], [59], [62], [122] to analyze malicious code due to the ease of deployment, scalability, isolation, and hardware utilization. However, this approach often comes with important visibility costs. In particular, environmentally sensitive samples launch several forms of fingerprinting checks to analyze the target host. They rarely load their actual malicious payloads in these environments, making analysis and reverse engineering a challenging task.

TABLE II: A subset of malware samples with the corresponding labels analyzed by SHADOWBOX-enabled threat analysis platform. The dataset contains four major malware types, 941 malware families, and 3,596 malware variants.

Malware	Families-Occurrences	Variants
Ransomware	64 - 472 (1.18%)	
StopCrypt	253 (0.63%)	48
Gandcrab	23 (0.06%)	13
Lockbit	20 (0.05%)	10
Hive	19 (0.05%)	6
MedusaLocker	15 (0.04%)	1
Filecoder	14 (0.03%)	8
Others	58 - 128 (0.32%)	71
RAT	602 - 31,158 (77.79%)	
Sabsik	3,474 (8.67%)	8
Berbew	3,116 (7.78%)	5
AgentTesla	2,362 (5.90%)	457
Wacatac	2,009 (5.01%)	3
Vindor	1,175 (2.93%)	3
RpcDcom based	1,069 (2.67%)	1
Others	596 - 17,953 (44.83%)	2,495
PUP	163 - 1,322 (3.30%)	
Zbot	92 (0.23%)	8
Infostealer	83 (0.21%)	1
KuaiZip	80 (0.20%)	1
AutoKMS	75 (0.19%)	3
Ymacco	72 (0.18%)	16
DarkStealer	67 (0.17%)	2
Others	157 - 853 (2.12%)	186
Self Replicating Malware	112 - 7,098 (17.72%)	
Sfone	982 (2.45%)	2
Ganelp	759 (1.89%)	7
Viking	752 (1.87%)	16
Xolxo	707 (1.76%)	1
Autorun	595 (1.48%)	13
Vobfus	492 (1.23%)	50
Others	106 - 2,811 (7.02%)	161
Total	941 - 40,050	3,596

In this case study, we study the feasibility of SHADOWBOX as a low-artifact engine and build a scalable baremetal analysis environment to collect artifacts about modern malware payloads. The core motivation behind this case study was to answer the question of how to build threat monitoring

platforms that can satisfy common fingerprinting techniques often used in modern malware samples. We started by asking how much visibility could be gained by running the evasive samples in the baremetal analysis environment.

This experiment was performed by executing 79,544 malware samples from across 941 malware families and analyzing the results. Table II shows the distribution of malware samples across different families for the 40,050 labels obtained. Each sample was executed on both the virtual and baremetal machines for five (5) minutes [69], during which run-time behavior artifacts were collected. The infrastructure used is equipped with high-bandwidth network switches to enable transferring Windows 10 images to the device in each run. That is, in each run, a new SHADOWBOX-enabled OS image was distributed and loaded into each machine. The baremetal machines had similar processing, memory, and storage specifications.

Figure 3 shows the summary of our analysis. While SHADOWBOX records 45 various low-level OS operations, we have provided six major operations to show the run-time behavior of a given sample. As shown, while there are samples with a similar number of key operations in both environments, there are still a large number of samples that show significantly more activities in baremetal compared to virtualized environments. For instance, the number of write operations (Figure 3b) is almost 64% more than the ones in a virtualized environment for 40,050 samples. The difference in the type of dropped payloads in the two environments suggests that the malware might have taken different execution paths in the two environments. We noticed the download of exe files, as a result of launching the malware, had a significant decrease (93.9%) in the virtualized environment. JavaScript files (JS) are the second most common type of malware payload downloaded, accounting for over 12.68% of all payloads. This is most likely due to the growing popularity of web-based attacks, as JS is the most often used scripting language on websites [28]. PDF documents are the most common type of non-executable payload downloaded, accounting for over 30% of all non-executable document payloads. The prevalence of PDFs as a document-sharing format is a probable cause, as they can serve as a vehicle for embedding and executing malicious code upon opening [51], [81]. Table VIII in the appendix shows a more comprehensive version of the dropped files in the experiments. We performed an analysis of the fingerprinting checks used by samples in the dataset that cause the differences in the execution traces of SHADOWBOX-enabled machines in baremetal and virtual environments. Our analysis shows that 1,187 samples from at least five malware families call Windows Management instrumentation API to get the actual hardware-level driver information before loading their actual payloads. Direct access to CPU clocks, bios, and timing checks have also been seen across various malware families in the test, as shown in table III. An immediate conclusion from this study is that a low-artifact analysis environment that can satisfy the hardware-based fingerprinting is critical to achieving more visibility in today’s attack landscape. That said, a baremetal-assisted

TABLE III: A subset of environment fingerprinting checks satisfied by SHADOWBOX, compared to other core technologies. ✓=Satisfied, ✗=Failed

FP Check	Samples	Baremetal-Based	VM-Based	Top Families
calls-wmi	1,187	✓	✗	AgentTesla, Sabsik, Redline, Leone, Znyon
direct-cpu-clock-access	9,254	✓	✗	AgentTesla, Sabsik, Wacatac, Woreflint, FormBook
checks-bios	6,859	✓	✗	AgentTesla, Leone, Znyon, Formbook, Wacatac
GetTickCount	1,288	✓	✗	AgentTesla, Sabsik, Woreflint, Wacatac, FormBook

solution should not be considered as an alternative solution to the current hypervisor-based solution that can offer significant reverse engineering flexibility. Instead, SHADOWBOX-enabled malware analysis can serve as the first line of defense in malicious code analysis by satisfying fingerprinting checks that are impossible or very costly in other defense systems. Furthermore, the results also show the need for a more effective analysis environment for the research community that forces a trade-off between achieving more fine-grained visibility while leaving minimally detectable artifacts on the analysis machines.

B. Case Study 2: A Distributed Threat Intelligence to Collect Evidence on Modern Cyber Crimes

Adversaries are continuously becoming more effective in launching low-profile attacks on critical systems. The question that arises is how to collect relevant evidence about these attacks and their strategies at the early stages of their operations. Lack of proper evidence on attack strategies can potentially put defenders in a highly disadvantaged position because understanding attack tactics and lateral movements becomes complex. While tools and services such as Event Tracing Windows (ETW) offer critical insights on systems, performing root cause analysis and formulating a proper response often requires more fine-grained data. In this case study, we developed a distributed deception-based threat intelligence infrastructure, using SHADOWBOX to gather evidence on real-world attacks on an intentionally exposed service.

To run this experiment, we deployed SHADOWBOX-enabled cloud instances on Windows 11 x64 with 2vCPUs. Run-time traces were collected via a high-privilege triage engine to manage collection, packing, and subsequent posting to our remote servers. By design, we made our vulnerable systems detectable by automated scanners that actively scan network-based services for weak Remote Desktop Protocol (RDP) credentials. Intrusions were detected by monitoring a subset of high-valued files; once touched, we received an immediate notification of a successful intrusion. In an effort to lessen the likelihood of encountering the same attack campaigns repeatedly, we reset the servers and assign a random IP address after each successful compromise.

We executed the experiments for a duration of 46 days, from December 19, 2023, to February 3, 2024. During this time, there were interactions with the environment from 398 unique IP addresses originating within 175 geographical locations. Figure 4 shows the distribution of successful attack origins during the experiment timeline. Our data shows that the exposed, vulnerable services were discovered on average 6 hours

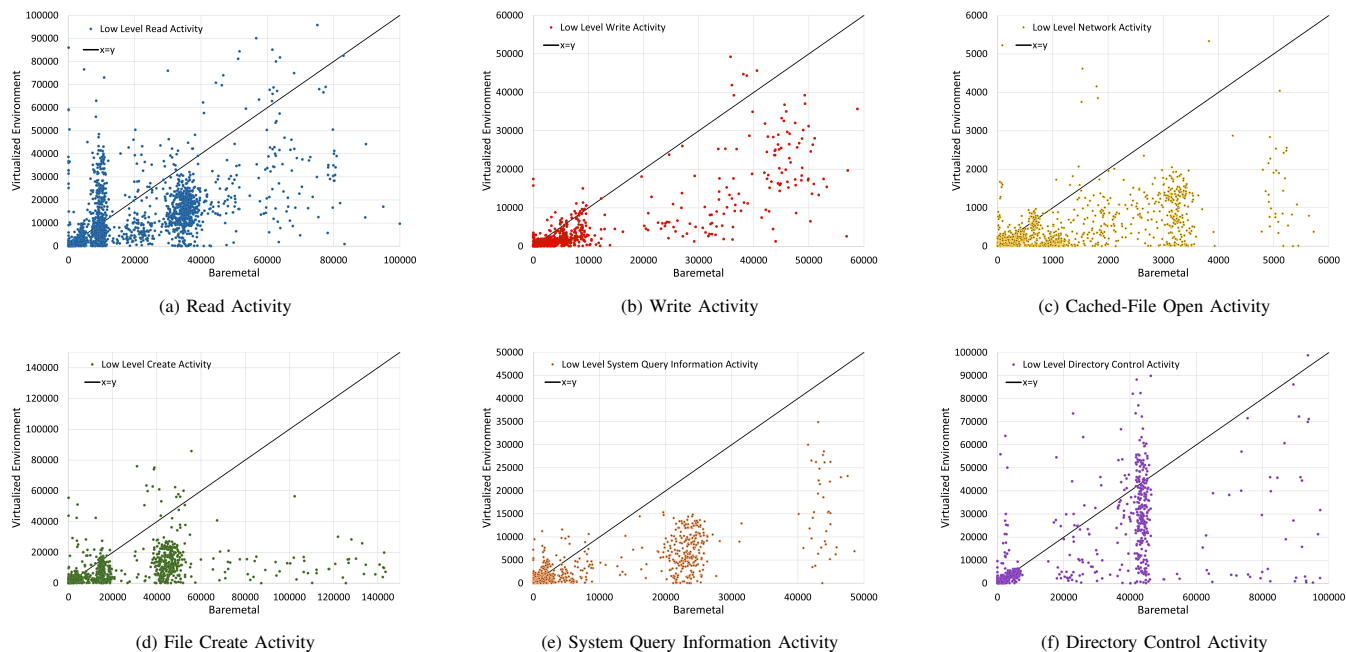


Fig. 3: File system activities in SHADOWBOX-enabled machines in baremetal and virtual environments. The results show that malware samples manifest more activities in the baremetal environment, resulting in more dropped files and execution payloads.

after publishing the instances, and successful exploitation occurred on average 13 hours after publishing the decoy server. We noticed that the high-valued files were opened in 32 of the experiments. Privilege escalation was inferred in three cases where we observed that files were written in the system32 folder (an operation requiring elevated permissions) using the logs generated by the forensic engine. In 51 experiments, at least one file was retrieved from remote servers and dropped on the machines. In particular, SHADOWBOX collected 196,774 dropped files, including 1,221 executable files, 4,586 DLL files, 401 shell scripts, 14,706 program source code files, 235 installers, and 438 digital certificates.

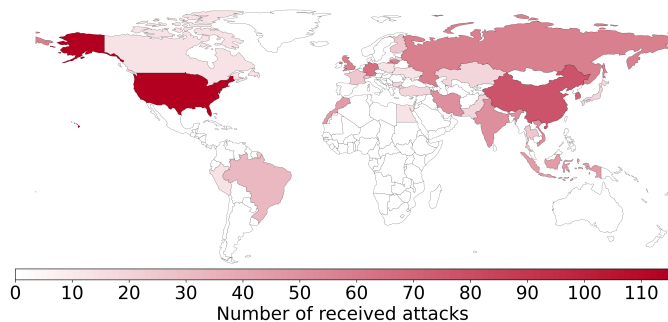


Fig. 4: Distribution of the origin of successful attacks. While US-based attacks are the largest by volume, we observed that most of them are being abused as proxy nodes.

We identified attack strategies that allowed adversaries to monetize the compromised machine or run code, which fa-

cilitated further exploitation or persistence. In the following, we describe some of the attack strategies we observed.

Using the compromised machine as a proxy. Traffic jacking is the process of converting a compromised machine into a proxy server and renting the server’s bandwidth for adversarial purposes [83]. To initiate traffic jacking, we observed that the adversary dropped an executable that loaded 183 DLLs into the memory, including several cryptographic and socket communication libraries. Once loaded, the process established a connection with a remote server to schedule tasks. Forensics trace data from SHADOWBOX shows that the parent process initiated hundreds of threads on the infected machine. The machine established a large number of TCP connections with hundreds of remote IP addresses to relay traffic. The adversary ran the malicious process in user-mode, and the forensic data captured by SHADOWBOX did not indicate any traces of establishing virtual interfaces for relaying packets. Our analysis suggests that the compromised machine was indeed converted to an exit node in this experiment.

Account Creation and Modification. In 15 of the exploitations, we observed indications of account creation and privilege escalation attempts using `net user` and `net localgroup` to gain persistent access. Another popular operation was group enumeration (e.g., `WMIC Group where "SID = 'X' Get Name /Value — Find "=)` to identify group memberships, specifically for admins and remote desktop users to potentially manipulate user privileges.

Backup Erasure We observed a pattern across 12 incidents where Powershell scripts attempted to delete the Volume Shadow Copies and backup catalog. We observed that adversaries were using commands such as `'vssadmin delete`

shadows /all /quiet', 'wbadmin delete catalog -quiet', or 'wmic shadowcopy delete' to prevent the recovery of data, making it more difficult to restore from backup after an attack. This practice is common among adversaries to either force the victim to pay a ransom fee or destroy the system logging evidence to complicate incident analysis.

Persistence Threats. We observed several attempts among adversaries to stay persistent on the exploited machines. We also observed 5 cases where the adversaries tried to modify the password policies to set passwords to never expire (e.g., net accounts /MAXPWAGE:unlimited), making it easier for adversaries to maintain access. In 10 incidents, we observed listing and manipulation queries where adversaries attempted to add or modify registry entries to hide user accounts from the login screen, making the accounts less noticeable to the casual user. In 5 cases, we observed no significant operation during the initial compromise. However, we observed setting scheduled tasks for future logins. In particular, we observed the use of `schtasks /create` to create tasks that execute at specific times or system events.

V. DISCUSSION

In this paper, we aimed to show that modern threat analysis highly depends on defining robust malicious code analysis engines that force a reliable trade-off between providing fine-grained behavioral insights and minimizing detectable artifacts. We posit that a solution that can achieve these two goals has several use cases on the defense side. As a first step in this direction, we defined two main application scenarios for SHADOWBOX and showed how it can add a new lens to the analysis and detection of security incidents. We summarize our main findings along the following four points:

More Visibility over Malware Behavior: While prior work shows modern malware samples are getting less sensitive to virtualized environment [69], our data suggests that such environmental checks are still prevalent across different RAT, droppers, and PUP samples. We observed that 64% (22,021) of the samples with file write activities had at least 70% more dropped or modified files in SHADOWBOX compared to cases in virtualized environments. We acknowledge that code coverage can have various definitions in the context of program analysis. However, the fact that more filesystem activities were observed in head-to-head experiments and a larger number of payloads were delivered on average suggests that the malware sample was designed to take a different path in environments close to real-world settings. In a broader context, this procedure can be augmented with static analysis approaches to systemically analyze malicious code more deeply. We believe systems similar to SHADOWBOX should be more accessible to threat researchers and malware analysts for further investigations in malicious code analysis due to the scarcity of real-world run-time behavior datasets on baremetal systems.

SHADOWBOX in a Broader Context. In this work, we aimed to answer how to improve behavioral visibility in modern evasive malicious code attacks while making them more robust to

evasion. That is a critical question in malicious code analysis and the generalizability of the results to other analysis engines and sandboxes is important. For instance, in case study 1, we showed that our baremetal solution works more effectively in comparison to the virtualized environment in satisfying almost all debugging and VM-based checks. We expect similar results if the experiments are conducted in other virtualization technologies (e.g., hypervisor-based). In particular, we observed that it is quite common across modern attacks to call Windows Management Instrumentation API in order to gain information on system hardware, applications, networks, devices, and other managed components. While manipulating the return values in WMI API to report fake responses is possible, it often leads to serious system crashes if the system has to operate for a longer period (case study 2) because the healthy operation of many system drivers depends on the responses received from the WMI API, and a fake response can easily cause a system crash. Unfortunately, hypervisor-based methods are still vulnerable to this fingerprinting method, although they offer significant freedom for other forms of analysis.

Unique Artifact Catalogs. The dataset we created in this experiment is the largest catalog of threat artifacts collected from a baremetal environment. While there are some available artifacts [6], [41], [59] that helped other work to have basic benchmarks, we are not aware of any large-scale run-time baremetal artifacts from prior work or commercial sandboxes on different forms of malware samples.

The high cost of building robust solutions and the scarcity of high-quality data can potentially impact the development of new data-driven defenses in this dynamic ecosystem where collecting evidence on emerging trends is critical. Lack of access to malicious data and the inherent complexity of defining a robust defense mechanism make it extremely difficult to establish a common ground for comparison. Given the ability to create a usable, timely, and longitudinal database of threat artifacts, we propose a set of research applications that would enable lower-cost threat intelligence. For instance, having longitudinal access to the artifacts of real-world attacks can be useful in testing new defense tools, studying attack evolution, and evaluating defense models with reference datasets.

Towards More Robust Cyber Deception Technologies.

The ability to deploy SHADOWBOX in different computing environments, including baremetal systems, end-points, and cloud environments, allows the deployment of a low-artifact deception-based threat intelligence infrastructure on the Internet. That is, deploying SHADOWBOX-enabled machines will serve as distributed threat intelligence sensors, offering a unique opportunity to collect evidence on new attacks. Consider the recent attack on the MOVEit transfer service as an example [22]. A large-scale honeypot enabled by SHADOWBOX could offer significant visibility on who first started the attack, how the attack took place, and a chronological order of the steps taken to perform the attack. As a next step, we plan to systematize the development of a vulnerability insertion module when a zero-day vulnerability is announced to measure the scope and scale, the evolution of the payloads and attacks,

as well as the adversarial campaigns behind the attacks.

VI. RELATED WORK

Forensics Services: Many of the prior works utilized the built-in Windows OS, Event Tracing for Windows (ETW), for their security analysis. For instance, Lee et al. [79] used the cursory data provided by ETW to detect APTs, thus lowering the overhead for log storage. Other approaches by Hassan et al. [50] look at the problem using static analysis to identify structures from built-in operating system audits. Rapsheet's [47] proposed approach involves the detection of APTs by condensing the acquired low-level data into graphical representations and subsequently comparing the observed steps with the publicly-accessible MITRE ATT&CK knowledge base [125]. This knowledge base is curated by domain experts who analyze real-world APT attacks. In fact, ETW is designed to primarily serve as a telemetry collection tool that gathers high-level event data. Although this is helpful, its use in forensic analysis is constrained because of the level and granularity of reports. Additionally, ETW is susceptible to fingerprinting and can be disabled by malware [21], [102]. Contributions from Bakshi et al. [10] built upon ETW to augment its ability to collect event activity logs. However, its ability to collect data at the granularity needed for a detailed low-level analysis still falls short [100]. Lower-level artifacts such as monitoring file systems, networks, and processes are more useful in this context [100].

At the kernel level, data can be collected using global operating system-level abstractions such as processes, files, and network interactions [48]. There have been several works in which the implementation of lower-level artifact collection is centered around the file system alone [3], [24], [59], [60], [71], [111], [112]. This is advantageous for detecting changes to the filesystem, especially in cases of ransomware in which many files are modified during a relatively short window [59], [111], [121]. A further extension incorporating data collection at a process level is included in RwGuard [84] whenever it was determined that the interactions were suspicious. It approaches the problem by deploying a logger in the kernel space and adding a process monitoring aspect within the user space. Although the methodology is adept at collecting process information, it is still tightly coupled with their logger. Still in the kernel at a memory level, Shah et al. [113] and Kara [55] proposed systems for detecting malware that resides in computer memory. After executing malware, they extract a memory dump that is further converted into an image for processing using computer vision and machine learning.

One gap in this domain is that the core technology used in prior work to collect threat intelligence data either offered relevant but high-level insights about incidents or was primarily focused on one class of attacks (e.g., ransomware). To our knowledge, SHADOWBOX is the only open-source solution that collects kernel-level data points (processes and threads, filesystem, registry) and generates GBs of data on the collected artifacts.

Sandbox and behavioral Analysis. A drawback for any system collecting data in this adversarial space is its susceptibility to being fingerprinted [30], [36], [103]. Advances in malware construction have seen the advent of anti-debugging and anti-sandboxing techniques to evade analysis and detection by security researchers trying to analyze malware behavior [31], [36]. These techniques are principally designed to make analysis time-consuming and tedious for researchers. It has been observed that malware behaves differently when it perceives that the environment is sandboxed for analysis or executed in a virtualized environment [9], [63], [69], [72], [74], [75], [120]. In some cases, the malware behaves benignly or does not execute if it detects a sandboxed environment [44], [95]. Proposals for anti-sandboxing techniques Liu et al. [75] build upon the works of Miramirkhani et al. [95] and Hu et al. [117] in creating believable system artifacts to mitigate system fingerprinting.

As a technique for sidestepping the issue of virtualized systems being detected, several works have proposed the use of physical machines for the collection of forensic data [63], [64], [117]. BareBox [63] uses a restoration method in which snapshots of memory and disks are used simultaneously with parallel operating systems. BareCloud [64] uses a distributed model of disks, restoring them after each execution. LO-PHI [117], like BareCloud uses machines that do not contain instrumentation for fingerprinting. SHADOWBOX is designed to complement all these methods by being portable, allowing deployment of the SHADOWBOX as an OS service in any computing environment.

Fingerprinting and Evasive Techniques. Resistance to profiling in malware is presented in different forms where malicious authors develop a multitude of techniques to hide the true behavior of their applications whenever an analysis environment is detected. To this end, several studies [36], [39], [80], [101], [131] have focused on understanding the techniques used to evade analysis. Maffia et al. [80] investigate each technique, utilizing publicly available data to create an evasive program profiler to detect and circumvent evasive measures and collect statistical data on evasion methods. Similar research was carried out by Garollo et al. [36] to produce statistically significant results on the link between malware families, the commonality of evasion techniques, and modifications to evasion techniques within families. They investigated evasive techniques in legitimate software and found that they are employed less frequently. Classification of malware by their evasive behaviors is studied by Yin and Nunes [101], [131]. Nunes et al. [101] note that applications are classified as malicious simply based on evasive methods taken.

VII. CONCLUSIONS

In this paper, we propose SHADOWBOX, a low-artifact in-kernel forensics analysis system that aims to optimize a balance between offering fine-grained visibility while minimizing detectable artifacts. We evaluated SHADOWBOX by

running two case studies. In the first case study, we integrated SHADOWBOX in a baremetal analysis environment and analyzed more than 79K malware samples. We collected GBs of artifacts on modern malware threats. In the second case study, we deployed the system as a deception-based threat intelligence, aiming to collect real-world threat artifacts. Our analysis showed that SHADOWBOX was successful in collecting artifacts about hundreds of binaries and shell scripts developed and executed by adversaries on the threat infrastructure.

ACKNOWLEDGMENT

We would like to thank the anonymous reviewers for their thoughtful feedback. This project was supported by Microsoft AI Security, US National Security Agency (NSA) under Grant No. H98230-21-1-0324, US National Science Foundation (NSF) under Grant No. 2219920, and Intergovernmental Personnel Act Independent Research & Development Program. Opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding agencies.

REFERENCES

- [1] Impeding automated malware analysis with environment-sensitive malware. In *7th USENIX Workshop on Hot Topics in Security (HotSec 12)*, Bellevue, WA, August 2012. USENIX Association.
- [2] Iozone filesystem benchmark. www.iozone.org, 2024.
- [3] Muhammad Shabbir Abbasi, Harith Al-Sahaf, Masood Mansoori, and Ian Welch. Behavior-based ransomware classification: A particle swarm optimization wrapper-based approach for feature selection. *Applied Soft Computing*, 121:108744, 2022.
- [4] Amir Afianian, Salman Niksefat, Babak Sadeghiyan, and David Baptiste. Malware dynamic analysis evasion techniques: A survey. *ACM Comput. Surv.*, 52(6), nov 2019.
- [5] Saleh Alzahrani, Yang Xiao, and Wei Sun. An analysis of conti ransomware leaked source codes. *IEEE Access*, 10:100178–100193, 2022.
- [6] Andrea Continella. ShieldFS: a self-healing, ransomware-aware filesystem. <http://shieldfs.necst.it/>, Accessed: 05-06-2024.
- [7] Giovanni Apruzzese, Pavel Laskov, and Johannes Schneider. Sok: Pragmatic assessment of machine learning for network intrusion detection. In *2023 IEEE 8th European Symposium on Security and Privacy (EuroS&P)*, pages 592–614. IEEE, 2023.
- [8] Cristian Assaiante, Simone Nicchi, Daniele Cono D’Elia, and Leonardo Querzoni. Evading userland api hooking, again: Novel attacks and a principled defense method. In *Detection of Intrusions and Malware, and Vulnerability Assessment: 21st International Conference, DIMVA 2024, Lausanne, Switzerland, July 17–19, 2024, Proceedings*, page 150–173, Berlin, Heidelberg, 2024. Springer-Verlag.
- [9] Erin Avllazagaj, Ziyun Zhu, Leyla Bilge, Davide Balzarotti, and Tudor Dumitras. When malware changed its mind: An empirical study of variable program behaviors in the real world. In *30th USENIX Security Symposium (USENIX Security 21)*, pages 3487–3504. USENIX Association, August 2021.
- [10] Akshay Bakshi, Tanish Sawant, Prasad Thakare, Azeez Dandawala, Manjesh K. Hanawal, and Atul Kabra. Improving threat detection capabilities in windows endpoints with osquery. In *2023 15th International Conference on COMMUNICATION SYSTEMS & NETWORKS (COMSNETS)*, pages 432–435, 2023.
- [11] Thomas Barabosch, Niklas Bergmann, Adrian Dombeck, and Elmar Padilla. Quincy: Detecting host-based code injection attacks in memory dumps. In *Detection of Intrusions and Malware, and Vulnerability Assessment: 14th International Conference, DIMVA 2017, Bonn, Germany, July 6-7, 2017, Proceedings 14*, pages 209–229. Springer, 2017.
- [12] Thomas Barabosch and Elmar Gerhards-Padilla. Host-based code injection attacks: A popular technique used by malware. In *2014 9th International Conference on Malicious and Unwanted Software: The Americas (MALWARE)*, pages 8–17. IEEE, 2014.
- [13] RR Branco, GN Barbosa, and PD Neto. Scientific but not academical overview of malware anti-debugging. *Anti-Disassembly and Anti-VM Technologies, Black Hat USA*, 2012, 2012.
- [14] Doug Brubacher. Detours: Binary interception of win32 functions. In *Windows NT 3rd symposium (windows NT 3rd symposium)*, 1999.
- [15] Alexei Bulazel and Bülent Yener. A survey on automated dynamic malware analysis evasion and counter-evasion: Pc, mobile, and web. In *Proceedings of the 1st Reversing and Offensive-Oriented Trends Symposium, ROOTS*, New York, NY, USA, 2017. Association for Computing Machinery.
- [16] Gregory T Byrd and Michael J Flynn. Producer-consumer communication in distributed shared memory multiprocessors. *Proceedings of the IEEE*, 87(3):456–466, 1999.
- [17] Ping Chen, Christophe Huygens, Lieven Desmet, and Wouter Joosen. Advanced or not? a comparative study of the use of anti-debugging and anti-vm techniques in generic and targeted malware. In *ICT Systems Security and Privacy Protection: 31st IFIP TC 11 International Conference, SEC 2016, Ghent, Belgium, May 30-June 1, 2016, Proceedings 31*, pages 323–336. Springer, 2016.
- [18] Xu Chen, Jon Andersen, Z Morley Mao, Michael Bailey, and Jose Nazario. Towards an understanding of anti-virtualization and anti-debugging behavior in modern malware. In *2008 IEEE international conference on dependable systems and networks with FTCS and DCC (DSN)*, pages 177–186. IEEE, 2008.
- [19] Binlin Cheng, Jiang Ming, Jianmin Fu, Guojun Peng, Ting Chen, Xiaosong Zhang, and Jean-Yves Marion. Towards paving the way for large-scale windows malware analysis: Generic binary unpacking with orders-of-magnitude performance boost. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, pages 395–411, 2018.
- [20] Binlin Cheng, Jiang Ming, Erika A Leal, Haotian Zhang, Jianming Fu, Guojun Peng, and Jean-Yves Marion. {Obfuscation-Resilient} executable payload extraction from packed malware. In *30th USENIX Security Symposium (USENIX Security 21)*, pages 3451–3468, 2021.
- [21] Adam Chester. Hiding your .net - etw. <https://blog.xpnsec.com/hiding-your-dotnet-etw/>, 2020.
- [22] Cisco Talos. Active exploitation of the MOVEit Transfer vulnerability by Clop ransomware group. <https://blog.talosintelligence.com/active-exploitation-of-moveit/>, Accessed: 06-30-2023.
- [23] Claudiu Teodorescu, Igor Korkin, Andrey Golchikov. Veni, No Vidi, No Vici: Attacks on ETW Blind EDR Sensors. <https://i.blackhat.com/EU-21/Wednesday/EU-21-Teodorescu-Veni-No-Vidi-No-Vici-Attacks-On-ETW-Blind-EDRs.pdf>, Accessed:01-08-2025.
- [24] Andrea Continella, Alessandro Guagnelli, Giovanni Zingaro, Giulio De Pasquale, Alessandro Barengi, Stefano Zanero, and Federico Maggi. Shieldfs: a self-healing, ransomware-aware filesystem. In *Proceedings of the 32nd annual conference on computer security applications*, pages 336–347, 2016.
- [25] Léo Cosseron, Louis Rilling, Matthieu Simonin, and Martin Quinson. Simulating the network environment of sandboxes to hide virtual machine introspection pauses. In *Proceedings of the 17th European Workshop on Systems Security, EuroSec ’24*, page 1–7, New York, NY, USA, 2024. Association for Computing Machinery.
- [26] Crispin Cowan. Turing around the security problem. In *15th USENIX Security Symposium (USENIX Security 06)*. USENIX Association, 2006.
- [27] Cybrary. Windows Commands Most Used by Attackers. <https://www.cybrary.it/blog/windows-commands-used-attackers>, Accessed: 05-06-2024.
- [28] Kyle Daigle. Octoverse: The state of open source and rise of ai in 2023. <https://github.blog/2023-11-08-the-state-of-open-source-and-ai/>, 2023.
- [29] Daniele Cono D’Elia, Emilio Coppa, Federico Palmaro, and Lorenzo Cavallaro. On the dissection of evasive malware. *IEEE Transactions on Information Forensics and Security*, 15:2750–2765, 2020.
- [30] Daniele Cono D’Elia, Emilio Coppa, Federico Palmaro, and Lorenzo Cavallaro. On the dissection of evasive malware. *IEEE Transactions on Information Forensics and Security*, 15:2750–2765, 2020.
- [31] Daniele Cono D’Elia, Lorenzo Invidia, Federico Palmaro, and Leonardo Querzoni. Evaluating dynamic binary instrumentation sys-

- tems for conspicuous features and artifacts. *Digital Threats*, 3(2), feb 2022.
- [32] Waeland Elder. Automatic extraction of vulnerability information for security operators using gpt models. 2024.
- [33] Jason Franklin, Mark Luk, Jonathan M McCune, Arvind Seshadri, Adrian Perrig, and Leendert Van Doorn. Remote detection of virtual machine monitors with fuzzy benchmarking. *ACM SIGOPS Operating Systems Review*, 42(3):83–92, 2008.
- [34] Zhuoqun Fu, Mingxuan Liu, Yue Qin, Jia Zhang, Yuan Zou, Qilei Yin, Qi Li, and Haixin Duan. Encrypted malware traffic detection via graph-based network analysis. In *Proceedings of the 25th International Symposium on Research in Attacks, Intrusions and Defenses*, pages 495–509, 2022.
- [35] Hisham Shehata Galal, Yousef Bassyouni Mahdy, and Mohammed Ali Atia. Behavior-based features model for malware detection. *Journal of Computer Virology and Hacking Techniques*, 12:59–67, 2016.
- [36] Nicola Galloro, Mario Polino, Michele Carminati, Andrea Continella, and Stefano Zanero. A systematic and longitudinal study of evasive behaviors in windows malware. *Computers & Security*, 113:102550, 2022.
- [37] Varun Gandhi, Sarbartha Banerjee, Aniket Agrawal, Adil Ahmad, Sangho Lee, and Marcus Peinado. Rethinking system audit architectures for high event coverage and synchronous log availability. In *32nd USENIX Security Symposium (USENIX Security 23)*, pages 391–408, 2023.
- [38] Tal Garfinkel, Keith Adams, Andrew Warfield, Jason Franklin, et al. Compatibility is not transparency: Vmm detection myths and realities. In *HotOS*, 2007.
- [39] Jiakuan Geng, Junfeng Wang, Zhiyang Fang, Yingjie Zhou, Di Wu, and Wenhan Ge. A survey of strategy-driven evasion methods for pe malware: Transformation, concealment, and attack. *Computers & Security*, 137:103595, 2024.
- [40] Shahram Ghandeharizadeh and Hieu Nguyen. Design, implementation, and evaluation of write-back policy with cache augmented data stores. *Proceedings of the VLDB Endowment*, 12(8):836–849, 2019.
- [41] Giorgio Severi. MALREC: Compact Full-Trace Malware Recording for Retrospective Deep Analysis. <https://giantpanda.gtisc.gatech.edu/malrec/dataset/>, Accessed: 05-06-2024.
- [42] Akul Goyal, Xueyuan Han, Gang Wang, and Adam Bates. Sometimes, you aren't what you do: Mimicry attacks against provenance graph host intrusion detection systems. In *30th ISOC Network and Distributed System Security Symposium (NDSS'23)*, San Diego, CA, USA, 2023.
- [43] Yeming Gu, Hui Shu, Pan Yang, and Rongkuan Ma. Minsib: Minimized static instrumentation for fuzzing binaries. *Computers & Security*, 122:102894, 2022.
- [44] Francis Guibernau and Ayelen Torello. 'catch me if you can!—detecting sandbox evasion techniques. *Proc. USENIX Assoc*, 2020.
- [45] Rajat Gupta, Lukas Patrick Dresel, Noah Spahn, Giovanni Vigna, Christopher Kruegel, and Taesoo Kim. Popcorn: Popping windows kernel drivers at scale. In *Proceedings of the 38th Annual Computer Security Applications Conference*, pages 854–868, 2022.
- [46] Tran Hoang Hai, Vu Van Thieu, Tran Thai Duong, Hong Hoa Nguyen, and Eui-Nam Huh. A proposed new endpoint detection and response with image-based malware detection system. *IEEE Access*, 11:122859–122875, 2023.
- [47] Wajih Ul Hassan, Adam Bates, and Daniel Marino. Tactical provenance analysis for endpoint detection and response systems. In *2020 IEEE Symposium on Security and Privacy (SP)*, pages 1172–1189. IEEE, 2020.
- [48] Lori Whippler Hollasch, Andrew Kim, Sameer Saiya, and Matt. File systems driver design guide. <https://learn.microsoft.com/en-us/windows-hardware/drivers/ifs/>, 2022.
- [49] Ashkan Hosseini. Ten process injection techniques: A technical survey of common and trending process injection techniques, endgame. <https://www.elastic.co/blog/ten-process-injection-techniques-technical-survey-common-and-trending-process>, 2018.
- [50] Hassaan Irshad, Gabriela Ciocarlie, Ashish Gehani, Vinod Yegneswaran, Kyu Hyung Lee, Jignesh Patel, Somesh Jha, Yonghui Kwon, Dongyan Xu, and Xiangyu Zhang. Trace: Enterprise-wide provenance tracking for real-time apt detection. *IEEE Transactions on Information Forensics and Security*, 16:4363–4376, 2021.
- [51] Maryam Issakhani, Princy Victor, Ali Tekeoglu, and Arash Habibi Lashkari. Pdf malware detection based on stacking learning. In *ICISSP*, pages 562–570, 2022.
- [52] Jacob Smith. BlueSpawn. <https://github.com/ION28/BLUESPAWN>, Accessed: 26-05-2024.
- [53] Justin Jones and Narasimha Shashidhar. Ransomware analysis and defense. *Journal of Colloid and Interface Science*, 374(1):45–53, 2012.
- [54] Rémi Jullian. In-depth formbook malware analysis – obfuscation and process injection. <https://www.stormshield.com/news/in-depth-formbook-malware-analysis-obfuscation-and-process-injection/>, 2023.
- [55] Ilker Kara. Fileless malware threats: Recent advances, analysis approach through memory forensics and research challenges. *Expert Systems with Applications*, 214:119133, 2023.
- [56] Mohammad Sina Karvandi, MohammadHossein Gholamrezaei, Saleh Khalaj Monfared, Soroush Meghdadizanjani, Behrooz Abbassi, Ali Amini, Reza Mortazavi, Saeid Gorgin, Dara Rahmati, and Michael Schwarz. Hyperdbg: Reinventing hardware-assisted debugging (extended version), 2022.
- [57] Yuhei Kawakoya, Makoto Iwamura, Eitaro Shioji, and Takeo Hariu. Api chaser: Anti-analysis resistant malware analyzer. In *Research in Attacks, Intrusions, and Defenses: 16th International Symposium, RAID 2013, Rodney Bay, St. Lucia, October 23-25, 2013. Proceedings 16*, pages 123–143. Springer, 2013.
- [58] Yuhei Kawakoya, Eitaro Shioji, Makoto Iwamura, and Jun Miyoshi. Api chaser: Taint-assisted sandbox for evasive malware analysis. *Journal of Information Processing*, 27:297–314, 2019.
- [59] Amin Kharaz, Sajjad Arshad, Collin Mulliner, William Robertson, and Engin Kirda. {UNVEIL}: A {Large-Scale}, automated approach to detecting ransomware. In *25th USENIX security symposium (USENIX Security 16)*, pages 757–772, 2016.
- [60] Amin Kharraz and Engin Kirda. Redemption: Real-time protection against ransomware at end-hosts. In *Research in Attacks, Intrusions, and Defenses: 20th International Symposium, RAID 2017, Atlanta, GA, USA, September 18–20, 2017. Proceedings*, pages 98–119. Springer, 2017.
- [61] Amin Kharraz, William Robertson, Davide Balzarotti, Leyla Bilge, and Engin Kirda. Cutting the gordian knot: A look under the hood of ransomware attacks. In *Detection of Intrusions and Malware, and Vulnerability Assessment: 12th International Conference, DIMVA 2015, Milan, Italy, July 9-10, 2015. Proceedings 12*, pages 3–24. Springer, 2015.
- [62] Danny Kim, Amir Majlesi-Kupaei, Julien Roy, Kapil Anand, Khaled ElWazeer, Daniel Buettner, and Rajeev Barua. Dynodet: Detecting dynamic obfuscation in malware. In Michalis Polychronakis and Michael Meier, editors, *Detection of Intrusions and Malware, and Vulnerability Assessment*, pages 97–118, Cham, 2017. Springer International Publishing.
- [63] Dhilung Kirat, Giovanni Vigna, and Christopher Kruegel. Barebox: efficient malware analysis on bare-metal. In *Proceedings of the 27th Annual Computer Security Applications Conference*, pages 403–412, 2011.
- [64] Dhilung Kirat, Giovanni Vigna, and Christopher Kruegel. Barecloud: Bare-metal analysis-based evasive malware detection. In *23rd USENIX Security Symposium (USENIX Security 14)*, pages 287–301, 2014.
- [65] Clemens Kolbitsch, Engin Kirda, and Christopher Kruegel. The power of procrastination: detection and mitigation of execution-stalling malicious code. In *Proceedings of the 18th ACM conference on Computer and communications security*, pages 285–296, 2011.
- [66] David Korczynski and Heng Yin. Capturing malware propagations with code injections and code-reuse attacks. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pages 1691–1708, 2017.
- [67] Igor Korkin. Protected process light is not protected: Memoryranger fills the gap again. In *2021 IEEE Security and Privacy Workshops (SPW)*, pages 298–308. IEEE, 2021.
- [68] Kenichi Kourai and Kousuke Nakamura. Efficient vm introspection in kvm and performance comparison with xen. pages 192–202, 11 2014.
- [69] Alexander Kuechler, Alessandro Mantovani, Yufei Han, Leyla Bilge, and Davide Balzarotti. Does Every Second Count? Time-based Evolution of Malware Behavior in Sandboxes. In *Network and Distributed System Security (NDSS) Symposium*, NDSS 21, February 2021.
- [70] Lawrence Abrams. US govt offers \$10 million bounty for info on Clop ransomware. <https://www.bleepingcomputer.com/news/security/us->

- govt-offers-10-million-bounty-for-info-on-clop-ransomware/, Accessed: 06-30-2023.
- [71] Seungkwang Lee, Nam su Jho, Doyoung Chung, Yousung Kang, and Myungchul Kim. Rcryptect: Real-time detection of cryptographic function in the user-space filesystem. *Computers & Security*, 112:102512, 2022.
- [72] Young Bi Lee, Jae Hyuk Suk, and Dong Hoon Lee. Bypassing anti-analysis of commercial protector methods using dbi tools. *IEEE Access*, 9:7655–7673, 2021.
- [73] Martina Lindorfer, Clemens Kolbitsch, and Paolo Milani Comporetti. Detecting environment-sensitive malware. In *Recent Advances in Intrusion Detection: 14th International Symposium, RAID 2011, Menlo Park, CA, USA, September 20-21, 2011. Proceedings 14*, pages 338–357. Springer, 2011.
- [74] Songsong Liu, Pengbin Feng, Shu Wang, Kun Sun, and Jiahao Cao. Enhancing malware analysis sandboxes with emulated user behavior. *Computers & Security*, 115:102613, 2022.
- [75] Songsong Liu, Pengbin Feng, Shu Wang, Kun Sun, and Jiahao Cao. Enhancing malware analysis sandboxes with emulated user behavior. *Computers & Security*, 115:102613, 2022.
- [76] Yushan Liu, Mu Zhang, Ding Li, Kangkook Jee, Zhichun Li, Zhenyu Wu, Junghwan Rhee, and Prateek Mittal. Towards a timely causality analysis for enterprise security. In *25th Annual Network and Distributed System Security Symposium, NDSS 2018, San Diego, California, USA, February 18-21, 2018*. The Internet Society, 2018.
- [77] Juan Lopez, Leonardo Babun, Hidayet Aksu, and A Selcuk Uluogac. A survey on function and system call hooking approaches. *Journal of Hardware and Systems Security*, 1:114–136, 2017.
- [78] R. Ma. Anomaly detection for linux system log, August 2020.
- [79] Shiqing Ma, Kyu Hyung Lee, Chung Hwan Kim, Junghwan Rhee, Xiangyu Zhang, and Dongyan Xu. Accurate, low cost and instrumentation-free security audit logging for windows. In *Proceedings of the 31st Annual Computer Security Applications Conference, ACSAC '15*, page 401–410, New York, NY, USA, 2015. Association for Computing Machinery.
- [80] Lorenzo Maffia, Dario Nisi, Platon Kotzias, Giovanni Lagorio, Simone Aonzo, and Davide Balzarotti. Longitudinal study of the prevalence of malware evasive techniques. *arXiv preprint arXiv:2112.11289*, 2021.
- [81] Davide Maiorca, Battista Biggio, and Giorgio Giacinto. Towards adversarial malware detection: Lessons learned from pdf-based attacks. *ACM Computing Surveys (CSUR)*, 52(4):1–36, 2019.
- [82] Mohd Fadzli Marhusin, Henry Larkin, Chris Lokan, and David Cornforth. An evaluation of api calls hooking performance. In *2008 International Conference on Computational Intelligence and Security*, volume 1, pages 315–319. IEEE, 2008.
- [83] Naif Mehanna, Walter Rudametkin, Pierre Laperdrix, and Antoine Vastel. Free proxies unmasked: A vulnerability and longitudinal analysis of free proxy services. *arXiv preprint arXiv:2403.02445*, 2024.
- [84] Shagufta Mehnaz, Anand Mudgerikar, and Elisa Bertino. Rvguard: A real-time detection system against cryptographic ransomware. In Michael Bailey, Thorsten Holz, Manolis Stamatogiannakis, and Sotiris Ioannidis, editors, *Research in Attacks, Intrusions, and Defenses*, pages 114–136, Cham, 2018. Springer International Publishing.
- [85] Microsoft. Pssetcreateprocessnotifyroutineex function (ntddk.h). <https://learn.microsoft.com/en-us/windows-hardware/drivers/ddi/ntddk/nf-ntddk-pssetcreateprocessnotifyroutineex>, 2022.
- [86] Microsoft. Pssetcreatethreadnotifyroutineex function (ntddk.h). <https://learn.microsoft.com/en-us/windows-hardware/drivers/ddi/ntddk/nf-ntddk-pssetcreatethreadnotifyroutineex>, 2022.
- [87] Microsoft. Pssetloadimagenotifyroutineex function (ntddk.h). <https://learn.microsoft.com/en-us/windows-hardware/drivers/ddi/ntddk/nf-ntddk-pssetloadimagenotifyroutineex>, 2022.
- [88] Microsoft. ntddk.h header. <https://learn.microsoft.com/en-us/windows-hardware/drivers/ddi/ntddk/>, 2023.
- [89] Microsoft. Overview of Early Launch AntiMalware. <https://learn.microsoft.com/en-us/windows-hardware/drivers/install/early-launch-antimalware>, Accessed: 10-15-2023.
- [90] Microsoft. Plug and Play Minor IRPs. <https://learn.microsoft.com/en-us/windows-hardware/drivers/kernel/plugin-and-play-minor-irps>, Accessed: 10-15-2023.
- [91] Microsoft Learn Challenge. Operations That Can Be IRP-Based or Fast I/O. <https://learn.microsoft.com/en-us/windows-hardware/drivers/ifs/operations-that-can-be-irp-based-or-fast-i-o>, Accessed:01-08-2025.
- [92] Microsoft Learn Challenge. Synchronous and Asynchronous I/O. <https://learn.microsoft.com/en-us/windows/win32/fileio/synchronous-and-asynchronous-i-o>, Accessed:01-08-2025.
- [93] Microsoft Learn Challenge. Windows Management Instrumentation API. <https://learn.microsoft.com/en-us/windows/win32/wmisdk/wmi-start-page>, Accessed:01-08-2025.
- [94] Aleksandar Milenkoski. *ELAM: The Windows Defender ELAM Driver*. PhD thesis, ERNW Enno Rey Netzwerke GmbH, 2019.
- [95] Najmeh Miramirkhani, Mahathi Priya Appini, Nick Nikiforakis, and Michalis Polychronakis. Spotless sandboxes: Evading malware analysis systems using wear-and-tear artifacts. In *2017 IEEE Symposium on Security and Privacy (SP)*, pages 1009–1024. IEEE, 2017.
- [96] Dimitris Mitropoulos and Diomidis Spinellis. Fatal injection: A survey of modern code injection attack countermeasures. *PeerJ Computer Science*, 3:e136, 2017.
- [97] Abhijit Mohanta, Anoop Saldanha, Abhijit Mohanta, and Anoop Saldanha. Code injection, process hollowing, and api hooking. *Malware Analysis and Detection Engineering: A Comprehensive Approach to Detect and Analyze Modern Malware*, pages 267–329, 2020.
- [98] KA Monnappa. Detecting deceptive process hollowing techniques usind hollowfind volatility plugin. <https://cysinfo.com/detecting-deceptive-hollowing-techniques/>, 2017.
- [99] Nedim Sabic. Fibratus. <https://github.com/rabbitstack/fibratus>, Accessed: 26-05-2024.
- [100] Matthew Nunes, Pete Burnap, Omer Rana, Philipp Reinecke, and Kaelon Lloyd. Getting to the root of the problem: A detailed comparison of kernel and user level data for dynamic malware analysis. *Journal of Information Security and Applications*, 48:102365, 2019.
- [101] Matthew Nunes, Pete Burnap, Philipp Reinecke, and Kaelon Lloyd. Bane or boon: Measuring the effect of evasive malware on system call classifiers. *Journal of Information Security and Applications*, 67:103202, 2022.
- [102] Odzhan. Another method of bypassing etw and process injection via etw registration entries. <https://modexp.wordpress.com/2020/04/08/red-teams-etw/>, 2020.
- [103] Ori Or-Meir, Nir Nissim, Yuval Elovici, and Lior Rokach. Dynamic malware analysis in the modern era—a state of the art survey. *ACM Computing Surveys (CSUR)*, 52(5):1–48, 2019.
- [104] osquery. osquery. <https://github.com/osquery/osquery>, Accessed: 26-05-2024.
- [105] François Plumerault and Baptiste David. Dbi, debuggers, vm: gotta catch them all! How to escape or fool debuggers with internal architecture cpu flaws? *Journal of Computer Virology and Hacking Techniques*, 17:105–117, 2021.
- [106] Mario Polino, Andrea Continella, Sebastiano Mariani, Stefano D’Alessio, Lorenzo Fontana, Fabio Gritti, and Stefano Zanero. Measuring and defeating anti-instrumentation-equipped malware. In *Detection of Intrusions and Malware, and Vulnerability Assessment: 14th International Conference, DIMVA 2017, Bonn, Germany, July 6-7, 2017, Proceedings 14*, pages 73–96. Springer, 2017.
- [107] David B Probert. Windows kernel internals cache manager. *Microsoft Corporation*, page 48, 2010.
- [108] RawSec. Whids. <https://github.com/0xrawsec/whids>, Accessed: 26-05-2024.
- [109] Donald Ray and Jay Ligatti. Defining code-injection attacks. *Acm Sigplan Notices*, 47(1):179–190, 2012.
- [110] Mark E Russinovich, David A Solomon, and Alex Ionescu. *Windows internals, part 2*. Pearson Education, 2012.
- [111] Nolen Scaife, Henry Carter, Patrick Traynor, and Kevin RB Butler. Cryptolock (and drop it): stopping ransomware attacks on user data. In *2016 IEEE 36th international conference on distributed computing systems (ICDCS)*, pages 303–312. IEEE, 2016.
- [112] Daniele Sgandurra, Luis Muñoz-González, Rabih Mohsen, and Emil C Lupu. Automated Dynamic Analysis of Ransomware: Benefits, Limitations and use for Detection. *arXiv preprint arXiv:1609.03020*, 2016.
- [113] Syed Shakir Hameed Shah, Abd Rahim Ahmad, Norziana Jamil, and Atta ur Rehman Khan. Memory forensics-based malware detection using computer vision and machine learning. *Electronics*, 11(16):2579, 2022.
- [114] Syed Zainudeen Mohd Shaid and Mohd Aizaini Maarof. In memory detection of windows api call hooking technique. In *2015 International conference on computer, communications, and control technology (I4CT)*, pages 294–298. IEEE, 2015.

- [115] Michael Sikorski and Andrew Honig. *Practical malware analysis: the hands-on guide to dissecting malicious software*. no starch press, 2012.
- [116] David A Solomon, Mark E Russinovich, and Alex Ionescu. *Windows internals*. Microsoft Press, 2009.
- [117] Chad Spensky, Hongyi Hu, and Kevin Leach. Lo-phi: Low-observable physical host instrumentation for malware analysis. In *NDSS*, 2016.
- [118] Jerre Starink, Marieke Huisman, Andreas Peter, and Andrea Continella. Understanding and measuring inter-process code injection in windows malware. In *19th International Conference on Security and Privacy in Communication Networks, SecureComm 2023*, 2023.
- [119] Steven White, Kent Sharkey, David Coulter, Dan Mabee, Drew Batchelor, Mike Jacobs, Michael Satran. About WMI. <https://learn.microsoft.com/en-us/windows/win32/wmisdk/about-wmi>, Accessed: 10-15-2023.
- [120] Ming-Kung Sun, Mao-Jie Lin, Michael Chang, Chi-Sung Lai, and Hui-Tang Lin. Malware virtualization-resistant behavior detection. In *2011 IEEE 17th International Conference on Parallel and Distributed Systems*, pages 912–917, 2011.
- [121] Ruimin Sun, Marcus Botacin, Nikolaos Sapountzis, Xiaoyong Yuan, Matt Bishop, Donald E. Porter, Xiaolin Li, Andre Gregio, and Daniela Oliveira. A praise for defensive programming: Leveraging uncertainty for effective malware mitigation. *IEEE Transactions on Dependable and Secure Computing*, 19(1):353–369, 2022.
- [122] Yixin Sun, Kangkook Jee, Suphanee Sivakorn, Zhichun Li, Cristian Lumezanu, Lauri Korts-Parn, Zhenyu Wu, Junghwan Rhee, Chung Hwan Kim, Mung Chiang, and Prateek Mittal. Detecting Malware Injection with Program-DNS Behavior. In *2020 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 552–568, Los Alamitos, CA, USA, September 2020. IEEE Computer Society.
- [123] Rashid Tahir, Muhammad Huzaifa, Anupam Das, Mohammad Ahmad, Carl Gunter, Fareed Zaffar, Matthew Caesar, and Nikita Borisov. Mining on someone else’s dime: Mitigating covert mining operations in clouds and enterprises. In *International Symposium on Research in Attacks, Intrusions, and Defenses*, pages 287–310. Springer, 2017.
- [124] Dmitry Tanana. Behavior-based detection of cryptojacking malware. In *2020 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBREIT)*, pages 0543–0545. IEEE, 2020.
- [125] The MITRE Corporation. MITRE ATTACK. <https://attack.mitre.org/>, Accessed: 09-20-2023.
- [126] wazuh. <https://github.com/wazuh>, Accessed: 26-05-2024.
- [127] Carsten Willems, Thorsten Holz, and Felix Freiling. Toward automated dynamic malware analysis using cwsandbox. *IEEE Security & Privacy*, 5(2):32–39, 2007.
- [128] Hongyan Xia, David Zhang, Wei Liu, Istvan Haller, Bruce Sherwin, and David Chisnall. A secret-free hypervisor: Rethinking isolation in the age of speculative vulnerabilities. In *2022 IEEE Symposium on Security and Privacy (SP)*, pages 370–385, 2022.
- [129] Yarden Shafir. ETW internals for security research and forensics. <https://blog.trailofbits.com/2023/11/22/etw-internals-for-security-research-and-forensics/>, Accessed: 01-08-2025.
- [130] Yazid. BEOTM. <https://github.com/Xacone/BestEdrOfTheMarket>, Accessed: 26-05-2024.
- [131] Haikuo Yin, Brandon Lou, and Peter Reiher. A method for summarizing and classifying evasive malware. In *Proceedings of the 26th International Symposium on Research in Attacks, Intrusions and Defenses*, pages 455–470, 2023.
- [132] Akira Yokoyama, Kou Ishii, Rui Tanabe, Yinmin Papa, Katsunari Yoshioka, Tsutomu Matsumoto, Takahiro Kasama, Daisuke Inoue, Michael Brengel, Michael Backes, and Christian Rossow. Sandprint: Fingerprinting malware sandboxes to provide intelligence for sandbox evasion. volume 9854, 09 2016.
- [133] Pavel Yosifovich, David A Solomon, Mark E Russinovich, and Alex Ionescu. *Windows Internals: System architecture, processes, threads, memory management, and more, Part 1*. Microsoft Press, 2017.
- [134] Michael Zipperle, Florian Gottwalt, Elizabeth Chang, and Tharam Dillon. Provenance-based intrusion detection systems: A survey. *ACM Computing Surveys*, 55(7):1–36, 2022.

A. Limitations

Section IV demonstrates that SHADOWBOX achieves practical and useful detection results on a large, real-world dataset. Unfortunately, adversaries continuously observe defensive advances and adapt their attacks accordingly. In the following, we discuss the limitations of SHADOWBOX and potential evasion strategies.

1) *Delay Injection.*: In case study 1, injecting a significant delay before loading the malicious payload is another method to bypass detection. The experiments were designed to record the behavior for five minutes, so if the malicious code stays dormant for a long period of time, SHADOWBOX cannot record actions associated with the malicious activity. Note that these limitations are not specific to SHADOWBOX, and almost any dynamic analysis system may be impacted in some way by these evasion methods. Prior work has discussed this evasion mechanism. Note that incorporating these techniques can also potentially make detection easier in static analysis since these approaches require calling specific functions in the operating system. The presence of these mechanisms in the initial binary is being used currently by malware defense solutions to identify suspicious binaries in the wild.

2) *Trusted Computing Model.*: SHADOWBOX operates at the kernel level. If the target malicious code is a kernel-level attack, it can potentially thwart some of the hooks SHADOWBOX uses to monitor run-time behavior. Analyzing high-privilege attacks is out of the scope of this paper. These attacks can be better analyzed in a hypervisor-based environment where debugging kernel-level code is possible. We assume that the trusted computing base includes the display module, OS kernel, and underlying software and hardware stack. Therefore, we can safely assume that the components of the system are free of malicious code and that normal user-based access control prevents attackers from running malicious code with superuser privileges. This is a fair assumption considering the fact that most malicious operations (e.g., malware attacks and software vulnerability checking) are often initiated in the user mode.

B. Reference Tables

TABLE IV: Distribution of a subset of 2.6 TBs of dataset submitted as artifacts in this submission. This represents 2% of the artifacts. Each log contains traces by SHADOWBOX from baremetal and the virtual environment.

Malware	Occurrences
Ransomware	1 (0.13%)
StopCrypt	1 (0.13%)
RAT	669 (83.83%)
Banload	1 (0.13%)
CoinMiner	35 (4.39%)
CryptInject	401 (50.25%)
Emotet	3 (0.38%)
Farfli	1 (0.13%)
Glupteba	1 (0.13%)
Killav	2 (0.25%)
Musecador	4 (0.50%)
Phonzy	1 (0.13%)
Plyromt	1 (0.13%)
Pykspa	1 (0.13%)
Sabsik	69 (8.65%)
Trickbot	1 (0.13%)
Wabot	18 (2.26%)
Wacatac	2 (0.25%)
Ymacco	128 (16.04%)
PUP	2 (0.25%)
GameBox	2 (0.25%)
Self Replicating Malware	126 (15.79%)
Andriod	5 (0.63%)
Cambot	6 (0.75%)
Canbis	5 (0.63%)
Floxif	1 (0.13%)
Gogo	10 (1.25%)
Morefi	5 (0.63%)
Neshta	5 (0.63%)
Nuqel	1 (0.13%)
Pidgeon	36 (4.51%)
Ramnit	1 (0.13%)
Sality	1 (0.13%)
Sfone	22 (2.76%)
Shodi	11 (1.38%)
Sivis	4 (0.50%)
Small	2 (0.25%)
Viking	1 (0.13%)
Virus	1 (0.13%)
Vobfus	9 (1.13%)
Total	798

TABLE V: Disk I/O performance in a standard and host running SHADOWBOX

Operation	Standard	SHADOWBOX	Overhead
Writer			
Small	788,641 KB/s	808,474 KB/s	2.51%
Large	893,921 KB/s	909,448 KB/s	1.74%
Re-writer			
Small	1,036,665 KB/s	1,05,9468 KB/s	2.20%
Large	1,054,297 KB/s	1,092,756 KB/s	3.65%
Reader			
Small	3,564,507 KB/s	3,492,892 KB/s	2.01%
Large	2,841,287 KB/s	2,928,931 KB/s	3.08%
Re-Reader			
Small	4,228,550 KB/s	4,452,886 KB/s	5.31%
Large	3,643,169 KB/s	3,791,870 KB/s	4.08%

TABLE VI: Comparing SHADOWBOX with current engines

Recorded Activities	SHADOWBOX	BestEDR [130]	Fibratus [99]	Osquery [104]	Wazuh [126]	BlueSpawn [52]	Whids [108]
Technology	kernel driver	API Hooking	ETW	App log collection	App log collection	Win32 API	ETW, Sysinternal
Run-time Behavior							
Process-level Information	system-wide	single proc.	system-wide	single proc.	single proc.	single proc.	system-wide
Thread-level Information	system-wide	no thread	system-wide	no thread	no thread	single proc.	system-wide
DLL Injection	system-wide	single proc.	not supported	not supported	not supported	not supported	not supported
Inter-process Communication	system-wide	not supported	not supported	not supported	not supported	not supported	not supported
I/O Activities							
Filesystem I/O	system-wide	not supported	system-wide	not supported	not supported	single proc.	system-wide
Synchronous I/O	system-wide	not supported	system-wide	not supported	not supported	single proc.	system-wide
Asynchronous I/O	system-wide	not supported	not supported	not supported	not supported	not supported	not supported
Fast I/O	system-wide	not supported	not supported	not supported	not supported	not supported	not supported
Analysis Agents							
Logging Service	kernelmode	usermode	usermode	usermode	usermode	usermode	usermode
Scheduler Agent	kernelmode	usermode	usermode	usermode	usermode	usermode	usermode

TABLE VII: SHADOWBOX artifacts for a multi-stage macro-malware attack (md5:7ccf88c0bbe3b29bf19d877c4596a8d4). The malware incorporates several techniques for fetching malicious payload, information grabbing, persistence, and camouflage.

#	Operation	Time	Global #	PPID	PID	TID	Image path	Command-line args	File path
1	Pr Create	20:51:45:628	183668	5480	10092	0	%MSOffice%\EXCEL.EXE	/dde	
2	Pr Create	20:51:48:141	205076	10092	1612	0	%MSOffice%\EXCEL.EXE	/Embedding	
3	Pr Create	20:52:16:301	346364	916	7028	0	%MSOfficeCommon%\eqned32.exe	-Embedding	
4	Pr Create	20:52:16:579	348340	7028	6344	0	%SysWOW64%\cmd.exe	/c REN mp\q v& WSCripT mp\v?.wsf C	
5	Pr Create	20:52:16:668	350587	6344	12148	0	%SysWOW64%\wscript.exe	%TEMP%\v?.wsf C	
6	Pr Create	20:52:17:084	358524	12148	4324	0	%SysWOW64%\cmd.exe	/c cscript %TEMP%\xx.vbs	
7	Pr Create	20:52:17:149	359955	4324	3800	0	%SysWOW64%\cscript.exe	%Temp%\xx.vbs	
8	Pr Exit	20:52:17:520	376991	916	7028	0	%MSOfficeCommon%\eqned32.exe		
9	Pr Exit	20:52:18:638	380165	10092	1612	0	%MSOffice%\EXCEL.EXE		
10	Pr Create	20:52:19:456	381227	916	11916	0	%SysWOW64%\wbem\WmiPrivSE.exe	-secured -Embedding	
11	Pr Create	20:53:05:751	407732	10092	5228	0	%SysWOW64%\cscript.exe	C:\programdata\asc.txt:script1.vbs	
12	Pr Exit	20:57:10:385	589717	5480	10092	0	%MSOffice%\EXCEL.EXE		
13	Pr Create	20:57:42:131	666687	916	7660	0	%MSOfficeCommon%\eqned32.exe	-Embedding	
14	Pr Create	20:57:42:305	668655	7660	2128	0	%SysWOW64%\cmd.exe	/c REN mp\q v& WSCripT mp\v?.wsf C	
15	Pr Create	20:57:42:358	671600	2128	5368	0	%SysWOW64%\wscript.exe	%Temp%\v?.wsf C	
16	Pr Create	20:57:42:399	675850	7660	10328	0	%SysWOW64%\WerFault.exe	-u -p 7660 -s 1172	
17	Pr Create	20:57:42:474	678590	5368	12032	0	%SysWOW64%\cmd.exe	/c cscript %Temp%\xx.vbs	
18	Pr Create	20:57:42:516	679047	12032	10464	0	%SysWOW64%\cscript.exe	%Temp%\xx.vbs	
19	Ld Image	20:57:42:527	679237	12032	10464	0	%SysWOW64%\cscript.exe		%SysWOW64%\bcryptprimitives.dll
20	Ld Image	20:57:43:645	705589	12032	10464	0	%SysWOW64%\cscript.exe		%SysWOW64%\winhttp.dll
21	Ld Image	20:57:43:652	705594	12032	10464	0	%SysWOW64%\cscript.exe		%SysWOW64%\mswsock.dll
22	Tr Create	20:57:44:210	707985	12032	10464	2844	%SysWOW64%\cscript.exe		
23	Ld Image	20:57:44:226	708195	12032	10464	0	%SysWOW64%\cscript.exe		%SysWOW64%\urlmon.dll
24	Ld Image	20:57:44:232	708207	12032	10464	0	%SysWOW64%\cscript.exe		%SysWOW64%\virdisk.dll
25	Tr Exit	20:57:44:249	708455	12032	10464	2844	%SysWOW64%\cscript.exe		

(a) SHADOWBOX artifacts on process activities.

#	Operation	Time	Duration	Global #	PID	TID	Image path	File Path
1	IRP_Create	20:51:45:636	69	183856	10092	1504	%MSOffice%\EXCEL.EXE	C:\Users\grace\Downloads\ORDER SHEET & SPEC.xlsm
2	IRP_Write	20:51:48:590	1757	211615	10092	3620	%MSOffice%\EXCEL.EXE	C:\Users\grace\AppData\Local\Temp\DED9E0FE.xlsm
3	IRP_Write	20:51:48:755	448	213650	1612	12240	%MSOffice%\EXCEL.EXE	C:\Users\grace\AppData\Local\Packages\...AC\Temp\560D285B.emf
4	IRP_Write	20:52:13:024	501	296839	10092	1504	%MSOffice%\EXCEL.EXE	C:\Users\grace\AppData\Local\Microsoft\...1983A0E7.png
5	IRP_Write	20:52:16:156	468	345357	10092	1504	%MSOffice%\EXCEL.EXE	C:\Users\grace\AppData\Local\Microsoft\Windows\...1959A28D.emf
6	IRP_Write	20:52:16:230	886	345950	10092	1504	%MSOffice%\EXCEL.EXE	C:\Users\grace\AppData\Local\Temp\q
7	IRP_Write	20:52:16:270	551	346113	10092	1504	%MSOffice%\EXCEL.EXE	C:\Users\grace\AppData\Local\Temp\xx
8	IRP_Set_Information	20:52:16:932	748	355451	12148	8056	%SysWOW64%\wscript.exe	C:\Users\grace\AppData\Local\Temp\xx
9	IRP_Close	20:52:16:933	17	355455	12148	8056	%SysWOW64%\wscript.exe	C:\Users\grace\AppData\Local\Temp\xx.vbs
10	IRP_Write	20:53:05:692	385	407107	10092	1504	%MSOffice%\EXCEL.EXE	C:\ProgramData\asc.txt:script1.vbs
11	IRP_Read	20:57:42:541	73	679583	10464	2584	%SysWOW64%\cscript.exe	C:\Users\grace\AppData\Local\Temp\xx.vbs
12	IRP_Write	20:57:44:237	3432	708409	10464	2844	%SysWOW64%\cscript.exe	C:\ProgramData\Podaliri4.exe
13	IRP_Close	20:57:44:248	20	708591	10464	2844	%SysWOW64%\cscript.exe	C:\ProgramData\Podaliri4.exe

(b) SHADOWBOX artifacts on filesystem activities.

TABLE VIII: Top dropped extensions in the two experiments. A malware sample tends to drop more files in SHADOWBOX when compared to a virtualized environment.

Extension	SHADOWBOX	Virtualized Environment	Difference (# — %)	Description
Archive	12,309	9,195		
cab	3,504	1,989	1,515 — 76.2%	Microsoft cabinet file - compressed archive
pak	8,805	7,206	1,599 — 22.2%	Game Archive/Skype Language Pack
Binary	3,132,3425	1,402,507		
api	22,594	7,480	15,114 — 202.1%	Adobe Acrobat Plug-in/WebObjects API File
appx	2,553	557	1,996 — 358.3%	Microsoft Windows 8 app package
bin	4,891	2,053	2,838 — 138.2%	Binary archive
cur	6,772	2,070	4,702 — 227.1%	Windows custom cursor
dat	5,600	3,156	2,444 — 77.4%	General data file
dll	333,946	71,361	262,585 — 368.0%	Windows Dynamic Linked Library
exe	2,160,744	1,114,617	1,046,127 — 93.9%	Executable file
js	586,494	197,952	388,542 — 196.3%	JavaScript file
pmp	2,674	917	1,757 — 191.6%	AutoCAD plot model parameter file
pyd	1,278	789	489 — 62.0%	Windows binary that contains compiled Python code
sequ	4,879	1,555	3,324 — 213.8%	Adobe Acrobat Batch sequence
Document	74,340	34,706		
html	32,586	13,380	19206 — 143.5%	Hypertext Markup Language Document
mpp	2,888	924	1,964 — 212.6%	Microsoft Project document
pdf	19,595	13,321	6,274 — 47.1%	Portable Document Format Document
rtf	1,355	910	445 — 48.9%	Rich Text Format document
x3d	4,662	1,391	3,271 — 235.2%	X3D (XML) scene to represent 3D graphics
xml	13,254	4,780	8,474 — 177.3%	Extensible Markup Language File
Font	30,719	10,164		
eot	2,697	871	1,826 — 209.6%	Microsoft Embedded OpenType font
otf	25,360	7,586	17,774 — 234.3%	Font file
woff	2,662	1,707	955 — 55.9%	Web Open Font Format
Image	735,210	306,284		
bmp	1,383	169	1,214 — 718.3%	Bitmap image
gif	75,298	24,718	50,580 — 204.6%	Bitmap image
ico	28,806	18,479	10,327 — 55.9%	Icon file
jpg	16,865	5,462	11,403 — 208.8%	JPEG Bitmap image
png	266,327	136,027	130,300 — 95.8%	Portable Network Graphics image
svg	346,531	121,429	225,102 — 185.4%	Scalable Vector Graphic
Raw	328,714	106,443		
aapp	35,181	10,773	24,408 — 226.6%	Adobe Acrobat AcroApp script
css	46,271	15,218	31,053 — 204.1%	Cascading style sheets
dic	1,873	627	1,246 — 198.7%	Microsoft Office custom dictionary
ini	13,668	4,937	8,731 — 176.8%	Initialization file
json	21,681	1,430	20,251 — 1416.2%	JavaScript Object Notation
lnk	1,160	912	248 — 27.2%	Windows shortcut
log	15,285	7,987	7,298 — 91.4%	General log
tmp	84,564	34,789	49,775 — 143.1%	General temporary file
txt	109,031	29,770	79,261 — 266.2%	Plain text document
Audio	2,924	2,843		
wav	2,924	81	2,843 — 3509.9%	Windows waveform sound

TABLE IX: I/O Request Packets (IRPs) Types supported by SHADOWBox

Major	Description	Minor	Description
Standard		IRP_MN_REGINFO	Driver registry path
IRP_MJ_CREATE	Open object handle	IRP_MN_QUERY_DIRECTORY	Get files in directory
IRP_MJ_CLOSE	Create or open named pipe	IRP_MN_NOTIFY_CHANGE_DIRECTORY	Notify directory changes
IRP_MJ_READ	Close open handle	IRP_MN_USER_FS_REQUEST	Filesystem requests
IRP_MJ_WRITE	Data read	IRP_MN_MOUNT_VOLUME	Requested volume mount
IRP_MJ_QUERY_INFORMATION	Data write	IRP_MN_VERIFY_VOLUME	Mounted volume integrity check
IRP_MJ_SET_INFORMATION	Query object information	IRP_MN_LOAD_FILE_SYSTEM	Load driver from filesystem
IRP_MJ_QUERY_EA	Modify object information	IRP_MN_TRACK_LINK	Watch for device events
IRP_MJ_SET_EA	Query object extended attributes	IRP_MN_LOCK	Acquire file lock
IRP_MJ_FLUSH_BUFFERS	Set object extended attributes	IRP_MN_UNLOCK_SINGLE	Release file lock
IRP_MJ_QUERY_VOLUME_INFORMATION	Flush cached data to disk	IRP_MN_UNLOCK_ALL	Release all file locks
IRP_MJ_SET_VOLUME_INFORMATION	Query volume information	IRP_MN_UNLOCK_ALL_BY_KEY	Release locks associated with key
IRP_MJ_DIRECTORY_CONTROL	Modify volume property	IRP_MN_NORMAL	Perform device management tasks via messages
IRP_MJ_FILE_SYSTEM_CONTROL	Directory management operations	IRP_MN_DPC	Deferred procedure calls (DPC)
IRP_MJ_DEVICE_CONTROL	Advance file system operations	IRP_MN_MDL	Manage memory descriptors lists (MDL)
IRP_MJ_INTERNAL_DEVICE_CONTROL	Use device driver functions	IRP_MN_COMPLETE	Mark filesystem operation complete
IRP_MJ_SHUTDOWN	Midware between device and kernel	IRP_MN_COMPRESSED	Read and transfer compressed data
IRP_MJ_LOCK_CONTROL	System shutdown	IRP_MN_MDL_DPC	Merge DPC with completion alert
IRP_MJ_CLEANUP	Managing byte-range locks	IRP_MN_QUERY_ALL_DATA	Retrieve comprehensive device information
IRP_MJ_CREATE_MAILSLOT	Release process file resources	IRP_MN_COMPLETE_MDL_DPC	Merge DPC, MDL with completion alert
IRP_MJ_QUERY_SECURITY	Manage mailslots for communication	IRP_MN_SCSL_CLASS	Manage SCSI devices
IRP_MJ_SET_SECURITY	Query security info about objects	IRP_MN_START_DEVICE	Starting and initializing devices
IRP_MJ_POWER	Set object security information	IRP_MN_QUERY_REMOVE_DEVICE	Pending device removal
IRP_MJ_SYSTEM_CONTROL	Manage device power state	IRP_MN_CANCEL_REMOVE_DEVICE	Pending device removal tasks
IRP_MJ_DEVICE_CHANGE	Request system-wide operations from OS	IRP_MN_STOP_DEVICE	Cancel device removal
IRP_MJ_QUERY_QUOTA	Manage device events	IRP_MN_QUERY_STOP_DEVICE	Query stop device possibility
IRP_MJ_PNP	Manage device quota	IRP_MN_CANCEL_STOP_DEVICE	Cancel stop device
IRP_MJ_TRANSACTION_NOTIFY	Volume quota information	IRP_MN_QUERY_DEVICE_RELATIONS	Discover device relationships
Fast IO	Manage operations notifications	IRP_MN_QUERY_INTERFACE	Supported device interfaces
IRP_MJ_FAST_IO_CHECK_IF_POSSIBLE	Check object fast I/O capability	IRP_MN_SET_LOCK	Manage object lock requests
IRP_MJ_DETACH_DEVICE	Detach device and free resources	IRP_MN_QUERY_CAPABILITIES	Query device capabilities
IRP_MJ_NETWORK_QUERY_OPEN	Query network connection information	IRP_MN_QUERY_RESOURCES	Query device resource requirements
IRP_MJ_MDL_READ	Read from file or device via MDL	IRP_MN_QUERY_RESOURCE_REQUIREMENTS	Query device resource requirements
IRP_MJ_MDL_READ_COMPLETE	Notify MDL object read completion	IRP_MN_QUERY_DEVICE_TEXT	Query device description and location
IRP_MJ_PREPARE_MDL_WRITE	Prepare file for MDL write	IRP_MN_FILTER_RESOURCE_REQUIREMENTS	Modify device resource requirement
IRP_MJ_MDL_WRITE_COMPLETE	Complete and notify of MDL write	IRP_MN_READ_CONFIG	Read connected device configuration
IRP_MJ_VOLUME_MOUNT	Mount a volume	IRP_MN_WRITE_CONFIG	Write connected device configuration
IRP_MJ_VOLUME_DISMOUNT	Dismount a volume	IRP_MN_EJECT	Remove device from system
FsFilter		IRP_MN_DISABLE_COLLECTION	Prevent automatic resource collection for device
IRP_MJ_ACQUIRE_FOR_SECTION_SYNCHRONIZATION	Acquire memory block for synchronization	IRP_MN_EXECUTE_METHOD	Invoke device functions via its driver
IRP_MJ_RELEASE_FOR_SECTION_SYNCHRONIZATION	Release sync memory block	IRP_MN_QUERY_ID	Uncover the identity of a device
IRP_MJ_ACQUIRE_FOR_MOD_WRITE	Acquire memory section for modifier write	IRP_MN_QUERY_PNP_DEVICE_STATE	Query PnP device state
IRP_MJ_RELEASE_FOR_MOD_WRITE	Release modifier write memory	IRP_MN_QUERY_BUS_INFORMATION	Query connected bus information
IRP_MJ_ACQUIRE_FOR_CC_FLUSH	Acquire memory for cache coherency flush	IRP_MN_DEVICE_USAGE_NOTIFICATION	Notify of device power usage
IRP_MJ_RELEASE_FOR_CC_FLUSH	Release memory for cache coherency flush	IRP_MN_SURPRISE_REMOVAL	Notify of unexpected device removal
IRP_MJ_NOTIFY_STREAM_FO_CREATION	Notify driver of file create	IRP_MN_QUERY_LEGACY_BUS_INFORMATION	Query legacy bus hardware information
		IRP_MN_WAIT_WAKE	Enable wake from low-power state
		IRP_MN_POWER_SEQUENCE	Manage device power during transitions
		IRP_MN_SET_POWER	Change power state of device
		IRP_MN_QUERY_POWER	Query ability to move to a named power state
		IRP_MN_QUERY_SINGLE_INSTANCE	Get info about named instance of device
		IRP_MN_CHANGE_SINGLE_INSTANCE	Modify named instance of device
		IRP_MN_ENABLE_SINGLE_ITEM	Modify named device property
		IRP_MN_DISABLE_EVENTS	Enable specific event generation from device
		IRP_MN_ENABLE_COLLECTION	Disable device from events generation
			Enable automatic resource collection for device