Virtualization and Security Boundaries

by Mike Lococo

Introduction

This guide provides a framework for discussing the security ramifications of virtualization with regard to the enforcement of security boundaries. It presents a taxonomy of virtualization technologies that is helpful in security analysis, an overview of attacks that are possible in virtualized environments, a summary of current best practices, and a list of unresolved challenges.

Scope

This guide is the result of a survey of existing public literature. It does not present original technical research. It is also not a hardening guide for any particular virtualization technology, although VMWare ESX and related products do receive some special focus due to their current dominance in the market.

Feedback and Corrections

This document may contain errors. If you find one, please send feedback to mikelococo at gmail dot com.

Citations are provided to support statements which may be controversial whenever possible, but some assertions are supported primarily by first-hand experience detecting and responding to incidents on research and education networks. If you are aware of citations which support or challenge any of the assertions made in this guide, please send feedback.

The most recent version of this guide can be found at:

http://mikelococo.com/2009/06/virtualization-and-security-boundaries/

This version was last updated on 06/07/09.

Copyright

Virtualization as Security Boundary Enforcement

A broad definition provides a starting point for developing a comprehensive virtualization strategy:

> "Virtualization is a framework or methodology of dividing the resources of a computer into multiple execution environments, by applying one or more concepts or technologies such as hardware and software partitioning, time-sharing, partial or complete machine simulation, emulation, quality of service, and many others."[1]

The above definition can be summarized somewhat sloppily by saying that virtualization abstracts the operating system from the physical hardware on which it runs. Security practitioners have long held that the security boundary between two instances of an operating system is very strong. The perceived strength of this boundary is due in part to its enforcement by the physical separation of hardware, which strictly limits the types of attacks a malicious user can leverage to expand their scope of privilege. The introduction of virtualization into an environment changes the enforcement of that boundary in a fundamental way, and organizations must analyze and understand that change in order to determine how virtualization affects their security architecture.

On the other hand, security boundaries cannot always be enforced by physical separation. It would be impossibly expensive and complex to manage a network where the only trusted method of boundary enforcement was physical separation. Instead, organizations employ a variety of enforcement methods such as:

- Separation of application user privileges enforced by application logic
- Separation of operating-system user privileges enforced by the kernel
- Separation of operating systems enforced by physical separation of hardware
- Network separation enforced by network-firewalls

In order to reduce the complexity that results from employing many different types of security boundaries of varying strength, organizations typically classify their data and systems and issue guidelines about the enforcement of boundaries between resources of different classifications[2]. Organizations that have formalized the classification and separation of IT resources can successfully address virtualization as one of many boundary enforcement tools, and can issue guidelines that specify what security boundaries may (or may not) be enforced by virtualization technology. Organizations which have informal or poorly articulated guidelines on separation must navigate the risk of virtualization deployments that traverse those informal boundaries in ways that disrupt previous assumptions about boundary strength.

The remainder of this guide will present a framework for analyzing the relative ability of different virtualization platforms to provide security boundary enforcement, through the isolation of guest operating systems from each other and from the host under which they run.

[1] http://www.kernelthread.com/publications/virtualization/

[2] http://www.educause.edu/Resources/DataClassificationandPrivacyAF/162709

Beyond VMWare – A Taxonomy of Virtualization Technologies

This section presents a classification system which is useful for describing the rigor with which a given approach to virtualization enforces security boundaries. Although there are other existing taxonomies of virtualization technologies, they typically focus on performance or implementation details and are not suitable for classifying virtualization technologies according to their security properties.

Hardware Partitioning

Hardware-based partitioning is employed on expensive "enterprise" computing platforms such as Sun Fire/Enterprise systems and IBM P-series systems, and is not available on x86 systems. It is characterized by tight coupling of the partitioning primitives available to system administrators with hardware-based mechanisms to enforce guest isolation.

The following features or limitations may indicate that a hardware-based partitioning strategy is being employed:

- The virtualization technology is tied to a specific hardware platform (due to tight integration with hardware-based isolation mechanisms) and is not available on x86.
- Multiple physical computing and I/O resources are present in the system that are dedicated and assigned to guests rather than being virtualized and shared.
- Resource assignment is static and inflexible.
- Electrical faults are isolated between guests[3].

One example of hardware-based partitioning is Sun's Dynamic System Domains (DSDs) on Sun Fire hardware. Each guest is comprised of one or more system boards that contain separate CPU, memory, I/O, boot-disk and network resources[4]. These system boards are assigned to a particular guest, rather than shared between them, and isolation is enforced via electrical separation; the connections between different guests are physically severed.

Hardware partitioning schemes tend to provide very strong isolation between guests. An attacker must be well-resourced to have access to a test platform on which to design exploits for hardware partitioning technologies. Additionally, since many of the capabilities used to enforce guest isolation are implemented in hardware, the attack surface available to malicious software programs is comparatively small.

Software Partitioning

Software-based partitioning strategies are employed by many commodity server and workstation virtualization products. Many products employ this approach:

- VMWare ESX, Server, and Workstation
- Xen
- Linux Kernel-based Virtualization Machine (KVM)
- Parallels
- Microsoft Hyper-V

Although all virtualization strategies employ hardware and software in combination to enforce

[3] http://www.repton.co.uk/library/server_workload_consolidation.pdf

[4] http://www.sun.com/servers/white-papers/domains.html

guest isolation, software partitioning strategies are characterized by the general absence of hardware-based enforcement mechanisms. Guest isolation is enforced primarily by a software layer between the guest operating system kernels and the physical hardware.

The presence of the following features or limitations may suggest that a virtualization or partitioning technology utilizes a primarily software-based partitioning strategy:

- The virtualization technology is portable to multiple hardware platforms (due to loose coupling with hardware enforcement capabilities) or is available on x86.
- The physical computing and I/O resources available are not typically sufficient to assign a dedicated set to each guest operating system, so they must be virtualized and shared.
- Resource assignment is flexible, dynamic and automatic reallocation of arbitrary fractions of resources is common.
- Virtual network switching is present, or it is possible to create complex virtual network topologies.

Generally speaking, virtualization technologies that employ a primarily software-based partitioning strategy provide weaker guest isolation than their hardware-based counterparts. Device virtualization adds considerable complexity to virtualization infrastructure, and the robustness of guest isolation may vary considerably depending on the hardware platform and the specifics of the software partitioning implementation.

Single Kernel Partitioning

Single kernel partitioning strategies are employed primarily in lightweight server-consolidation technologies. Examples of single kernel technologies include:

- chroot jails
- BSD Jails
- User-Mode Linux
- Sun Zones/Containers

Single kernel partitioning is characterized by the use of single operating system kernel which is shared among the host and all guests. No hardware-based mechanisms to enforce guest isolation are used, and hardware is not virtualized since only a single operating system kernel is present. Guest isolation is enforced by kernel-level isolation primitives which provide each guest-application with the appearance of an isolated operating system instance.

The presence of the following features or limitations may suggest that a single kernel partitioning strategy is employed:

- All guests must run the same operating system kernel as the host, even if the underlying hardware is capable of running alternative kernel software.

Generally speaking, single kernel partitioning strategies provide the weakest guest isolation of the three strategies discussed in this paper. Although they have the advantage of a comparatively simple and auditable codebase, the shared operating system kernel is typically attackable from the network and provides a single point of failure from which guest isolation can be compromised.

Attacks in Virtualized Environments

When discussing the security merits of various virtualization strategies, it is useful to have in mind the types of attacks which are possible. Almost any attack which is possible in a physical environment is also possible in a virtual one, and there are some additional types of attacks that are only feasible in a virtual environment. The following list is adapted from presentations made by Thomas Ptacek[5].

Jailbreak Attacks, aka Escapes

Any action which results in a user or administrator of one guest gaining unauthorized access to the underlying host or to a different guest may be characterized as a jailbreak attack. They are not possible in a non-virtualized environment because they typically exploit the virtualization infrastructure itself, or weaknesses in guest isolation that have no counterparts in purely physical deployments. Attacks typically require the attacker to have user or administrative access to a guest. Upon success the attacker gains the ability to access memory or execute code on the host or on a different guest, which they are not authorized to access. For example, an attacker with administrative access to a guest could send malformed SCSI commands to a virtual SCSI disk controller in an attempt to exploit device virtualization code. An attacker with user-level access to a guest could abuse or exploit legitimate information-sharing tools such as shared-folder or drag-and-drop functionality[6,7].

At the time of this writing, there are no publicly known examples of jailbreak attacks being used in the wild against any major virtualization platform. Whether such attacks will become common in the next few years is a matter of speculation, but the rapid and widespread adoption of virtualization technology will make it an increasingly attractive target for attackers. Additionally, there is a substantial amount of security research activity which has yet to be weaponized:

- VMWare ESX and other well-known products that employ software partitioning have had escape vulnerabilities announced and patched[8,9,10,11]. For example, VMWare announced an escape vulnerability in April 2009 for ESX that would allow a guest to run arbitrary code in the context of the host[12]. It was weaponized into an exploit by Immunity for Workstation[13], but not for ESX. VMWare marketing materials often note that as of yet, none of the escape vulnerabilities for ESX have been weaponized.
- Tavis Ormandy presented a method to test the robustness of virtualization platforms by generating random byte activity on I/O ports and in program instruction streams[14]. His results demonstrated the presence of crash-bugs in every virtualization platform tested, including VMWare Workstation. These bugs were not studied to verify their

[5] http://searchsecurity.bitpipe.com/detail/RES/1213273947_134.html
[6] http://www.computerworld.com/action/article.do?command=viewArticleBasic&articleId=9064319&source=rss_news50
[7] http://www.foolmoon.net/cgi-bin/blog/index.cgi?mode=viewone&blog=1185593255
[8] http://www.securityfocus.com/bid/29183/info
[9] http://secunia.com/advisories/29129/
[10] http://seclists.org/fulldisclosure/2007/Sep/0355.html
[11] http://lists.vmware.com/pipermail/security-announce/2009/000055.html
[12] Ibid.
[13] http://www.immunityinc.com/documentation/cloudburst-vista.html
[14] http://taviso.decsystem.org/virtsec.pdf

exploitability, but the number of faults generated by the fuzzing exercise indicate that it is an area of research which is likely to uncover exploitable flaws.

Migration Attacks

Many virtualization technologies provide the capability to move a guest from one physical host to another physical host, either to adapt to changes in capacity requirements or to enable hardware maintenance without guest downtime. Attacks against this process are characterized as migration attacks. Migration attacks are possible only in virtualized environments, as they exploit virtualization infrastructure. Typical attacks require the attacker to have access to the network where migrations are occurring, and success results in the ability to arbitrarily read and modify the memory and disk contents of the guest while in transit.

At the time of this writing, there are no known public examples of migration attacks occurring in the wild, however researchers at the University of Michigan have built exploit tools for VMWare and Xen and have shown that a number of high-impact attacks are possible[15]:

- An attacker may exploit weak or missing integrity checks to trojan the guest in transit, incorporating their own code into the memory image which will be run by the guest when it arrives at its new host. For example, researchers trojaned a running sshd process during migration such that root logins would succeed without requiring a password or public-key credential.
- An attacker may exploit the absence of encryption to sniff confidential information in transit, such as encryption keys stored in guest memory.
- An attacker may exploit weak or missing authentication of the migration capability to induce migration of a guest to a malicious host. Once there, the attacker may arbitrarily read and modify the contents of memory and disk to extract confidential information or compromise the integrity of the guest.

Client-Side Attacks

Client-side attacks leverage common desktop software such as web-browsers, email-clients, and media-players. They are extremely common in both virtual and physical environments, and won't be discussed in detail here. It should be noted, though, that payloads leveraging virtualization management tools or scripting interfaces could have an extremely high impact if successfully executed on the workstation of a virtualization administrator.

Virtual or Physical Network Service Attacks

Network service attacks occur when a remote attacker gains unauthorized access to a guest or host through exploitation of a listening service. They are extremely common in both virtual and physical environments, and won't be discussed in detail here. However, virtualization hosts typically run network accessible services which must be protected as any high-value service would be. Additionally, snapshotting and rollback capabilities can introduce previously patched service vulnerabilities into guests. Patching systems should verify the current state of patch-clients, and not assume that systems that were patched once will remain so. Finally, system administrators should ensure that services which may be vulnerable after a rollback are patched before being exposed to potentially hostile networks.

[15] http://www.eecs.umich.edu/techreports/cse/2007/CSE-TR-539-07.pdf

Encryption Attacks

Encryption attacks have as their goal the recovery of encrypted data, often by recovering the key used to encrypt the data. The types of encryption attacks described here are not commonly attempted in the wild, there are typically far simpler ways to get unauthorized access to data than to use sophisticated encryption attacks. However, virtualization can decrease the effort required to compromise encryption keys significantly.

- Side-channel encryption attacks: By observing the system load, timing, and other characteristics of the system during encryption, an attacker may be able to recover encryption keys much more quickly and easily than would be possible using traditional “brute-forcing” methods[16]. Because hardware is shared among the host and multiple guests, it may be possible for an attacker with user-level access to a single guest to recover encryption keys used by the host or other guests.
- Replay encryption attacks: Many virtualization products include the capability to perform “snapshotting”, where the memory, disk, and cpu register state of a running guest is saved for restoration at a later time. Many encryption algorithms rely on the linear forward progression of time to ensure secrecy[17]. For example, pseudo-random number generators may repeat their output after rollback. This can allow an attacker to deduce the private key used to encrypt data in some public key encryption methods, including DSS. Repeated pseudo-random number generator output can also lead to predictable session keys in a variety of encryption technologies. Additionally, one-time passwords collected by an attacker may be reusable after rollback.
- Live migration attacks: As discussed previously, an attacker with sufficient network access can extract encryption keys in guest memory by sniffing migration traffic.

Best Practices

Like all complex systems, virtualization platforms are imperfect and can have exploitable security flaws. This guide has reviewed different categories of virtualization systems and the differing levels of guest isolation they provide. It has also described various attacks against virtualization systems, and cited proof-of-concept exploit tools or vulnerabilities for many major virtualization products, including VMWare ESX. Successful organizations will treat virtualization as part of a multi-layer partitioning approach, and seek to limit the impact of a security failure in any one piece. The best practices listed in this section serve that goal.

Classify Data and Systems

The primary security function of a virtualization platform is to enforce security boundaries between different guests, and between guests and the host. Security practitioners must formalize where important boundaries are in order to integrate virtualization into their security architecture.

- Classify data according to the risk of a breach in the confidentiality, integrity, accessibility, or auditability of that data.
- Classify virtualized guest operating systems according to the data that they store and process.

[16] http://www.schneier.com/crypto-gram-9806.html#side

[17] http://www.stanford.edu/~talg/papers/HOTOS05/virtual-harder-hotos05.pdf

- Classify hosts according to the guests that run on them. Group hosts such that only appropriately classified guests are migrated onto them, VMWare provides groupings such as “clusters” and “data-centers” that can be used for this purpose.
- Classify networks and storage appropriately. Use separate vlans and luns for systems with different classifications.

Evaluate Virtualization Products Before Deployment

Perform a risk assessment on new virtualization technologies before deploying them. Determine whether they take a hardware, software, or single kernel approach, and determine what boundaries they can be trusted to enforce. Periodically review these risk assessments to ensure they remain appropriate in relation to rapidly evolving product and threat landscapes.

Harden Virtualization Infrastructure

Virtualization infrastructure should be classified at least as high as the highest classified guest it runs, and should be protected appropriately. Many excellent hardening guides exist:

- VMWare Security Hardening Whitepaper[18]
- Tripwire Configcheck[19]
- Center for Internet Security ESX Server Benchmark[20]
- Virtual Computing STIG[21]

I will not repeat the contents of those hardening guides here, but will provide a short list of what I consider to be some of their key recommendations:

- Remove unnecessary virtual hardware from guests
- Disable copy/paste, mouse-takeover, file-sharing in VMWare tools
- Use signed certificates for Virtual Infrastructure clients to protect management traffic
- Send logs to a non-virtual remote host to facilitate forensics in the event of a compromise
- Configure virtual switches to reject MAC changes and forged transmissions
- Do not create a “Default Port Group” on VMWare hosts

Segregate Insecure Networks

Virtualization products often have poorly protected backend communications. Guest migration traffic (VMotion traffic in VMWare) and storage traffic tend to be unencrypted and weakly authenticated. Ensure that these insecure communications are segregated from guest traffic. An alternative to segregation is hardening via external encryption like IPSec, although this may have a performance impact and is not supported by all vendors. For example, VMWare ESX does not currently support IPSec.

[18] http://www.vmware.com/resources/techresources/726
[19] http://www.vmware.com/security/resources/configcheck.html
[20] http://www.cisecurity.org/bench_vm.html
[21] Http://iase.disa.mil/stigs/draft-stigs/Virtual-Computing-STIG-V1R01.doc

Challenges and Open Questions

Virtualization is a rapidly evolving sector of the industry, and standards for best-practices have not yet been developed for many issues. This section identifies key open questions on which broad agreement has not yet been reached.

Containing a Compromised Host

If used to consolidate systems of different classifications, virtualization platforms have the potential to aggregate an unprecedented amount of risk into a single system component. It is desirable, therefore, for that component to employ a layered security strategy that allows faults to be contained to a single host or set of hosts, even if those hosts are compromised.

Unfortunately, most vendors are currently focused on hardening the perimeter of their platforms and do not consider the containment of a malicious host to be a viable threat scenario. VMWare's work on the VMSafe API is an early step toward improving detection of compromised guests, but at this time no vendor using a software-based partitioning approach appears to have engaged the issue of containing a malicious host, or an attacker with access to the migration and storage networks. The inability to contain security faults, in combination with the large numbers of systems and sensitive data-elements in use in large virtualization deployments, means that a compromise of virtualization infrastructure could have an extremely high impact and associated recovery cost.

Security Sensors and Monitoring

Many organizations employ network-based security monitoring to gather trusted data about the activity of potentially compromised or malicious hosts, but virtual networking infrastructure may not provide the monitoring capabilities necessary to do so for virtual networks. Organizations that perform security monitoring primarily at their borders, and who have little or no visibility into activity on internal networks will not experience a loss of monitoring capability. Organizations that do perform management and monitoring of internal networks will encounter challenges maintaining those standards in virtualized environments, although VMWare does have plans to expand the range management and monitoring products available to customers on its platform.

Virtual switches do not currently provide management features which are common in physical switching products. VMWare has partnered with Cisco to produce managed switches that will integrate more seamlessly with Cisco management tools[22], but organizations that use other switching vendors may have trouble integrating those products into their environments.

Traditional network monitoring tools like flow-monitoring and intrusion detection/prevention cannot currently observe packets that traverse only the virtual switching infrastructure. VMWare is introducing the VMSafe API to enable such inspection by virtual security appliances[23]. The VMSafe API is said to allow trusted inspection of network, memory, disk, and processor activity of guests and will enable a variety of monitoring appliances to be created. Such security appliances will have a cost in terms of host load to perform the extra inspection duties, but should enable trusted out-of-band inspection combined with deep host-awareness.

[22] http://www.virtualization.info/2007/07/cisco-to-announce-first-3rd-party.html

[23] http://www.vmware.com/technology/security/vmsafe.html

Separation of Duties and Organizational Roles

Many organizations have staff to provision and maintain network infrastructure who are separate from system administration staff. Such organizations often rely on the separation of duties among these groups to provide policy enforcement and oversight. Organizational adjustments are sometimes required to ensure that oversight is maintained when appropriate, while unnecessary barriers to provisioning resources are eliminated.

Conclusions

Virtualized environments offer compelling benefits and are being rapidly adopted by many organizations, but also aggregate an unprecedented amount of risk into a single data-center component. Furthermore, security researchers have demonstrated that a wide variety of attacks are possible against virtualization technologies. Although there is currently little evidence of malicious attacks, other product spaces have shown that threat landscapes can evolve rapidly. As virtualization deployments increase in number and value, and as security research matures and is weaponized, there is a possibility that attacks against virtualization infrastructure will become commonplace. Organizations can begin preparing now to successfully manage risk in such a threat landscape by considering virtualization infrastructure as a tool to enforce security boundaries, and by articulating what boundaries may be virtually enforced and what boundaries must be enforced by other tools.