VMworld2005
virtualizenow

Security Management in a VMware Virtual Infrastructure Environment

John Y. Arrasjid
VCP, Professional Services
Jason Mills
Tech Support Engineer
VMware, Inc.

SLN 138

Part 2 Agenda: SLN138b

- Overview
- Regulatory compliance
- CPU isolation
- Memory isolation
- Network isolation
- Disk isolation
- Partitioning and management
- Special configurations

Part 1 Agenda: SLN138a

- Overview
- Methodology
- VirtualCenter
- ESX Server
- Virtual machines
- Logging and monitoring
- Security scanning
- Prevention, detection and forensics

1. Overview

Introduction: Who Are We?

- John Y. Arrasjid, BSCS, VCP
  - VMware PSO consulting architect
  - Developer of VMware PSO engagements on security, performance and disaster recovery
  - Has worked with both large and small companies
    - VMware, Roxio, AT&T Information Systems, Amdahl, 3Dfx, Chronologic Simulation, WebNexus Communications
- Jason Mills
  - VMware technical support engineer
  - Member VMTN infrastructure support group
  - Has worked with both large and small companies
    - VMware, Exodus Communications, Apple Computer, Integrated Device Technology

What Will Be Covered?

- Security best practices for virtual infrastructure
  - ESX Server
  - VirtualCenter
  - Virtual Machines
- Logging and monitoring
- Security audits of virtual infrastructure
- Regulatory compliance and resource management
- Note: This material is an extract from the VMware Professional Services Virtual Infrastructure Security (VIS) engagement. For details, please see www.vmware.com/services/consulting.html or your sales contact.

What Will Be Covered?

- The SLN138 session is broken into two parts:
- SLN138a: Is geared to infrastructure related topics
- SLN138b: Is geared to detailed discussion on securing production machines
- Note that there will be a 10 minute recess between sessions.
- Approximate time for each session is 50 minutes, including Q&A.

VMware Security Response Policy

- VMware has a strong corporate response policy:
  - Monitoring of public repositories such as CERT
  - Acknowledgement and initial analysis: Posting of KB article with mitigation or workaround
  - Fix and issuance of a patch if needed
  - Customer Notification: Customers with SNS (Subscription and Support) notified of patch via e-mail
- Code is audited regularly be external resources and resulting recommendations are implemented.
- VMware's security response policy can be found at: www.vmware.com/support/using/security_response.html
- Summary of security notifications for VMware products can be found in KB article #1107.

2. Methodology

1. Document
2. Document
3. Document
4. …

- Infrastructure security is nothing without documentation ☺

Security is Essential

- VMware can provide you with information on how to identify and mitigate potential security concerns relating to your virtual infrastructure.
  - Virtualization does not eliminate need for security controls
  - Greater flexibility and control over resources requires planning and monitoring to meet security goals
  - Regulatory compliance may require understanding how virtual infrastructure components relate to physical infrastructure components
- VMware cannot tell you how to secure your physical plant and remote KVM systems.
  - Systems security does not end at the datacenter cage door

Recommendations

- Understand security controls within your virtual infrastructure!
- Plan for security incident handling
- Keep security patches up-to-date
- Automate detection and analysis
- Monitor security websites and mailing lists including those maintained by VMware
- Ensure that your Security Engineering Team is represented when developing a Virtual Infrastructure Methodology Center of Excellence (VIM CoE)

Virtual Infrastructure Security Plan

- Foundations for a plan tied to Virtual Infrastructure are similar to Physical Infrastructure.
- The following represents the foundation and outline of a plan.

Misconceptions

Security is not only an IT problem but a corporate Intellectual Property (IP) problem. If you don't use security to protect your IP, it will soon become someone else's.

Todd Massey
CEO, Privacy Networks

- "Virtualization eliminates the need for security controls." – FALSE
- Virtualized assets are as important as physical assets when dealing with sensitive content.
- The same security controls used in physical systems should be applied to virtual machines (i.e.: antivirus software, permission controls, etc.).
- "Virtualization is going to make my job more difficult as a security engineer." – FALSE
- Virtualization prevents opportunities for standardization that are not available in the purely physical world. Virtual Infrastructure access controls reduce datacenter shared-cage risks.

3. VirtualCenter

- VirtualCenter security background
- VirtualCenter security setup
- VMotion security background

VirtualCenter Security Background

- VirtualCenter security relies on Windows security controls
- VirtualCenter is role-based and tied to Active Directory or legacy NT domains
- VirtualCenter authentication centralizes access controls and provides opportunities for role-based access controls (RBAC)

VirtualCenter Security Background

- The root password for a managed host is cached only long enough to enable VirtualCenter management functionality.
- The VirtualCenter management channel is encrypted using a pseudo-randomly generated password, which is unique to each ESX Server.

VirtualCenter Security Setup

- Ensure that correct security measures are used when provisioning database resources for the VirtualCenter server.
- The database security must have DB Owner role granted to VirtualCenter Server during installation or upgrade. During normal operation, the permissions may be restricted based on the following key permissions:
  - Invoke/Execute Stored Procedures
  - Select, Update, and Insert
  - Drop

VirtualCenter Security Setup

- On the VirtualCenter Management Server:
- Install virus scanning software
- The VirtualCenter service runs as a user that requires local administrator privileges, and must be installed as a local administrator
- Restrict local security policy
- Do not allow local users access (require use of the VirtualCenter client)

VMotion Security Background

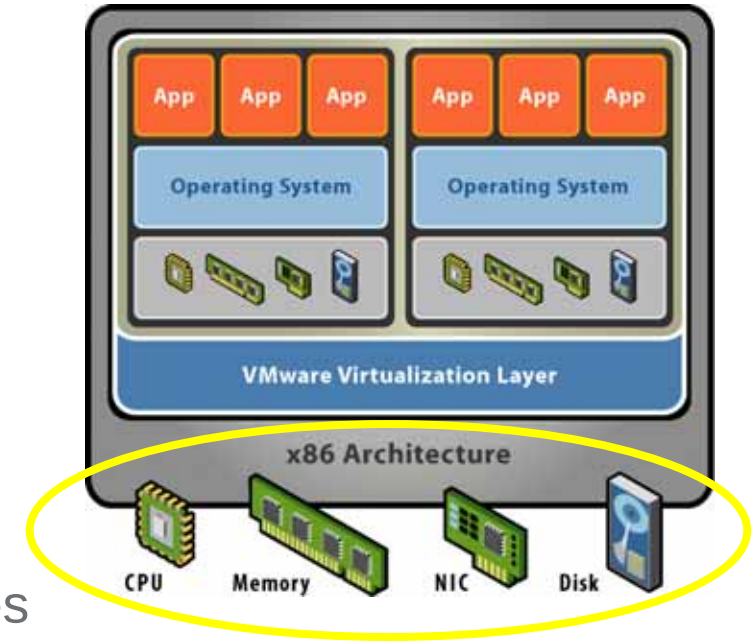
- VMotion transfer of virtual machine active memory content is performed over a non-encrypted path
- For optimal security, VMotion should be implemented on an isolated network. This may mean using 802.1q VLAN technology or a separate physical network as your security requirements dictate

4. ESX Server

- Security setup
- Security monitoring
- Security hardening

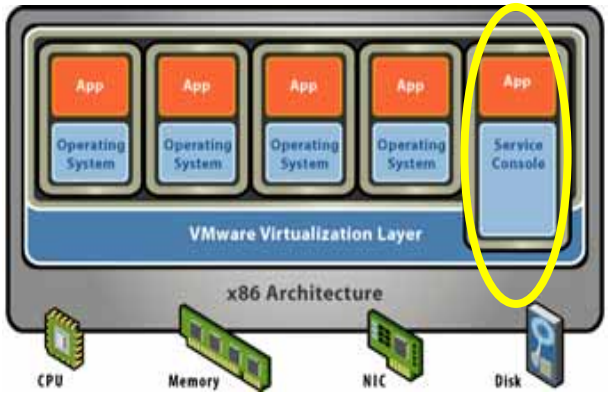
ESX Server Security Setup

- VMkernel
- Runs on and controls physical hardware
- Schedules resources for virtual machines
- Security features
  - No public interfaces
  - Cannot execute arbitrary processes
- Complete isolation of resources
  - Although they share physical resources each virtual machine cannot ‘see’ any other device other than the virtual devices made available to it by the virtual machine monitor.
  - Further discussion in part II of this presentation (SLN138b)



Service Console

- Bootstraps ESX Server and is its administrative interface
- Modified Red Hat Linux
- Security features
  - Apache modified to run only necessary features
  - Unnecessary services disabled
  - Insecure services disabled by default
  - Setuid/setgid programs limited
  - Permissions on files and commands limited



Network

- Leave ESX Server security settings at “High”
  - Limits ports to 22, 80, 443, 902 and 905
  - Forces use of SSH for access to shell and files
  - Forces use of HTTP over SSL (HTTPS) for web browser connections. Note that port 80 is automatically redirected to port 443 (HTTPS)
- Beware of promiscuous mode on network interfaces! Use for security monitoring only
- Label networks for virtual machines appropriately to prevent confusion or security compromises. Also required for compatibility with VirtualCenter and VMotion implementation

Network

- Use of a NIC on a dedicated management network for the Service Console is recommended to restrict access, provide secure connectivity and minimize the risk of DoS or DDoS attacks

Network Isolation

- Use VLANs for isolation if there are too few NICs
- Service console and virtual machines on separate networks
- Virtual machines from separate groups on separate networks
- Be aware that both SSL (HTTPS) and SSH/SCP are vulnerable to man-in-the-middle attacks leveraging a rogue server. Registered server certificates such as that from Verisign minimize this risk. The reason is that it is easy to perform ARP/DNS spoofing at the client side to intercept the traffic and get the logon credentials or the content of packets. (HIGH security settings on ESX Server still require other mechanisms especially in networks with lower levels of security mechanisms.)

File Integrity

- Install a checksum tool in the service console to monitor file and directory integrity
- Minimize additional, layered software in the service console to minimize risk of opening security holes. If adding software, analyze and lock down the installation
- ‘cksum’ and ‘md5sum’ is included with ESX Server for use to verify file integrity

Management User Interface (MUI)

- With SSL enabled, security certificates are created by, and stored on, the ESX Server. These are used for MUI access. These self-signed certificates do not provide end-to-end authentication. We recommend that you consider purchasing a commercial certificate from a trusted Certificate Authority
  - To use your own security certificate for your SSL connections, the certificate must be placed in /etc/vmware-mui/ssl. The management interface certificate includes 2 files:
    - mui.crt – the certificate
    - mui.key – the private key file
  - When upgrading the MUI, the certificate remains in place
  - Note that correct DNS resource records must be present and must match the certificate information to prevent warnings during browser certificate validation

NTP and SNMP

- Configure NTP on each ESX Server platform to ensure accurate clocks for logging purposes
- It is essential to have accurate clocks for event correlation under a forensic investigation. This also prevents problems with certain VMs, like Windows Domain Controllers, that are exposed when clocks get out of sync. (i.e.: Kerberos fails if there is more than 5 minutes difference between client and server.)
- ESX Server supports SNMP v1 only, in which community strings are sent in plaintext. A combination of IP tables and SSH tunneling may be set up to allow use of SNMP while maintaining the security of the SNMP transaction and the SNMP daemon
- The VMware MIB is all read-only. No portion can be set through SNMP management calls

Super-User Access

- Super-user (root) access can be provided in a restricted manner by using the sudo tool that is included with ESX Server. This is used for restricted root access within the Service Console. There is no equivalent in the MUI. The following are setup instructions:
  - *Setup user accounts on the ESX Server*
  - *Log in as root*
  - *Run: visudo*
  - *Add user accounts and permissions following sudo guidelines.*
  - *Save the file.*
- Modify /usr/lib/vmware-mui/apache/htdocs/vmware/src/login.js page if you want to deny specific users from accessing the MUI: (i.e. They only need Remote Console access)
  - *Pseudo-code: if loginname == “joe” deny access*

Syslog Logging

- Use remote as well as local syslog logging.
  - For entries where you wish logs to be sent to a remote log host, add the following within the /etc/syslog.conf file after the log name and restart syslogd:
    - , *@loghost*
  - Example:
    - *mail.* /var/log/maillog, @loghost.company.com*
- Use remote as well as local sudo logging.
  - local2.info /var/log/sudo, @loghost.company.com
- Note: Syslog is NOT encrypted and uses UDP (Best attempt delivery without guarantee). If you require encryption, consider using IPSec. Be aware that this is not currently on the supported list

Syslog Logging

- Note that you will want to update the /etc/logrotate.conf file to ensure log rotation to prevent filling up the log partition. This file defines log file rotation, log file compression, and time to keep the old log files. Processing the contents of /etc/logrotate.d directory is also defined here.
- The /etc/logrotate.d directory contains instructions service by service for log file rotation, log file compression, and time to keep the old files.
- For the three vmk* files, raise “size” to “4096k” and enable compression. This will allow greater logging with minimal impact to file system storage space.

Xinetd Logging

- Enable xinetd logging. This is disabled by default.
  1. Add the following option to the xinetd line in /etc/init.d/xinetd:
     -syslog auth
  2. Modify /etc/syslogd.conf appropriately
  3. Restart xinetd: xinetd restart

Service Console Best Practices

- Do not run the X Server within the service console
  - It may be acceptable in certain instances to use an X Windows client such as xterm with the display sent to an alternate X Windows Server, but the security implications of this should be weighed carefully
- Do not apply third-party patches to the ESX Server Service Console! Only use VMware authorized patches

Passwords

- Use VirtualCenter authentication
- Establish a password policy!
- Account entries for ESX Server are stored in /etc/passwd with the encrypted passwords stored in /etc/shadow
- Under the service console, based on Linux RedHat 7.2, the maximum password length is 128 characters
- There are tools that can be used to attack the passwords on the service console such as John the Ripper and Crack. Both require access to the encrypted password file /etc/shadow. This requires root access on the service console
- Never change access permissions for /etc/shadow
- Pick phrases to help with password selection and never use dictionary words
- Use grub password: prevents users with access to the keyboard from passing options to the kernel during the boot process

Storage and Network

- Mask and zone SAN resources appropriately
  - Minimize access to storage
  - Allow enough access for VMotion
- Create file systems for /home and others
  - Make sure / file system does not fill up. This can affect the operation of the Service Console under ESX Server. Please see VMware Knowledge Base article 904
  - Some file systems such as /home, /var and /vmimages will have different size requirements based on your use and any additional software added to the service console

ESX Server Security Hardening

- Use antivirus software within the Service Console. Do not scan VMFS volumes for performance reasons. Ensure that the virus definitions are kept up-to-date!
- Use antivirus software within your Guest Operating Systems
- Tighten SSH within the Service Console:
  - Modify the SSH configuration to disable ssh version 1 protocol. (Covered in VMworld LAB006)
  - Disable SSH login for the root user. (Note that this is not always possible)
  - Implement DSA/RSA authorized keys to authenticate users (prevents man-in-the-middle attacks)
  - Enable a login banner for law enforcement purposes

5. Virtual Machines

- Security Setup, Monitoring and Tightening

Virtual Machine Security Setup

- Each has its own set of virtual hardware
- Each isolated from the others like physical machines
  - Fault isolation: OS crashes in one do not crash the others
  - Security isolation: Viruses affecting a single virtual machine cannot infect the virtual disk file of another virtual machine
  - Resource isolation: Runaway applications in one do not cause the others to starve

Virtual Machine Security Setup

- Lockdown virtual machine related directories and files
  - Use disk permissions to modify read and write access to them
  - Configuration file permissions modify control of the virtual machine power and device functionality
- Provide network connections to appropriate segments
  - Remember that more than one network connection can result in the virtual machine becoming a router, gateway or bridge
  - Appropriate security should be applied

Virtual Machine Security Setup

- Label networks at the Virtual Infrastructure Node (VIN) level
  - This avoids confusion or security compromises at the virtual machine level
- Disconnect unused devices such as CD/DVD and floppy media

Virtual Machine Security Setup

- Minimize use of the VMware Remote Console
  - Use native remote management services such as terminal services and ssh instead to prevent performance impacts to the service console
  - The remote console provides power management and removable device connectivity controls
- The remote console utilizes TCP port 902 for access
  - Ensure that this port is reachable by any systems connecting to it
  - Optional: Install a stateful firewall, with stateful packet inspection, in between the management system and the ESX Server with the virtual machine to be managed

Virtual Machine Security Setup

- Prevent virtual machine from spoofing virtual MAC addresses
  - *Ethernet<n>.downWhenAddrMisMatch = TRUE*
  - *Ethernet<n>.noForgedSrcAddr = TRUE*
- Disable promiscuous mode for the virtual machine to prevent monitoring of all packets on a virtual switch (unless using a virtual machine for network based IDS, NIDS)
  - *Ethernet<n>.noPromisc = TRUE*
- Disable all ports on a virtual switch from being promiscuous. Does not prevent a guest OS from entering promiscuous mode
  - */proc/vmware/net/devname/config*
- Disable a particular port, based on MAC address, from being promiscuous. Does not prevent a guest OS from entering promiscuous mode
  - */proc/vmware/net/devname/macaddress*

Virtual Machine Security Setup

- Disable virtual machine logging to prevent risk of very large log files generated by malicious use of the virtual machine
  - *Isolation.tools.log.disable = TRUE*
- Disable cut and paste in virtual machines on hosted products
  - *isolation.tools.copy = FALSE*
  - *isolation.tools.paste = FALSE*

6. Logging and Monitoring

Logging and Monitoring

- Proactive, rather than reactive, monitoring of logging is essential to ensuring a secure environment. This includes using filters to eliminate non-important items and identify key issues.
- We are going to cover some of the areas of logging that are built-in to the VMware products as well as methods for monitoring these areas.

You can't defend what you can't monitor…
If you can't monitor for security, you probably can't monitor for performance reasons either.

Richard Bejtlich
Principal Consultant, Foundstone

Logging

- **ESX Server**
  - */var/log/{messages*,secure*,cron*,boot*}*
  - *Sudo logs*
- **GSX Server and Workstation**
  - *Event logs*
  - *C:\Documents and Settings\username\My Documents\My Virtual Machines\Vmname\vmware.log*
- **VirtualCenter**
  - *Event logs*
  - *Within the VirtualCenter agent, view the VirtualCenter "Tasks" tab under the "Reasons" column*
- **P2V**
  - *C:\Documents and Settings\username\Local Settings\Temp*

Monitoring

- Monitoring of log files as mentioned previously.
- Setup of SNMP monitoring.
- Implementation of rule sets to filter logging providing output to the monitoring team of relevance.
- Monitoring also includes Intrusion Detection Systems, or IDS. These can be Host (HIDS) or Network (NIDS) based. Each has it’s relative strengths and weaknesses.

7. Security Scanning

Scanning

- Scanning does not always provide valid results. You must perform further investigation to determine validity of the results to eliminate false-positives and identify non-reported issues.
- The following slides show a list of vulnerabilities that were discovered by a Nessus scan against ESX Server 2.0.1 and then explains what the results mean.

Nessus Scan Against ESX Server 2.0.1

- Note: This scan is against an older version of ESX Server. Newer versions do not show all of these as false positives.

1. **Vulnerability ssh**

   CAN-2003-0682, CAN-2003-0693, CAN-2003-0695

   VMware KB Article 1371

   VMware has patched this bug as of ESX Server 2.0.1 and higher.

Nessus Scan Against ESX Server 2.0.1

2. **Warning ssh**

   CAN-2003-0190

   Risk low: A remote attack could potentially use a brute force attack to determine usernames with this exploit. It would not give the user access to the system, but could facilitate a password guessing attack.

   Mitigation: This can be mitigated by changing the PAM configuration as follows:

   In /etc/pam.d/system-auth change line 2 to:

   auth sufficient /lib/security/pam_unix.so likeauth nullok nodelay

Nessus Scan Against ESX Server 2.0.1

3. **Warning ssh**

   CAB-2003-0386

   Risk low: A remote attack could bypass access controls. The default configuration of ESX Server does not utilize access restrictions based on source address. The attacker would need to control the DNS server of the network the attack was launched from to bypass these access controls. The attacker would still have to have a valid user id and password to launch a successful attack.

   Mitigation: Utilize access controls at the Firewall to prevent this attack from being attempted. Ensure user ids and passwords are limited to those with a valid need, and that good password policies are in place.

Nessus Scan Against ESX Server 2.0.1

4. **Vulnerability https**

   CAN-2004-0700

   VMware KB article 1429

   VMware does not use mod_proxy, thus exploit is not possible. Leaving mod_proxy disabled is recommended.

Nessus Scan Against ESX Server 2.0.1

5. Vulnerability https

   CAN-2003-0542

   VMware KB article 1259

   CERT VU#867593 Cross-site scripting attack vulnerability using HTTP TRACE

   "To prevent cross-site tracing attacks, you may wish to prevent the server from interpreting HTTP TRACE method. For example, you can place the following lines in your mod_rewrite module to deny all HTTP TRACE and TRACK requests:

   RewriteEngine On

   RewriteCond %{REQUEST_METHOD} ^(TRACE|TRACK)

   RewriteRule .* - [F]

Continued on next page…

Nessus Scan Against ESX Server 2.0.1

When deciding whether or not to deny all HTTP TRACE requests, keep in mind the following:

- To exploit the TRACE method, a malicious user must be able to access the ESX Server or GSX Server Web-based management interface
- The malicious user must have a Web browser that ignores the same-origin policy
- There is no more sensitive data in the Management Interface exposed to a cross-site tracing attack than would be exposed to a standard cross-site scripting attack that accesses the JavaScript document.cookie property

Nessus Scan Against ESX Server 2.0.1

- CAN-2003-0542: Multiple stack-based buffer overflows in (1) mod_alias and (2) mod_rewrite for Apache before 1.3.29 allow attackers to create configuration files to cause a denial of service (crash) or execute arbitrary code via a regular expression with more than 9 captures.
- "In order to expose this vulnerability, the attacker would need to access to the apache configuration file directory as root to modify or replace Apache configuration files that are owned by root. Non-root users cannot alter these files."

8. Prevention, Detection, and Forensics

Prevention

- Develop security policies and procedures
- Follow security best practices for each of your Virtual Infrastructure components
- Perform security scans of both your virtual machines and your virtual infrastructure platforms
- Identify network connections and their use by your virtual machines to eliminate tunneling security risks

Detection

- Implement release procedures for new software, patches and changes to the virtual infrastructure. This ensures you can track if an ‘unscheduled’ change has occurred that can lead to the source
- Implement a virtual honeypot to watch what is occurring in your network
- Setup an Intrusion Detection System (IDS) and be proactive in updates and review of logging. Fine tune if there are too many false positives. (Note: We see this all the time when a VMware customer has a scan done on their infrastructure yielding many false positives.)

Forensics
What can VMware do for you?

- Computer Forensics involves the following for gathering evidence and for root cause analysis
  - Detection: Use tools to monitor and detect intrusion.
  - Preservation: Save data for both evidence and for analysis.
  - Identification: What system and components were compromised.
  - Extraction: Pull the key items related to the intrusion for analysis.
  - Documentation: Identify methods, source, target, treasure. Document this and historical information on evidence.
  - Interpretation: Take the information gathered from the analysis phase and make conclusions that can be used to go to the next step. That may include criminal prosecution, further analysis or honeypot monitoring.

Forensics

1. Guaranteed preservation of evidence for forensic analysis.
   - Using VMware P2V the state of the original machine can be reproduced virtually for analysis by multiple investigators without contaminating the original evidence. (Be aware that certain components are changed in the virtualization reconfiguration.)
   - Clone of the configuration files and virtual disks of a virtual machine is 100% identical.
   - Clones can be analyzed by server, network, storage, and dedicated security groups at the same time with the original components 100% intact. No risk of damage to evidence.

Forensics

2. Use of virtual machine honeypots
   - A honeypot is an Internet host computer that appears to be simple to compromise but collects information on the criminals activity
   - Virtual machines make very simple deployments for gathering information and for luring criminals away from your critical systems
   - Search on Google for ‘VMware’ and ‘honeypots’!

Forensics

3. Deployment of latest security toolkits quicker
   - Try Knoppix-STD: Knoppix Security Tools Distribution CD. This contains open source tools for managing, preserving, monitoring, and analyzing your virtual infrastructure.
   - Use VirtualCenter to monitor and interact with your environment. Example: If a security compromise is identified, use (VINs) Virtual Infrastructure Networks and VirtualCenter Alerts to lock down connections between your network segments.

Questions?

Intermission

10 Minute Break

Part 2 Agenda: SLN138b

9. Overview
10. Regulatory compliance
11. CPU isolation
12. Memory isolation
13. Network isolation
14. Disk isolation
15. Partitioning and management
16. Special configurations

9. Overview

- -**- for each bullet point, discuss the gain and penalty
- All references to proc nodes are based on a root location of /proc/vmware/ unless otherwise noted

10. Regulatory Compliance

Regulations

- HIPAA
- GLB
- SAS-70
- Cobit
- ISO 17799
- COSO
- U.S. PATRIOT Act
- SOX
- FDA Chapter 11

Regulations

- Section 508
- CSRC
- FIPS
- RBAC
- NIST
- OASIS (eHealth/HSL)
- HL7

Safeguard Overlap

- Many of the regulatory compliance guidelines have overlapping areas:
- Administrative safeguards
- Physical safeguards
- Technical safeguards

Vendor Neutrality

- Regulatory guidelines are vendor neutral but the nature of the guidelines may influence vendor decisions.
- VMware simplifies the process as the underlying vendor hardware is abstracted away from what the guest OS sees.
- There are four hardware components to be aware of: CPU, RAM, network and disk

Control Weaknesses

1. Improper account provisioning with segregation of duties

   Use VirtualCenter to lock down access to Virtual Infrastructure Nodes (ESX Server and GSX Server) and virtual machines.

Control Weaknesses

2. Insufficient controls for change management

   Integrate VirtualCenter with existing change management processes.

   - Identify levels of change:
   - Notification to change management for VMotion
   - Signoff for change affecting network and disk access

Control Weaknesses

3. Lack of understanding around key system configurations

   VMware Certified Professional (VCP) certification to ensure understanding of configuration settings and implications of incorrect settings.

Control Weaknesses

4. Audit logs not being reviewed and tracking of audit reviews

   Ensure that your VIM CoE team reviews logs on a regular basis. These logs include VirtualCenter, MUI, and service console logs.

Control Weaknesses

5. Abnormal transactions not identified in a timely manner and/or violation of security policies within the network

   Identify changes in VirtualCenter access controls. These access controls are set using a combination of VirtualCenter and Domain controls (Active Directory/NT Domains).

Securing the Inner Network

- Define security relationships between virtual machines, virtual infrastructure components and staffing.
- Segregate the network and storage into security pods (or zones).
- Enforce established security relationships within and across security pods.
- Perform regular network audits to ensure security relationships are enforced.
- Update security relationships as business needs and regulatory compliance dictate.
- Provide audit trails and reporting to satisfy regulatory compliance audits.

11. CPU Isolation

- Ability to select various hyperthreading (HT) options on a per-chassis and per-virtual machine basis to address specialty workloads, HT Cache Snoop concerns, etc.: on a cost/need basis
- Ability to assign explicit CPU affinity on a cost/need basis (discussed in Section 15)
- Support for dual-core AMD Opteron in ESX Server 2.5.2 (released Sept. 15, 2005)

11. CPU Isolation

- Per-Chassis Level
  - HT Disabled
    - Physical CPU package remains in "Single-Task" mode at all times
    - Physical CPU package resources (L1/L2/L3 cache) are available for exclusive use by the scheduled vCPU
    - Higher ROI for low vCPU:pCPU ratios and cache-aggressive workloads
    - Coarse CpuSched granularity

11. CPU Isolation

- Per-Chassis Level
  - HT Enabled
    - Physical CPU package becomes a logical partner pair of CPUs (increased CpuSched granularity)
    - Physical CPU package switches between "Multi-Task" mode (default state) and "Single-Task" mode as directed by the VMkernel
    - Physical CPU package resources (L1/L2/L3 cache) are shared between logical partners
    - Base shares per Physical CPU package still 10k

11. CPU Isolation

- Auto HT quarantine by VMkernel based on known "HT Hostile" behavior patterns
  - Defaults to no HT quarantine
  - Applies short-term HT quarantine as needed
  - config/Cpu/MachineClearThreshold config/Misc/VmkperfPerWorld
  - vmkperf/worlds_machine_clear_any
  - Enable (1) or disable (0) counter:
    - echo -n '1' > vmkperf/enable/...

11. CPU Isolation

- Per-virtual machine level
  - No HTq
    - Any vCPU from any running world may occupy other half of logical partner pair
    - Physical CPU package remains in "Multi-Task" mode unless min CPU% commitments must be met by CpuSched
    - vCPU gets non-exclusive access to shared physical CPU package resources (L1/L2/L3 cache)

11. CPU Isolation

- Per-virtual machine level
  - Inter-virtual machine HTq (HT Sharing: "internal")
    - Only other vCPUs from current world (virtual machine instance) may occupy other half of logical partner pair
    - Physical CPU package remains in "Multi-Task" mode -IFF- the virtual machine is multi-vCPU

11. CPU Isolation

- Per-virtual machine level
  - Intra-virtual machine HTq (HT Sharing: "none")
    - Nothing may occupy other half of logical partner pair
    - Physical CPU package is dropped from "Multi-Task" mode to "Single-Task" mode
    - vCPU gets exclusive access to normally shared physical CPU package resources (L1/L2/L3 cache)
    - vCPU is "charged" at double the normal rate

11. CPU Isolation

- Per-virtual machine level
  - Functionality of "Isolate Virtual Machine from Hyper-Threading" UI checkbox is vCPU-count dependent
  - VMkernel will override virtual machine configuration if it violates CPU affinity constraints

11. CPU Isolation

- Per-virtual machine level
  - May be confirmed in .vmx configuration file:
    - (unset, or ...) sched.cpu.htsharing = "any"
    - sched.cpu.htsharing = "internal"
    - sched.cpu.htsharing = "none"
  - May be confirmed in proc node for relevant currently running World ID:
    - vm/WID#/cpu/hyperthreading
      htSharing: current mode
      maxSharing: most restrictive possible mode

12. Memory Isolation

- Each virtual machine operates in dedicated memory address space
- Transparent Page Sharing is VMkernel controlled: virtual machine cannot compromise it
- TPS may be disabled per-chassis
- TPS may be disabled per-virtual machine
- TPS scan rates may be tuned

12. Memory Isolation

- Ability to adjust virtual machines memory priority with respect to other active virtual machines
- Ability to allocate 100% of memory for individual virtual machines in physical RAM
- Ability to enable/disable/adjust behavior of Memory Balloon Driver per-virtual machine
- Ability to achieve near 2:1 memory over-commit if needed

13. Network Isolation

- Virtual switch defaults to no promisc
- Ability to create “local loop" virtual switch with zero pNICs
- Each virtual switch is isolated (no code-path between switches)
- Virtual switch will not bridge traffic between two pNICs

13. Network Isolation

- When using 802.1q at VMkernel layer (VST), virtual machines cannot inject untagged or improperly tagged packets
- Ability to apply basic packet shaping
- Ability to lock-down "risky" virtual machines based on additional MAC behaviors

13. Network Isolation

- Untagged packets delivered to separate layer of virtual switch w/ VST enabled
- UI deters attachment of vNIC to "untagged" portion of virtual switch w/ VST enabled
- Ability to "recable" a vNIC to another virtual switch without loss of Link state, useful for realtime HoneyWall quarantine
- -**- Need to confirm last point for sure -**-

14. Disk Isolation

- -??- Content still needed

15. Partitioning and Management

- Managing the "CHON Factor"
  - Enacting operational policies
  - Pre-emptive documentation
    - Runbook (NOC and chassis)
- Regulatory compliance
  - Ability to create resource over commit scenarios for compliance

15. Partitioning and Management

- Playing well with others
  - Proportional-share scheduling
  - Borrowing unused CPU cycles
  - Ability to define min/max CPU% per virtual machine
  - Ability to shift potential or known-hostile workloads to dedicated CPUs and/or VMkernel cell (Intel NUMA / AMD SUMO) using CPU and/or cell affinity

16. Specific Configurations

- Adjusting UI behavior
- Producing pre-customized .vmx files
- Producing pre-customized NVRAM files
- Understanding and armoring high-risk virtual machines
- Handling incident response scenarios
- Repurposing off-hours CPU time
- HoneyWall on ESX Server

Questions?

VMware Professional Services

- The following are consulting offerings from VMware PSO:
- Workshops
  - ESX Server Jumpstart
  - VirtualCenter Jumpstart
- Advanced Workshops
  - Virtual Infrastructure Security (VIS)
  - Disaster Recovery and Backups (DRB)
  - Server Consolidation
  - Virtual Infrastructure Performance Management (VIPM)
- Virtual Infrastructure Methodology Engagements
  - VIM Assess
  - VIM Plan
  - VIM Build
  - VIM Manage

Third-Party Security Patch Monitoring: Secunia

- The following web site provides security advisories and related patches for many vendors: http://secunia.com
- Included in their reports are VMware product reports using vendor ID 300.
- This is not a VMware sanctioned site but is used by many customers as a quick reference.
- The following are product Ids to use related to VMware:
  - 3979 – Virtual Center
  - 2125 – ESX Server 2.x
  - 3315 – GSX Server 3.x
  - 1445 – VMware Workstation 4.x

Related Security Sessions and Labs

- LAB006 – Securing and Monitoring Virtual Infrastructure
- LAB007 – Creating, Securing and Deploying ACE Desktops
- PAC103 – Best Practices for Securing VMware ESX Server
- PAC 117, 600, 601, 698 – VMware ACE presentations
- SLN068 – Important Tips for Securing the Virtual Datacenter
- SLN140 – NSA and NetTop
- SLN240 – How Virtualization Improves Security
- SLN241 – Virtualization Streamlines Regulatory Compliance
- SLN459 – Marine Corps Network Operations and Security Command Virtualization Network and Alternate Site Implementation
- SLN695 – VMware ACE Customer Best Practices for Secured Environments

Risk Equation of Threat Model

- The following is based on a formula from The TAO of Network Security Monitoring by Richard Bejtlich (Addison Wesley, ISBN 0-321-24677-2). It defines a formula for calculating both the total risk and the residual risk. This applies to physical and virtual infrastructure models.
  - Total Risk = Asset Value * Threat * Vulnerability
  - Residual Risk = Total Risk * Controls Gap
- Asset Value = cost of restoration + cost of lost revenue + cost of stolen IP
- Threat = attacker that can exploit an asset's vulnerability
- Vulnerability = weakness in an asset
- Controls Gap = items beyond your control
- Companies implement countermeasures to reduce overall risk to an acceptable level.
- There are four basic ways to deal with risk. These are transferring, rejecting, reducing, or accepting the risk.
- The risk equation above can also be applied to disaster recovery as well as security.

VMworld 2005

virtualize now

las vegas • october 18-20, 2005