SUMMIT JBoss WORLD

PRESENTED BY RED HAT

LEARN. NETWORK. EXPERIENCE OPEN SOURCE.

redhat.

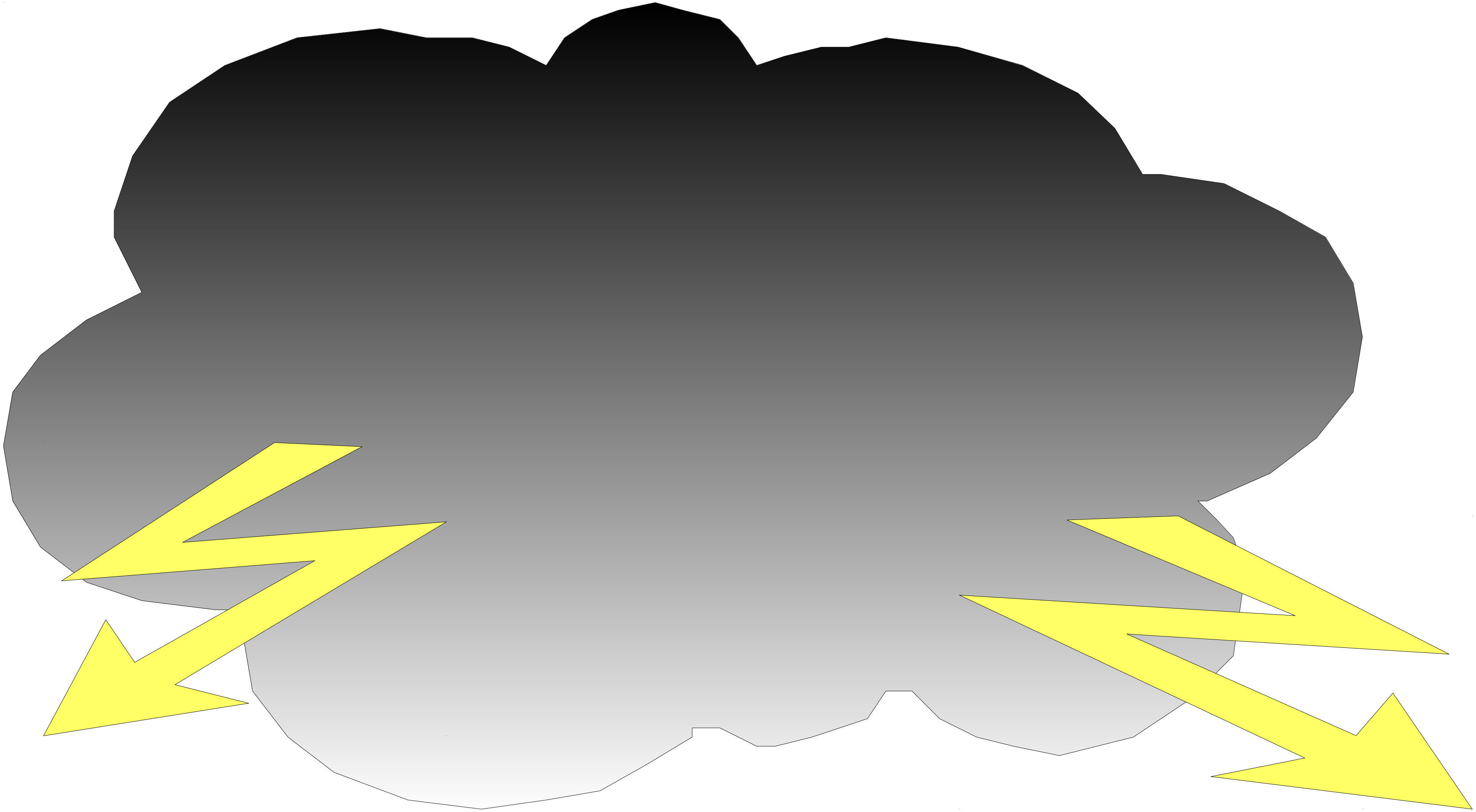
Multi-tenancy Virtualization

Challenges & Solutions

Daniel J Walsh
Mr SELinux, Red Hat

Date 06.28.12

What is Cloud?



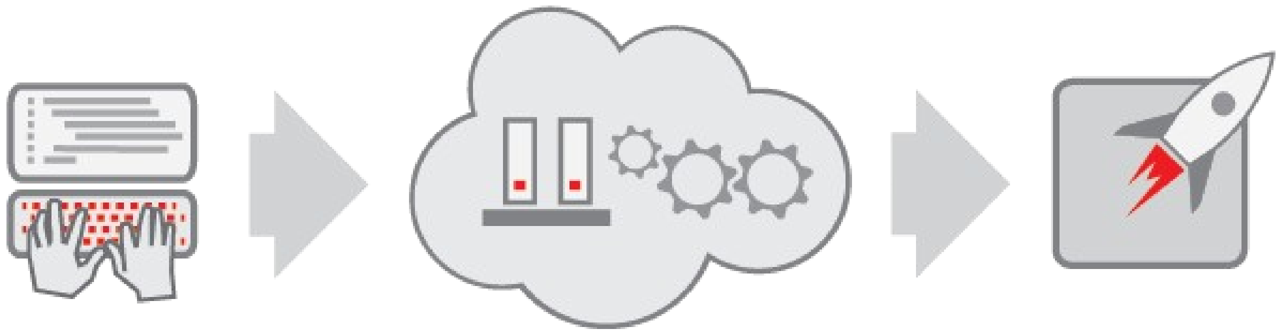
What is IaaS?

IaaS = Infrastructure-as-a-Service

What is PaaS?

PaaS = Platform-as-a-Service

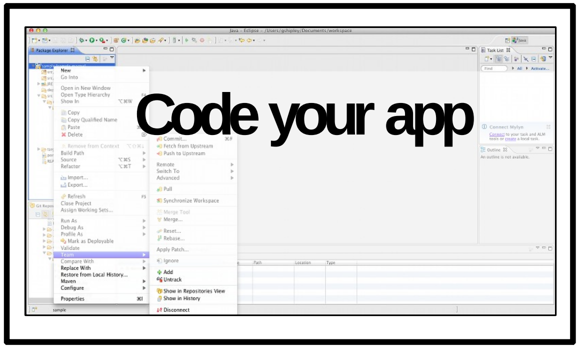
(AKA, a Cloud Application Platform)



Code

Deploy

Enjoy



Code your app

Push-button Deploy, and your App is running in the Cloud!

Save Time and Money

OPENSHIFT

OpenShift is PaaS by Red Hat

What should you look for when choosing where to live?

6 RED HAT ENTERPRISE LINUX

RED HAT ENTERPRISE VIRTUALIZATION 3

TRY RHEV3 NOW

OPENSHIFT

libvirt

the virtualization API

SELinux

cgroups

Quality???

Alicia Nijdam Flickr :Attribution 2.0 Generic (CC BY 2.0)

Quality!!!

6 RED HAT ENTERPRISE LINUX

Red Hat Enterprise Linux certified on more platforms than any other OS — from desktop to mainframe

IBM
hp invent
DELL
CISCO
Bull
egenera
FUJITSU
HITACHI
EMC²
intel
FUJITSU COMPUTERS SIEMENS
Sun microsystems
UNISYS
sgi
Stratus Technologies

Broad ISV Choice

SAP
VERITAS
bea
PeopleSoft
Borland
SYBASE
IBM
ORACLE
bmcsoftware
documentum
EMC² where information lives

Database Choice

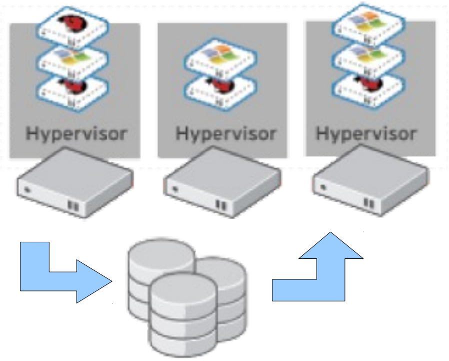
SAP
ORACLE
Ready for IBM DB2 software for Linux
SYBASE

Red Hat Enterprise Linux is Rock Solid

SCALABLE

- Systems to 108 cores, 2 TB RAM, 16 I/O slots
- Designed to scale to 4,096 cores and up to 64 TB RAM
- Industry benchmarks show near-linear scaling to 64+ cores

Flexible



- Resource management: cGroups
- Integrated hypervisor
- Migrate VMs regardless of hardware

Reliable

Self healing, automatic isolation of CPU/RAM

Improved hardware awareness of multi-core and NUMA

Energy efficient power management features

Maintenance ???

Maintenance !!!

nayukim Flickr :Attribution 2.0 Generic (CC BY 2.0)

Red Hat Enterprise Linux Updates are Great!!!

DON'T RIP out/replace Foundation but repair/Improve it.

- Released once or twice a year
- Bug fixes and hardware enablement
- New features in minor releases exception
- Extended Update Support (EUS) program.
- Security/Bugfixes for high-priority issues released asynchronously and don't wait for minor releases.
- Why risk your data with Knock-Offs

External Security ???

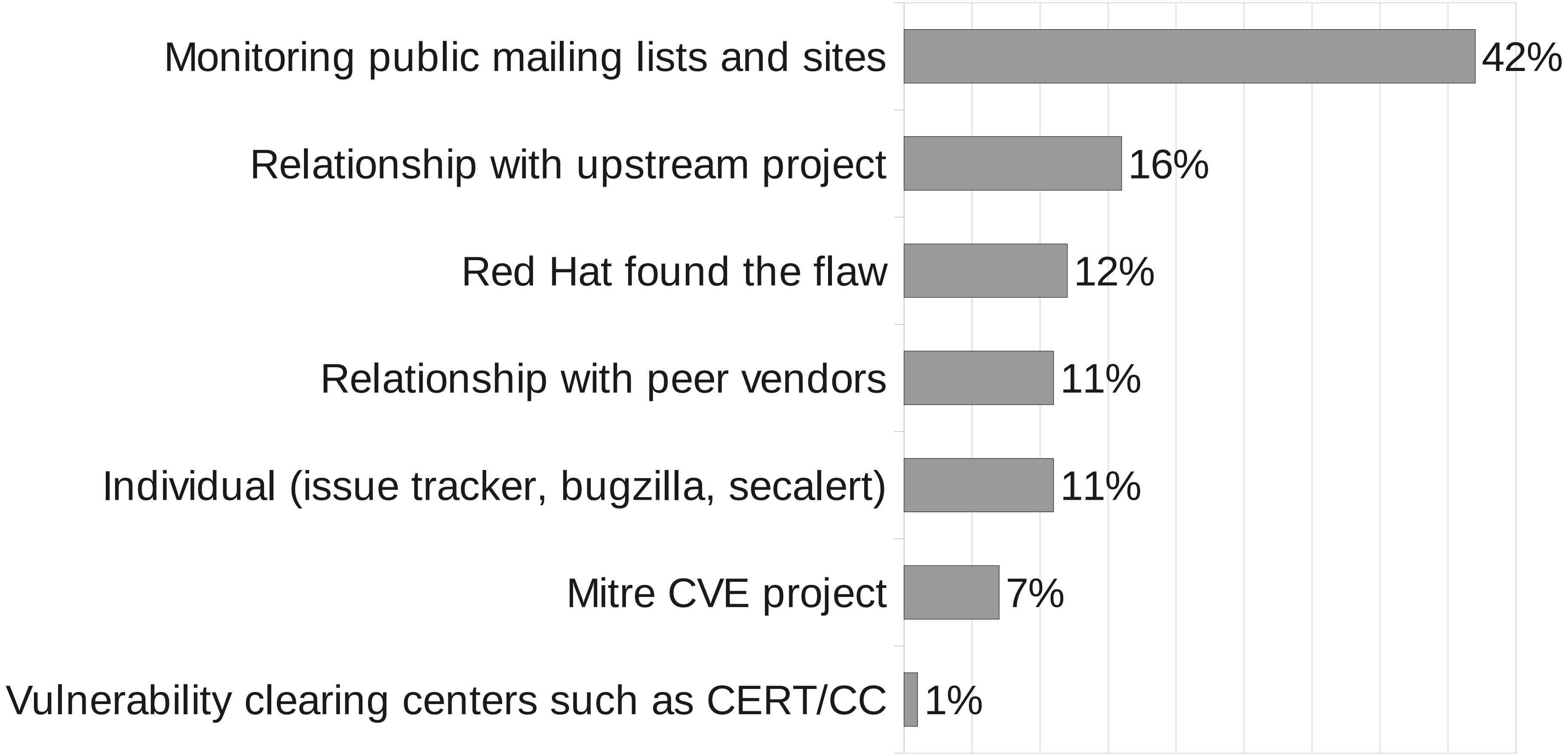
ogimogi Flickr :Attribution 2.0 Generic (CC BY 2.0)

External Security !!!

Red Hat Security Response Team

- Goal
  - Quickly address security issues that arise in products
- Established over 11 years, members span 10 countries
- Monitor vulnerabilities/threats from public/private sources
  - Triage vulnerability severity and determine fix strategy
  - Produce communications to customers
  - Manage process to get the right fix out at the right time
- 99.7% response within one business day of receipt

How we find out about the vulnerabilities



36% of the vulnerabilities reported to us in advance of public disclosure

data: 12 months to March 1 2012, 733 vulnerabilities

Internal Security
Controlling Tenants

nayukim Flickr :Attribution 2.0 Generic (CC BY 2.0)

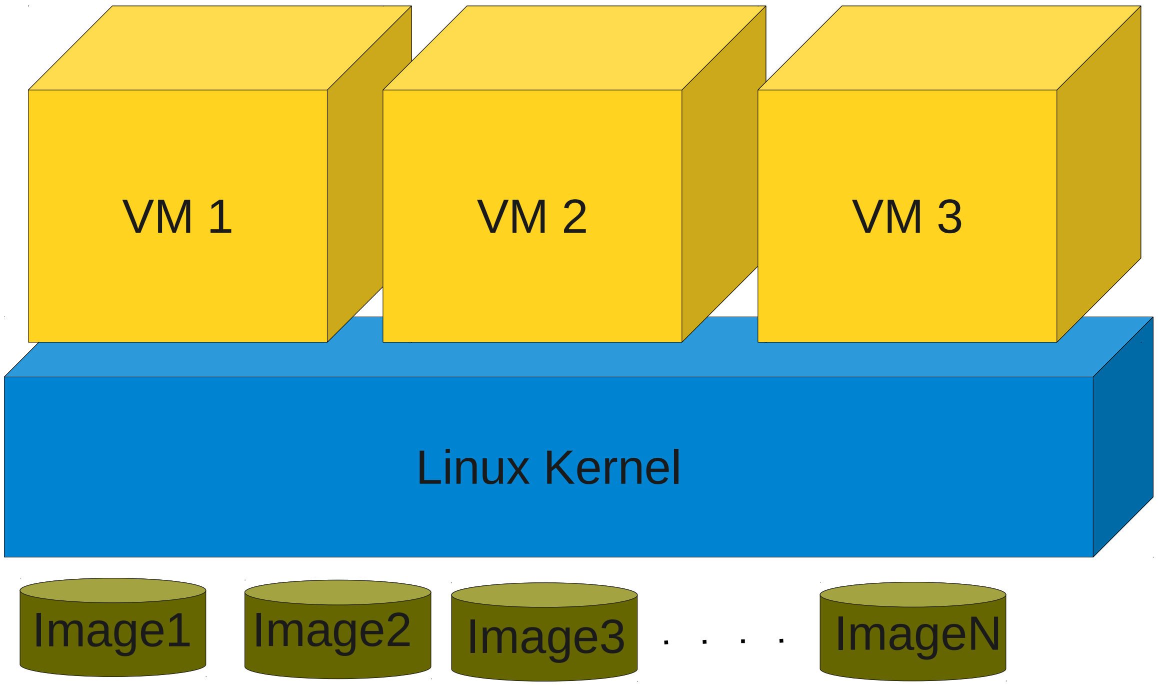
Internal Security Same Tools?

ktow Flickr :Attribution 2.0 Generic (CC BY 2.0)

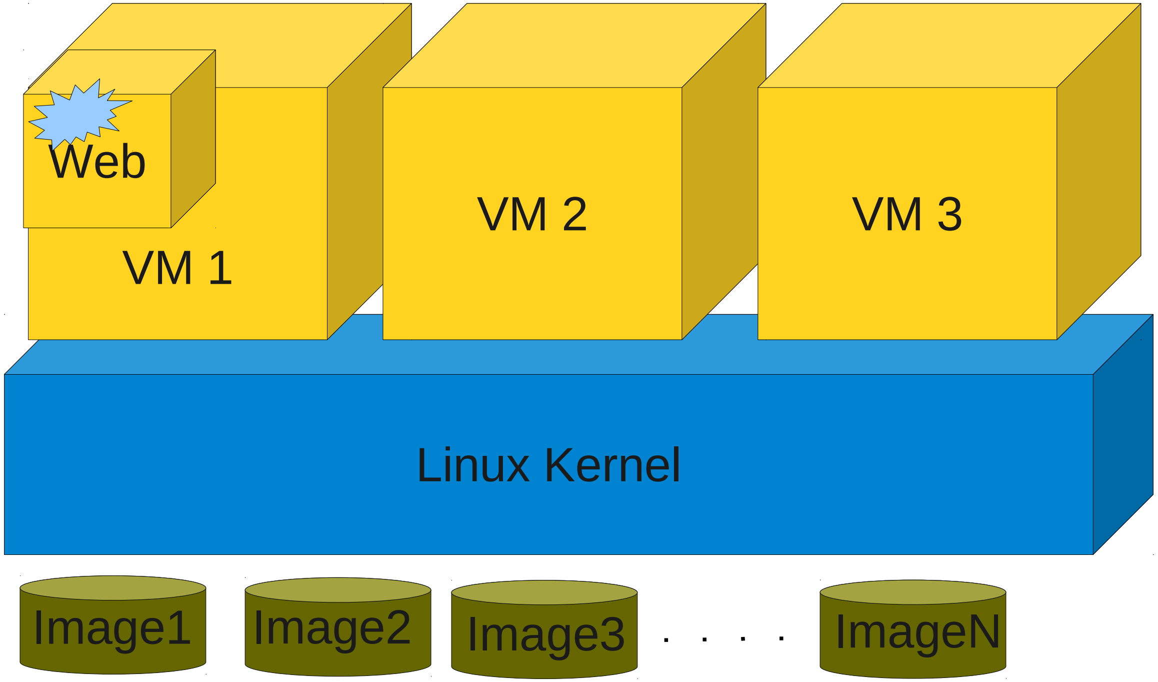
Hypervisor Vulnerabilities

- Hypervisor == All code used to run tenants
  - Not theoretical
  - Potentially Huge Payoffs
  - Xen Already Compromosed
    - Even Red Hat Entreprise Linux 5
  - Google “vmware vulnerabilies” - 500,000 Hits
  - Big topic at Black Hat conference

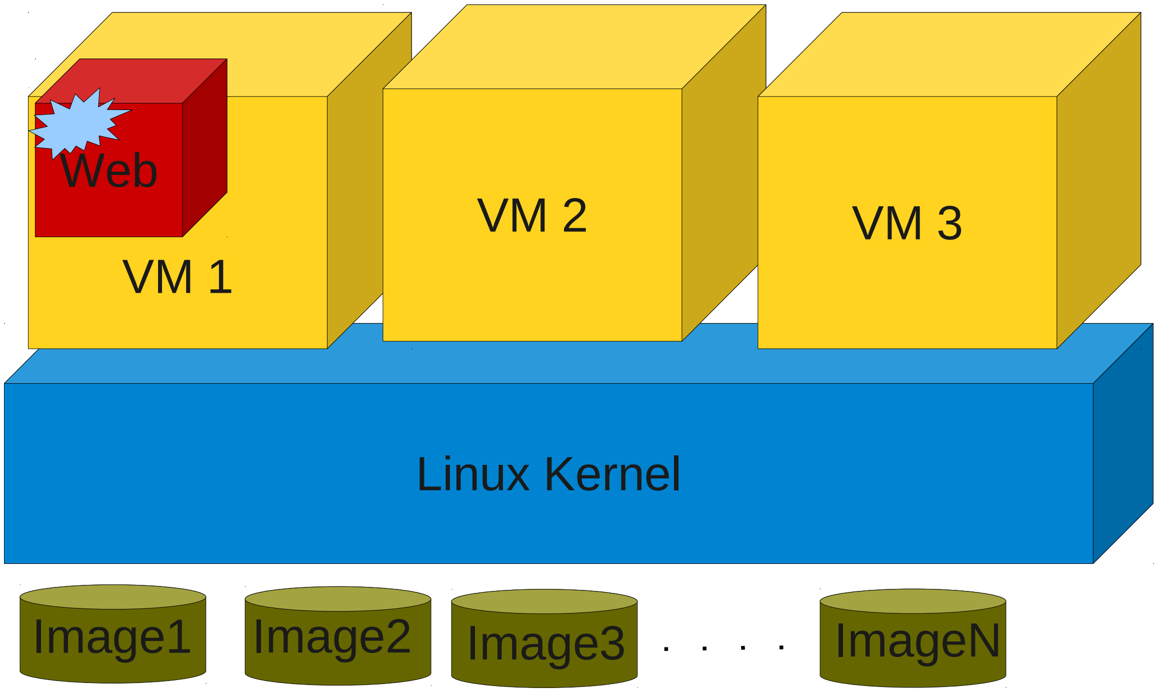
Virtual machine processes all have equal access to the system...



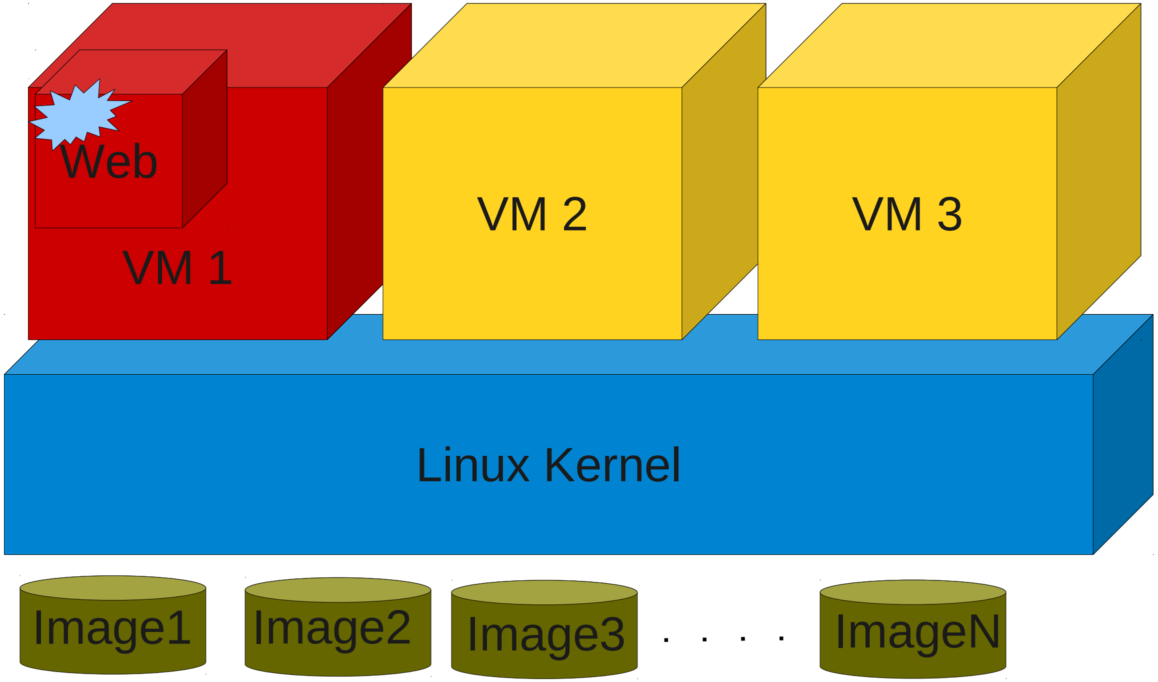
...if application on virtual machine is attacked...



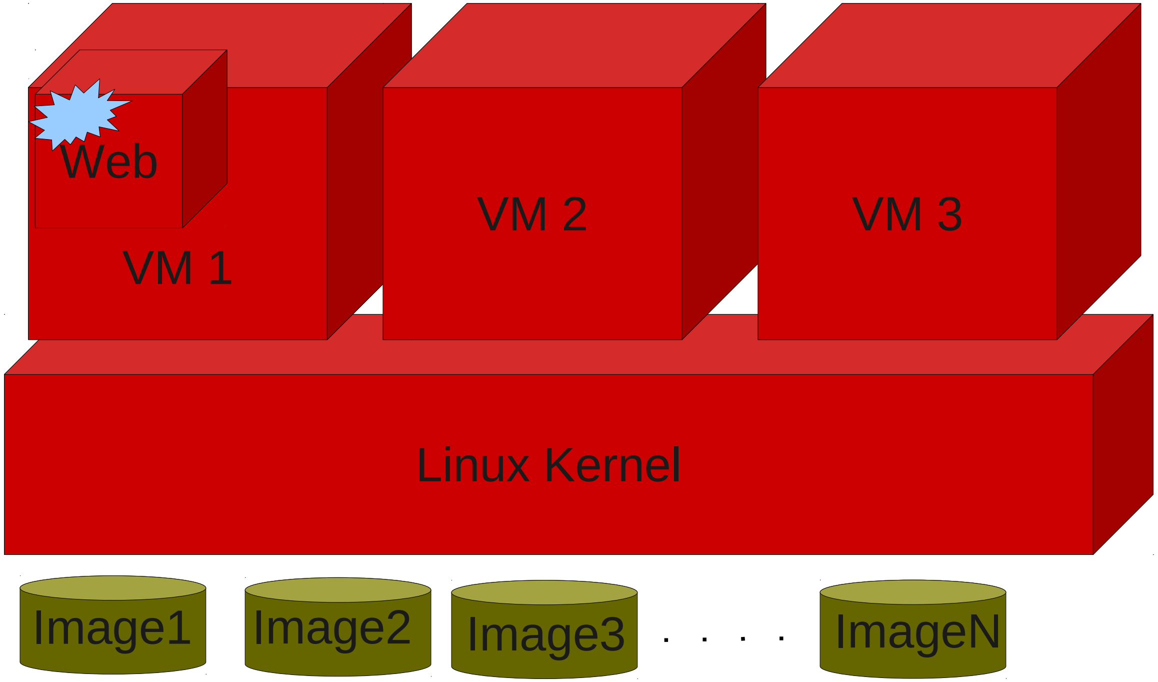
...compromised...



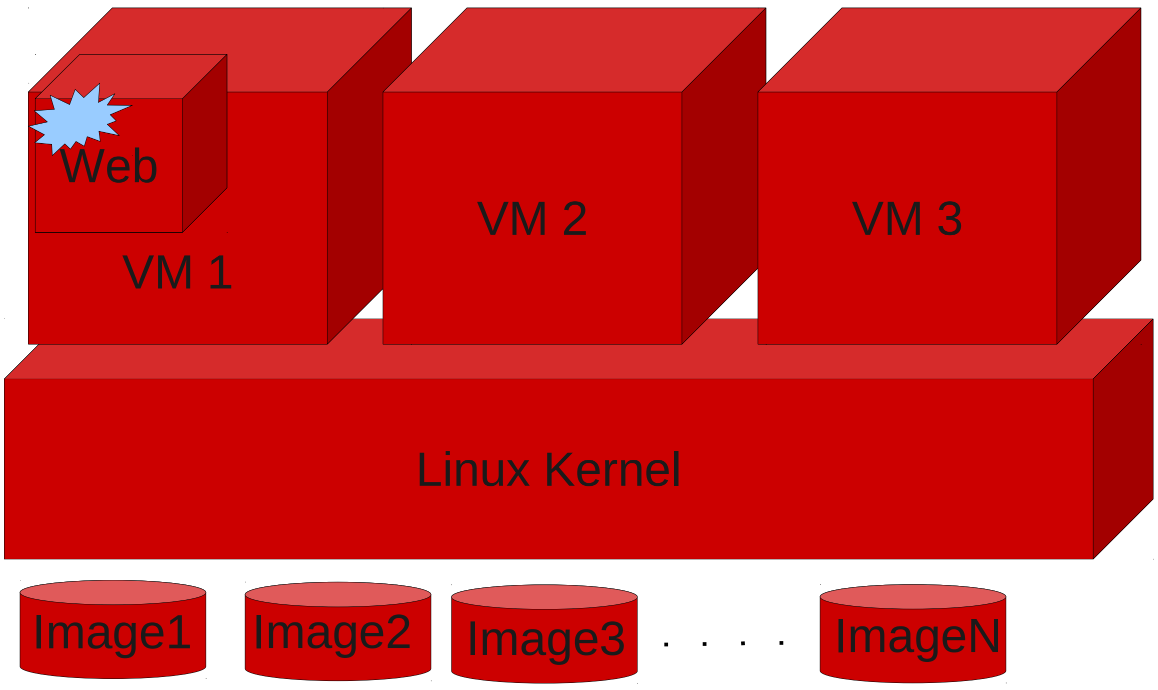
...and gets a privilege escalation...



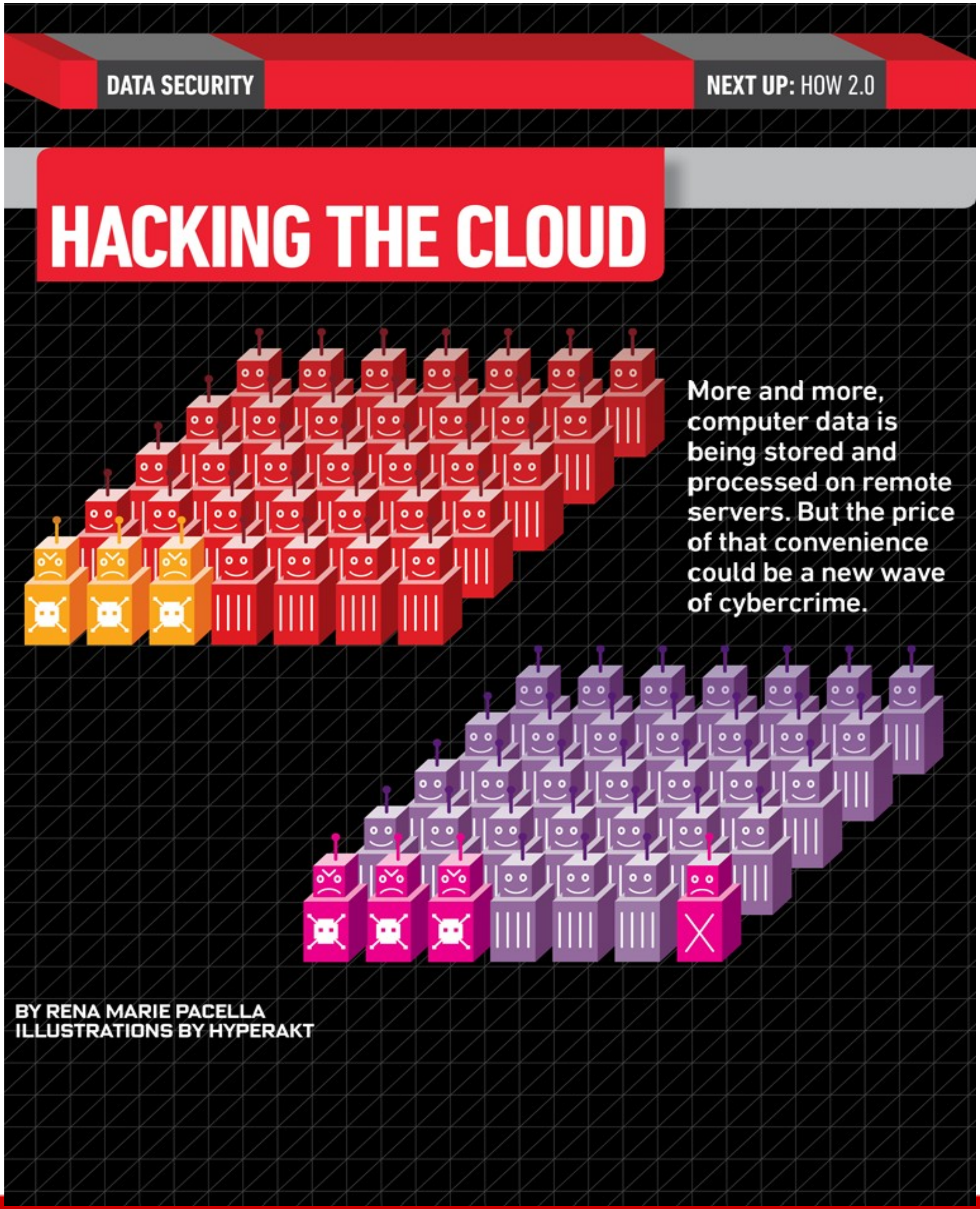
.. and your machine has a Hypervisor Vulnerability ...



.. But not just the running VM's and host, but all images ...



Popular Science Magazine April 2011



DATA SECURITY

NEXT UP: HOW 2.0

HACKING THE CLOUD

More and more, computer data is being stored and processed on remote servers. But the price of that convenience could be a new wave of cybercrime.

BY RENA MARIE PACELLA
ILLUSTRATIONS BY HYPERAKT

SELinux to the rescue

SELinux is all about labeling

SELinux – All Processes get labels

KVM VM's are processes!!!

SELINUX – All Files/Devices Get Labels

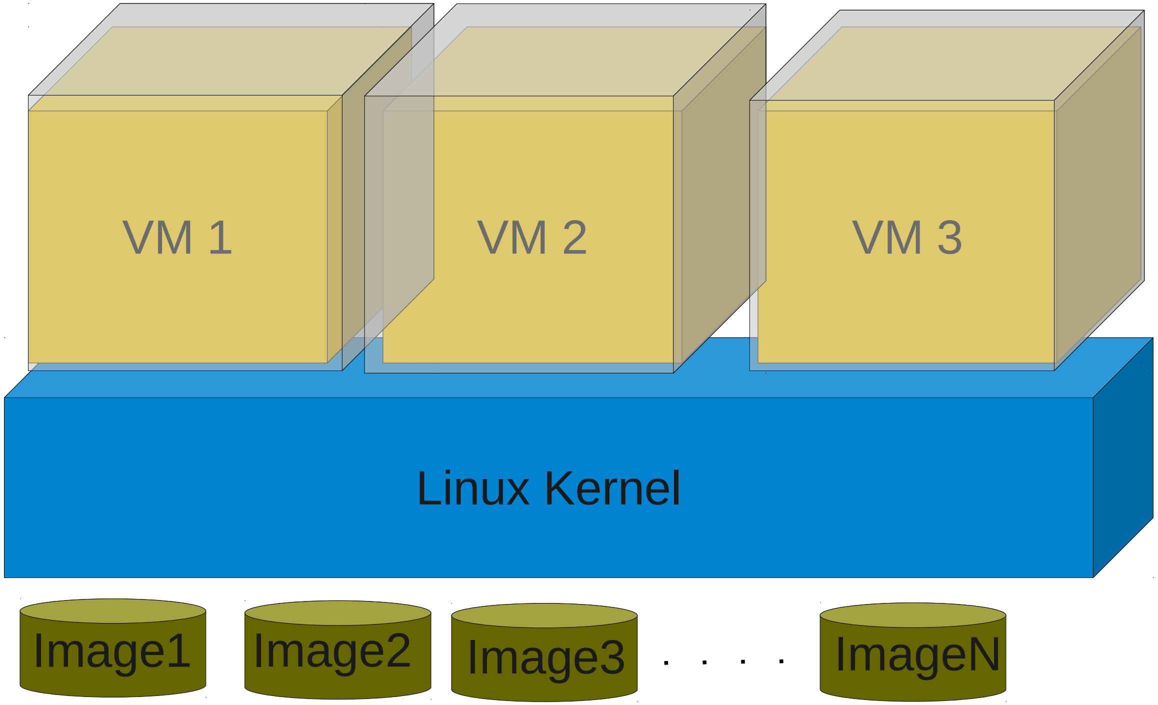
KVM Virtual images are stored on files/devices!!!!

SELinux Policy:

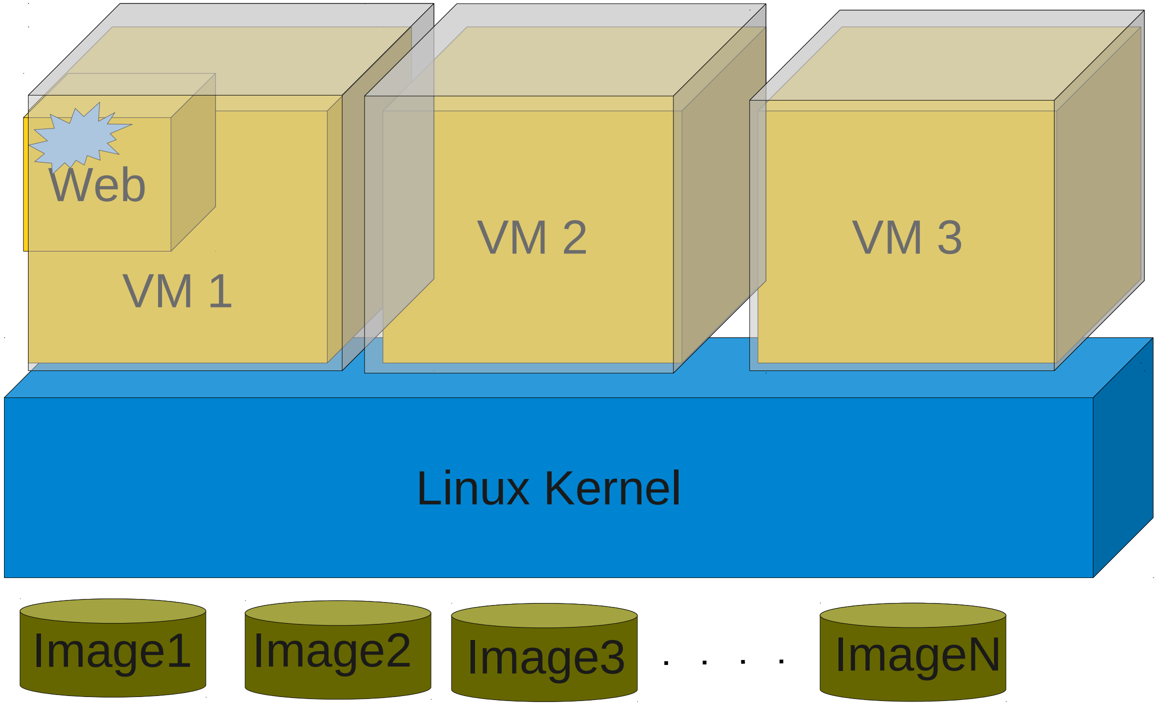
- Governs Process Labels access to Process/File Labels.

Kernel Enforces these Rules.

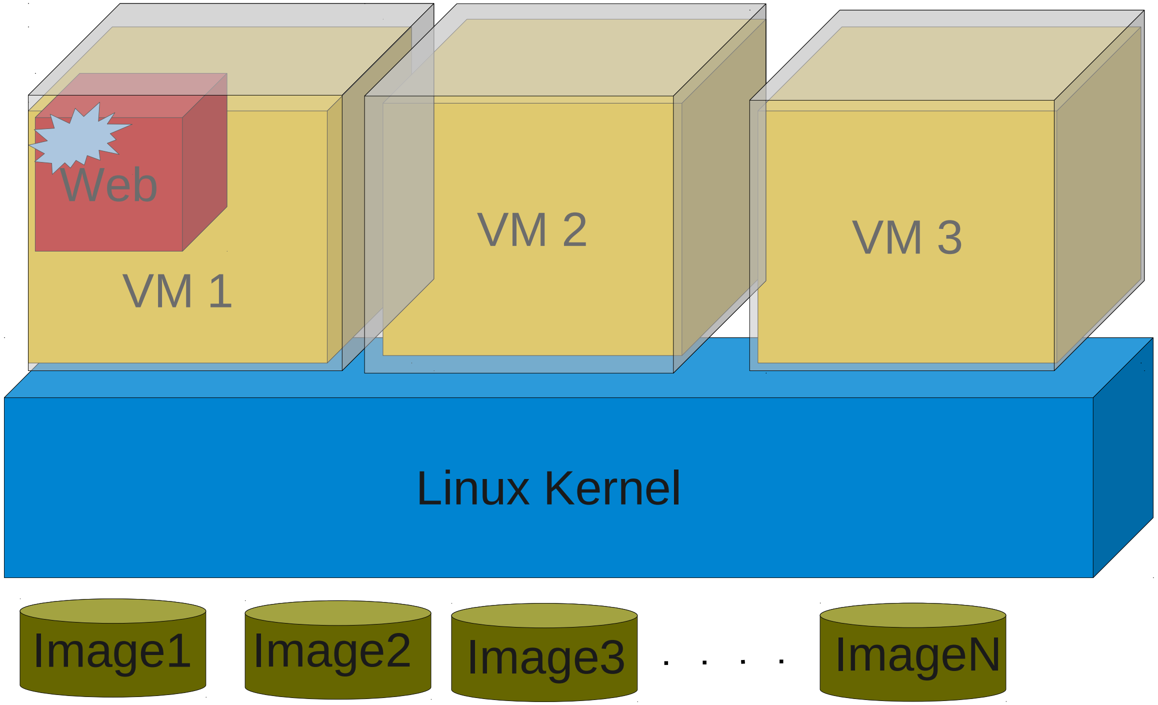
Virtual machine processes all have equal access to the system...



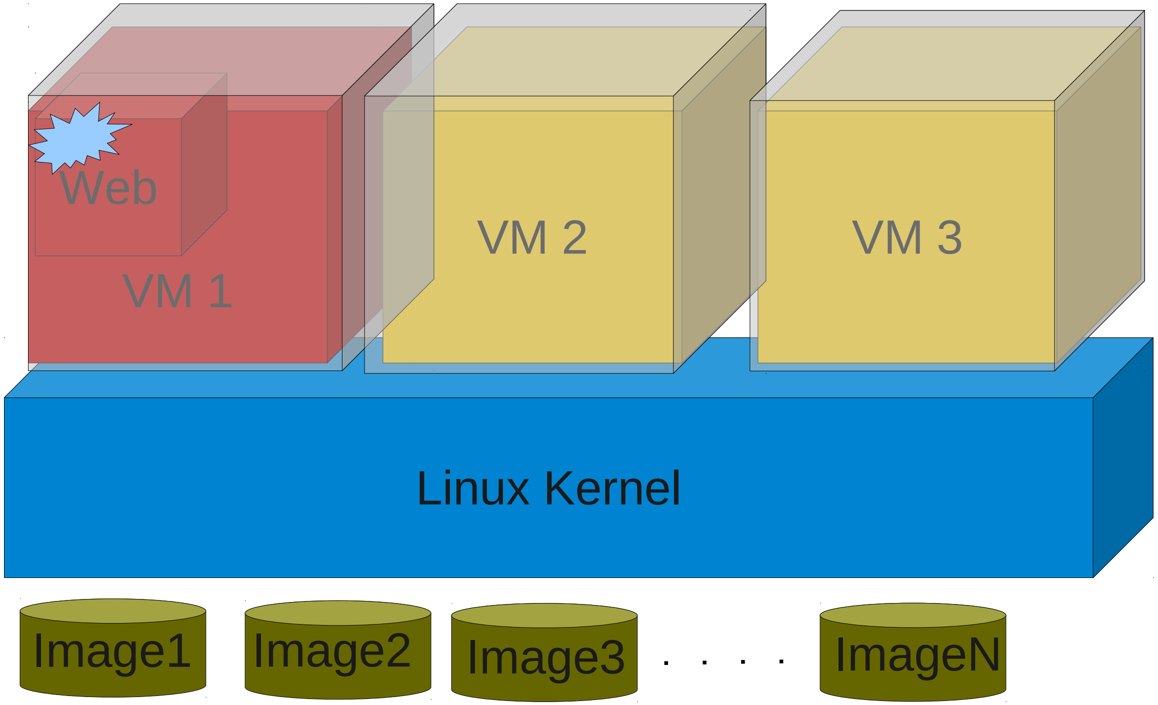
...if application on virtual machine is attacked...

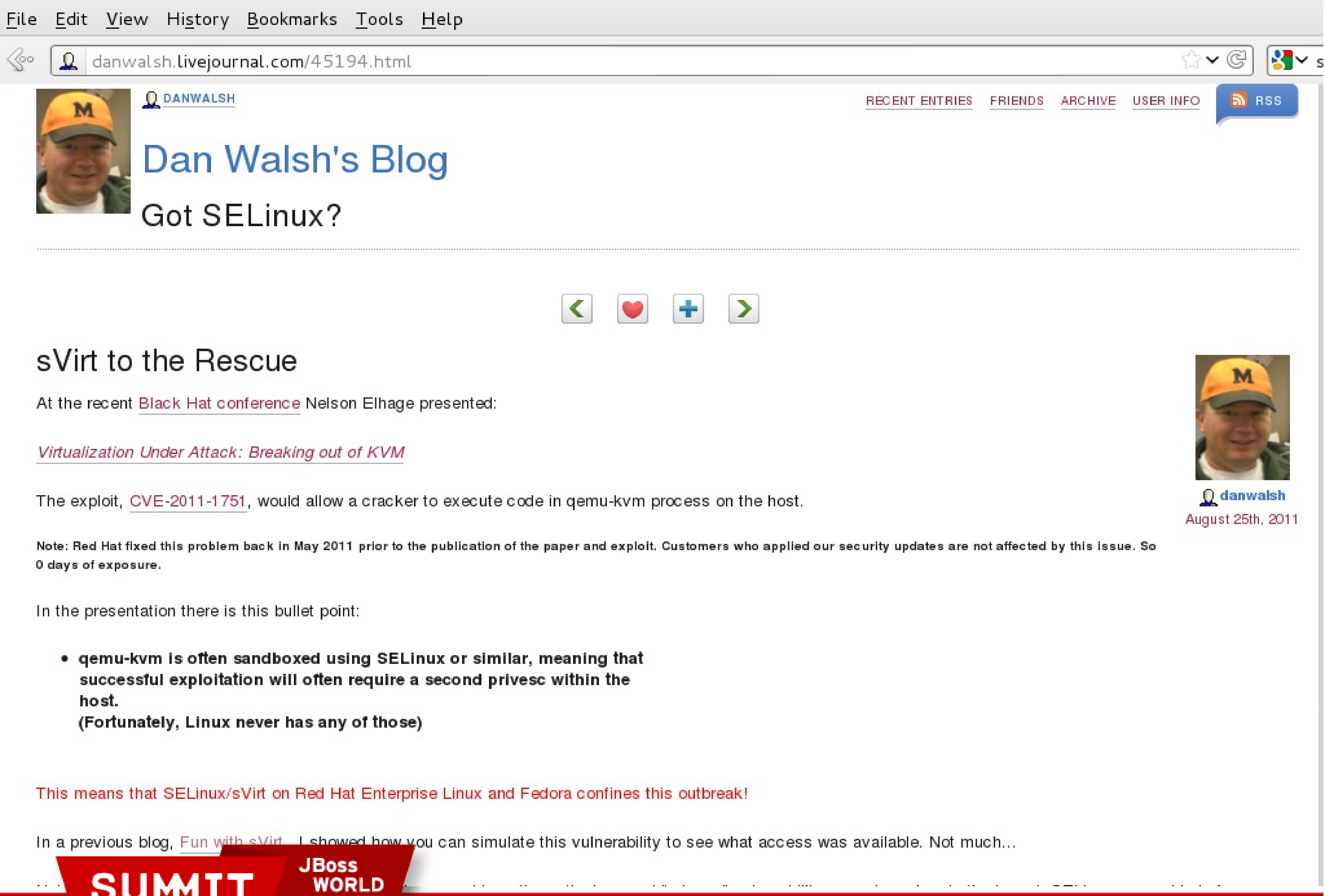


...compromised...



SELinux Force Fields...





File Edit View History Bookmarks Tools Help
danwalsh.livejournal.com/45194.html
DANWALSH
RECENT ENTRIES
FRIENDS
ARCHIVE
USER INFO
RSS

Dan Walsh's Blog

Got SELinux?

sVirt to the Rescue

danwalsh
August 25th, 2011

At the recent Black Hat conference Nelson Elhage presented:

Virtualization Under Attack: Breaking out of KVM

The exploit, CVE-2011-1751, would allow a cracker to execute code in qemu-kvm process on the host.

Note: Red Hat fixed this problem back in May 2011 prior to the publication of the paper and exploit. Customers who applied our security updates are not affected by this issue. So 0 days of exposure.

In the presentation there is this bullet point:

- **qemu-kvm is often sandboxed using SELinux or similar, meaning that successful exploitation will often require a second privesc within the host.**
  (Fortunately, Linux never has any of those)

This means that SELinux/sVirt on Red Hat Enterprise Linux and Fedora confines this outbreak!

In a previous blog, Fun with sVirt , I showed how you can simulate this vulnerability to see what access was available. Not much...

Svirt Demo

http://people.fedoraproject.org/~dwalsh/SELinux/Presentations/svirt.ogv

Shared Resources !

stevendepolo Flickr :Attribution-NoDerivs 2.0 Generic (CC BY-ND 2.0)

Shared Resources !

stevendepolo Flickr :Attribution 2.0 Generic (CC BY 2.0)

Control Group Overview

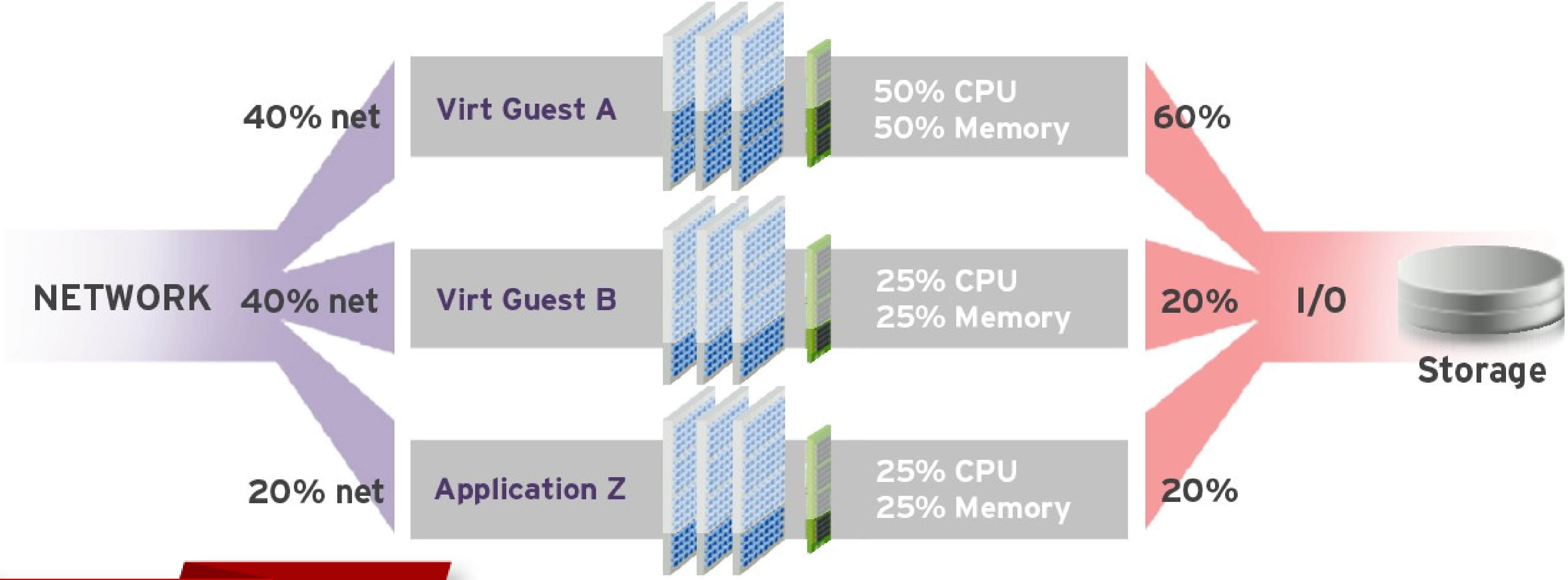
- Control Group is a generic framework where several "resource type of controllers" can be plugged into and managed different resources of the system such as process scheduling, memory allocation, network traffic, or IO bandwidth.
- Two types of control mechanisms
  - Proportional and Maximum Bandwidth Control
- Controller Types Supported
  - CPU/CPUset, Memory, Networking, Block IO, etc.

	RHEL6.2	RHEL6.3+	RHEL 7+
CPU	Proportional & Maximal	Proportional & Maximal	Proportional & Maximal
Memory	Maximal only	Maximal only	Maximal only
Networking	Proportional & Maximal	Proportional & Maximal	Proportional & Maximal
Block IO	Proportional & Maximal	Proportional & Maximal	Maximal [Proportional bandwidth will not work by default]

Resource Management: Control Groups

Ability to manage large system resources effectively

- Control groups (cgroups) for CPU/Memory/Network/Disk
- Benefit: guarantee Quality of Service & dynamic resource allocation
- Ideal for managing any multi-application environment
  - From backup jobs to the Cloud



Cgroups Demo

- http://people.fedoraproject.org/~dwalsh/SELinux/Presentations/cgroups.ogv

Internal Security Futures

SECCOMP/Libseccomp

- Selectively disable syscalls with seccomp
  - ~312 syscalls/x86_64, not including x86
  - Most applications use subset of all the syscalls
  - Reduces chance of kernel exploitation if app is exploited
- Some syscalls are “riskier” than others
  - Not fully protected by LSM/SELinux
  - History of vulnerabilities due to syscall complexity
- libseccomp makes seccomp easy to use
  - Simple architecture independent API for developers

Secure Linux Application Containers

- Run hundreds of servers simultaneously
  - Similar to Openshift
- Little overhead
- SELinux protections built in
- Uses all Namespaces

Verifying the Boot Sequence

- UEFI Secure Boot
- Trusted Boot
  - TXT
  - TPM