

# The Missing VMware Manual for Linux

*Krishna Raj Raja*

*Staff Engineer, VMware*



vmware®

# SYNOPSIS

- **Evolution of VMware and Linux**
- VMware initiatives for Linux
- Running Linux guest OS effectively
- Linux guest in the Cloud

# VMware and Linux

---

## ■ In 1998

- VMware pioneered x86 virtualization
- VMware's initial focus was only on desktop and developer market
- Adoption of Linux further enhanced the virtualization use cases

Date: Friday, April 16, 1999 6:53 AM

Subject: Helluva product

I just had to write and tell you what a hell of a product this is. My company started as an NT-based affair, but now everything we do is UNIX-based, and most of it is being ported to Linux... This thing could really change our engineering group's whole philosophy. Keep up the great work...

Steve

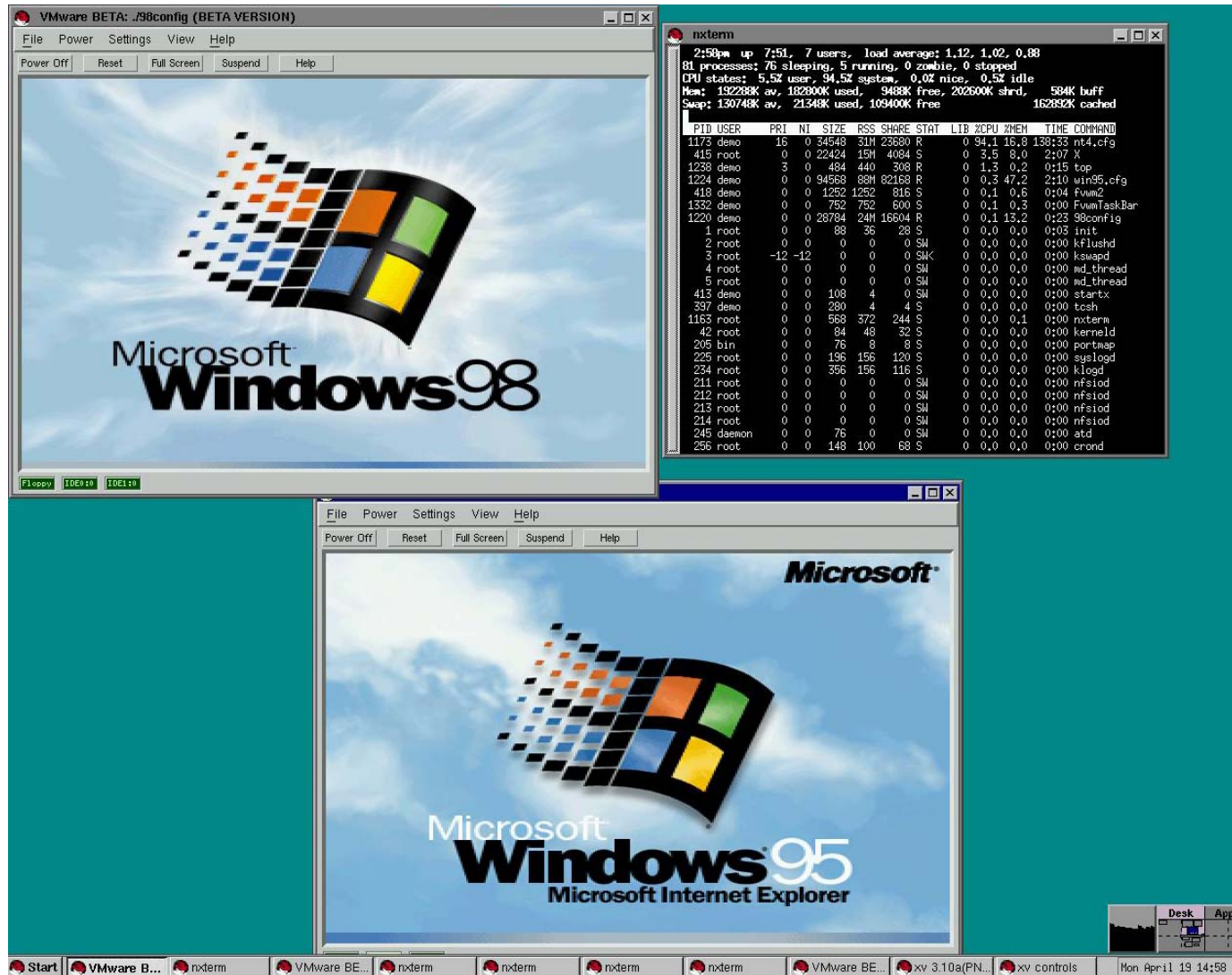
Date: Tuesday, August 17, 1999 5:32 AM

Subject: linux

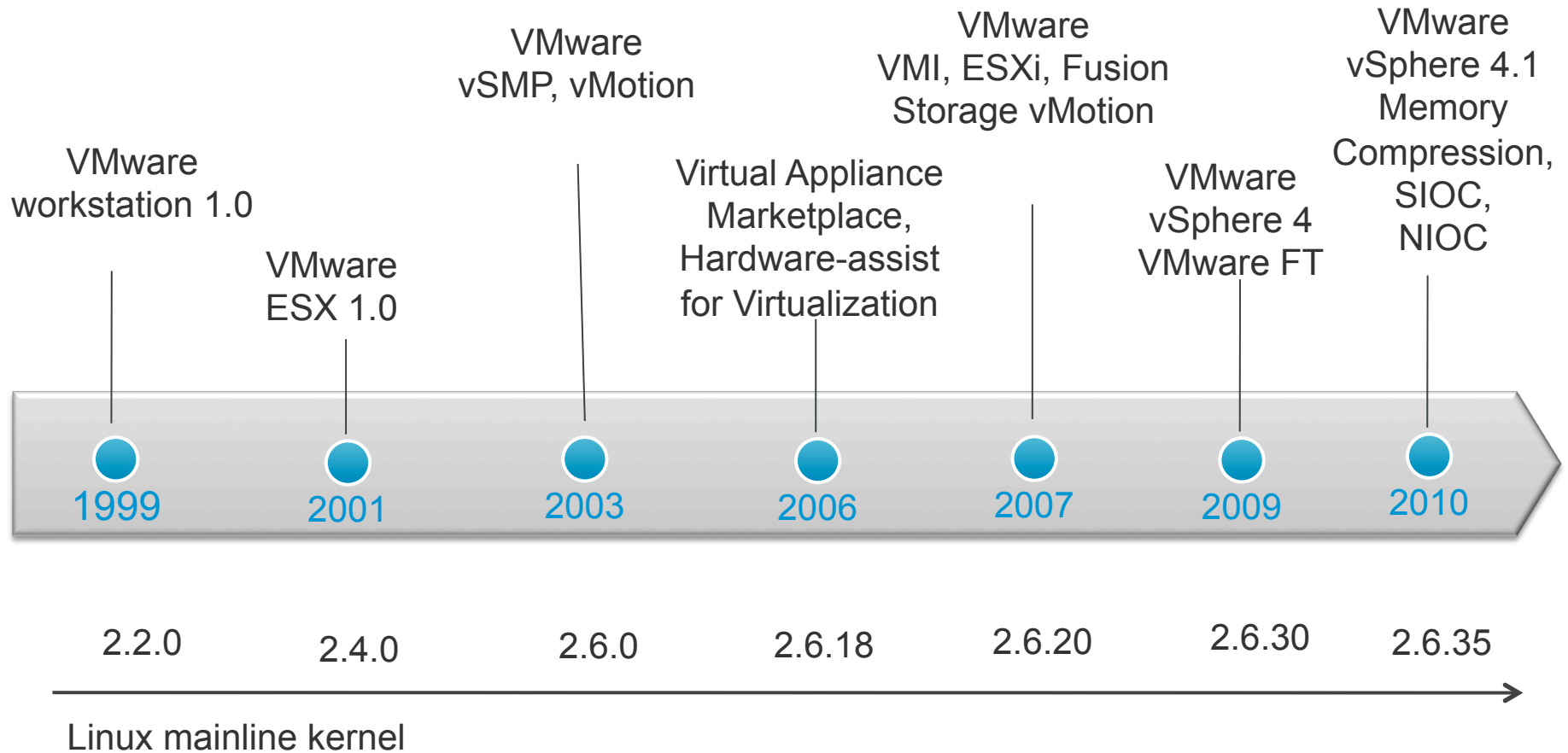
Wat Pah Nanachat is a Buddhist forest monastery for foreign Monks in Thailand..Your program would be very useful to appease the tension building between the Linux users (that's me) and the windows users (that's rest of the community)...Do you give any discount for religious non-profit institution? The monastery does not have an Internet connection so...should we send you a payment by mail?

Shaun Thitabho Bhikku

# VMware Workstation 1.0 for Linux on Redhat Linux 5.2



# VMware and Linux



PC Era

Cloud Computing Era

## Year 1999

- **Linux kernel 2.2.0**
  - 1.8 Million lines of Code
  - 11MB tar.gz download, 62MB uncompressed size
  
- **VMware workstation 1.0**
  - 3MB download
  - Supported 3 Linux distributions
  
- **The Matrix**
  - “Welcome to the virtual world”

## Year 2010

- **Linux kernel 2.6.34**
  - 13 Million Lines of Code
  - 65 MB tar.gz download, 448 MB uncompressed size
  
- **VMware workstation 7.1**
  - 314 MB download bundle
  - 514 MB disk space after install
  - Official support for over 60 Linux distributions
  
- **Inception**
  - “Dreams within Dreams”

# Why the VMware workstation has become so big?

---

## ■ New features

- Multiple Snapshots, Drag and Drop, Shared Folders, Debug register support, PXE boot
- vSMP support, 3D hardware acceleration (DirectX, OpenGL)
- Multi-monitor support, Record/Replay, Replay debugging, vProbes, Eclipse plugin support, Unity
- Virtual printing, virtual disk management, automatic tools update, virtual disk encryption



10 MB

## ■ Other bundled utilities

- VMware player, ovftool, vdiskmanager, Eclipse integration, help files, debug binaries



150 MB

## ■ OS support

- pre-compiled binary drivers and shared libraries



300 MB

# SYNOPSIS

- **Evolution of VMware and Linux**
- **VMware initiatives for Linux**
- Running Linux guest OS effectively
- Linux guest in the Cloud



## Out-of-the box experience

---

### ■ **VMware SVGA device**

- First device driver to be open-sourced, included with XFree86 version 4.1.0 (December 2000)
- Provided out of the box X-Windows graphics support

### ■ **NIC morphing**

- AMD PCnet32 Ethernet device is the default – has the widest compatibility with driver support for even really old operating systems
- PCne32 device automatically morphs into a high performance vmxnet device when the vmxnet driver loads

## Out-of-the box experience

---

### ■ **BusLogic, LSILogic and pvscsi virtual SCSI adapters**

- BusLogic SCSI device is supported by a lot of legacy operating systems
- LSILogic virtual device support was added later to workaround issues with legacy Buslogic device driver
- pvscsi adapter is para-virtualized and provides the best performance for I/O intensive workload

### ■ **Gallium3D driver**

- Unified interface for multiple graphics Compute APIs such as OpenGL, OpenCL, Direct3D etc
- Work in progress

# Open Virtual Machine Tools

---

- **VMware tools package for Linux, FreeBSD and Solaris was open-sourced at VMworld 2007**

- <http://open-vm-tools.sourceforge.net/>

- **The following components are open-sourced**

- Drivers for devices and file system access

- Memory ballooning

- Shared Folders

- Drag 'n'Drop Text and File Copy/Paste

- Clipboard sharing

- Disk wiping and shrinking

- Time synchronization

- Automatic guest resolution resizing

- GuestInfo (stats about guest environment)

- Guest SDK (provides stats about the VM)

- Soft power operations

- Multiple monitor support

- GTK Toolbox UI

- **Open-source drivers**

- vmxnet3 driver is merged into Linux kernel 2.6.32

- pvscsi driver is merged into Linux kernel 2.6.33

- Balloon driver is merged into Linux kernel 2.6.34

# Other VMware Open Source Projects

---

- **Virtual USB analyzer**
  - <http://vusb-analyzer.sourceforge.net/>
- **Xorg Video and Mouse Drivers**
  - <http://intr.overt.org/vmware-xorg-7.0/>
- **Libview**
  - <http://view.sourceforge.net/>
- **The VMware Infrastructure Perl ToolKit**
  - <http://sourceforge.net/projects/viperltoolkit>
- **The Virtual Infrastructure Java API**
  - <http://vijava.sourceforge.net/>
- **Zimbra**
  - <http://www.zimbra.com/community>
- **Redis**
  - <http://code.google.com/p/redis/>
- **VMware SVGA device developer kit**
  - <http://vmware-svga.sourceforge.net/>
- **vCloud open SDK**
  - <http://code.google.com/p/jclouds/>
- **DynamoRIO**
  - <http://code.google.com/p/dynamorio/>
- **VMware View open client**
  - <http://code.google.com/p/vmware-view-open-client/>
- **springsource**
  - <http://www.springsource.org/projects>

# Open Standards Initiatives

---

- **OVF – Open Virtual Machine Format**

- <http://www.vmware.com/appliances/getting-started/learn/ovf.html>

- **VDDK – Virtual Disk Development Kit**

- <http://www.vmware.com/support/developer/vddk/>

- **VMI – Virtual Machine Interface**

- <http://www.vmware.com/technical-resources/interfaces/paravirtualization.html>

- **vCloud API**

- <http://www.vmware.com/go/vcloudapi>

- **SPECvirt**

- [http://www.spec.org/virt\\_sc2010/](http://www.spec.org/virt_sc2010/)

- **CIM**

- <http://www.vmware.com/support/developer/cim-sdk/>

# VMI Standard

---

## ■ VMI – Virtual Machine Interface

- VMI interface works transparently both natively as well as in a VM, when run inside a VM provides increased performance
- VMware's VMI proposal led to the development of paravirt-ops with Linux Kernel 2.6.20, VMI backend code was Integrated with Linux Kernel 2.6.22
- Ubuntu Feisty Fawn (7.04) and Gutsy Gibbon (7.10) shipped with VMI enabled kernel

## ■ Performance gains of CPU para-virtualization approach is no longer relevant

- 64-bit computing and Hardware support for CPU and MMU virtualization offsets most of the virtualization overhead
- VMware I/O devices are already para-virtualized for better performance
- VMI will be deprecated

# Guest-Host communication

---

## ■ HGFS – Host-Guest File System

- Abstracts the underlying file system, similar to NFS but doesn't require networking

## ■ VMCI Sockets

- <http://www.vmware.com/support/developer/vmci-sdk/>
- Similar to Unix, BSD sockets

## ■ Guest API & SDK

- Guest API provides function calls to collect Virtual machine CPU and memory usage resource statistics from within the guest
- Guest SDK runtime component (shared library) is shipped with vmware-tools
- C and Java Programming languages supported
- SDK Could be downloaded from <http://www.vmware.com/support/developer/guest-sdk/>

# Virtual Appliances

---

- **Virtual Appliance Marketplace launched in 2006**
  - Pioneered a new model to ship software
  - Free VMware player was released in 2005
  - Powered by community and ISV
  
- **Within less than a year 500 virtual appliances were created**
  - All leveraging open source operating systems, with Linux being the number one choice
  
- **Currently there are over 1450 virtual appliances**
  - Over 500 of these appliances are pure Linux distributions
  - Ubuntu desktop is one of the most popular virtual appliances



# VMware Studio

---

## ■ Virtual Appliance authoring tool

- Supports Centos, RedHat Enterprise Linux, SUSE Linux Enterprise Server and Ubuntu distributions
- Handles package dependencies
- Supports OVF
- Eclipse IDE integration

## ■ Recommendations for building a Linux VM appliance

- Start with base installation (JeOS)
- You could trim down Linux kernel – remove irrelevant features
- Bundle open-vm-tools
- Apply best practice recommendations and tunings

## Other projects

---

### ■ VMware vCenter Converter

- Can convert physical machines into virtual machines (P2V)
- Linux P2V is supported for Redhat Enterprise Linux, SUSE Enterprise Linux and Ubuntu Linux distributions

### ■ vSphere Management Assistant (vMA)

- CentOS based Linux Appliance for managing VMware ESX hosts remotely

### ■ vCenter server for Linux

- [http://communities.vmware.com/community/beta/vcserver\\_linux](http://communities.vmware.com/community/beta/vcserver_linux)
- Currently in Technology preview

### ■ vSphere client for Linux

- Considering platform independent client for future versions

## VMware - Novell Partnership

---

- **Announced a joint support agreement recently**
  - Qualifying purchase of VMware vSphere automatically entitles support for SLES including subscription to patches at no additional cost
  - Direct SLES technical support through VMware

# SYNOPSIS

- Evolution of VMware and Linux
- VMware initiatives for Linux
- Running Linux guest OS effectively
- Linux guest in the Cloud

# Getting the best performance

---

## ■ Upgrade!

- CPU: Newer CPUs are not only vastly faster than the older ones but they also have greatly reduced virtualization overhead
- Guest operating system: performance issues are often ironed out in patch/update releases
- Hypervisor: We are constantly improving our efficiency and scalability
- VMware tools: new enhancements to the para-virtualized drivers are being constantly made

## Getting the best performance

---

### ■ Install/update vmware-tools

- Many customers forget to install vmware-tools since the VM works out-of-the box
- Kernel update breaks binary drivers, vmware-config-tools.pl needs to be run after kernel update
- vmxnet3 driver is NAPI compliant and supports TSO, LRO, Jumbo Frames
- pvscsi driver is efficient than LSI Logic and is recommended for I/O intensive VMs. You will need to create a new initrd image to use pvscsi for the boot disk

# Best practices

---

## ■ Setting the right guest OS type

- Some optimizations are based on the guest OS type, better time keeping for instance

## ■ Choosing the right kernel

- 64-bit kernel will generally give you better performance
- Avoid 4G/4G kernel for 32-bit VMs

## ■ Choosing the right virtual Hardware (vHw) version

- Changing virtual Hardware version is akin to changing motherboard
- Upgrading the virtual hardware upgrades the VM for more capabilities (for instance hot-add, pvscsi, vmxnet3 etc)

# Best practices

---

## ■ Swap partition

- Correct sizing is important to avoid OOM killer when the balloon driver kicks in
- You can also avoid OOM killing by disabling memory over-commit in the guest beyond the swap capacity (`/proc/sys/vm/overcommit_memory`)

## ■ Swappiness

- Default is 60 for most distributions
- You may see improved performance by setting it to zero

## ■ Large Pages

- VMware supports large pages – it improves performance greatly, but just like any other resource it should be used judiciously
- Linux mlocks large pages effectively and it cannot be swapped or used for other purposes



# Best practices

---

## ■ Scheduling of cron Jobs

- Scatter cron jobs to avoid consuming resources at the same time

## ■ Virtual SCSI adapter Queue depth

- SCSI adapter queue depth is optimized for local attached disks, if you use SAN to store your virtual disks, increasing queue depth will help

## ■ I/O Scheduler policies

- Guest I/O Scheduling policy is far too distant from the physical disk, Noop scheduling policy works better

## ■ Partition alignment

- Misaligned partitions doubles read I/Os and it can cause up to 30% performance degradation

# Linux Guest OS issues

---

## ■ **gettimeofday**

- Some kernels use CPUID instructions to serialize the processor pipeline before issuing RDTSC
- CPUID instructions causes VM exits and hurts performance in virtualized environment

## ■ **SCSI disk I/O retry timeout**

- Default timeout may not be sufficient (2.4 kernels have a hardcoded timeout value of 60 seconds, 2.6 kernels have a hardcoded timeout value of 30 seconds)
- On newer kernels (2.6.13 and above) VMware tools automatically increases the timeout to 180 seconds

## ■ **Congestion Control algorithm**

- Some congestion control algorithm are very sensitive to latency and constricts TCP window aggressively

# SYNOPSIS

- Evolution of VMware and Linux
- VMware initiatives for Linux
- Running Linux guest OS effectively
- Linux guest in the Cloud

# Transformation to Cloud Computing

---

- **Cloud computing is one more level of abstraction**
  - 1990 - 2000 – PC era - Single OS instance on a single Machine
  - 2000 - 2010 – Server consolidation era - Multiple OS instances on a single machine
  - 2010 and beyond – Multiple datacenters on multiple machines - Datacenter Consolidation

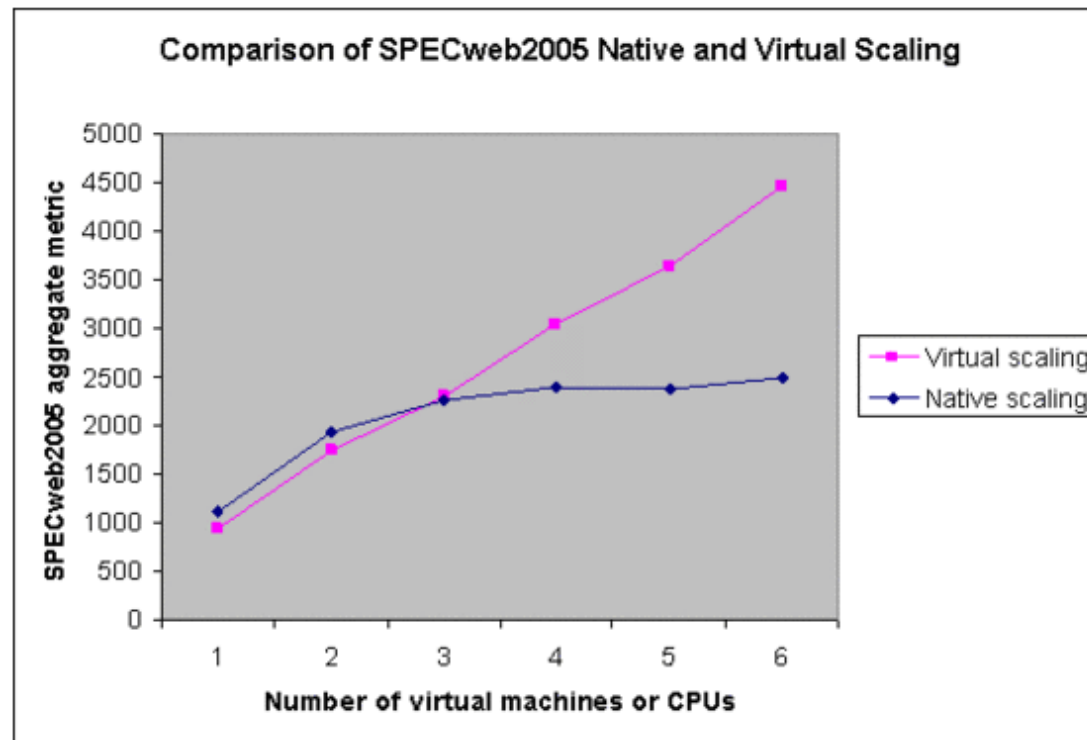
# Cloud Computing

---

- **Cloud computing is all about scale and efficiency**
  - VM density per host is an important metric
  - Even small efficiency gains play an important role at cloud scale
- **New ways of running workloads**
  - Horizontal scaling
  - NoSQL movement, in-memory database, distributed caching etc
- **On-demand and agile computing**
  - Faster provisioning, vMotion, Storage vMotion etc.

# Horizontal Scaling

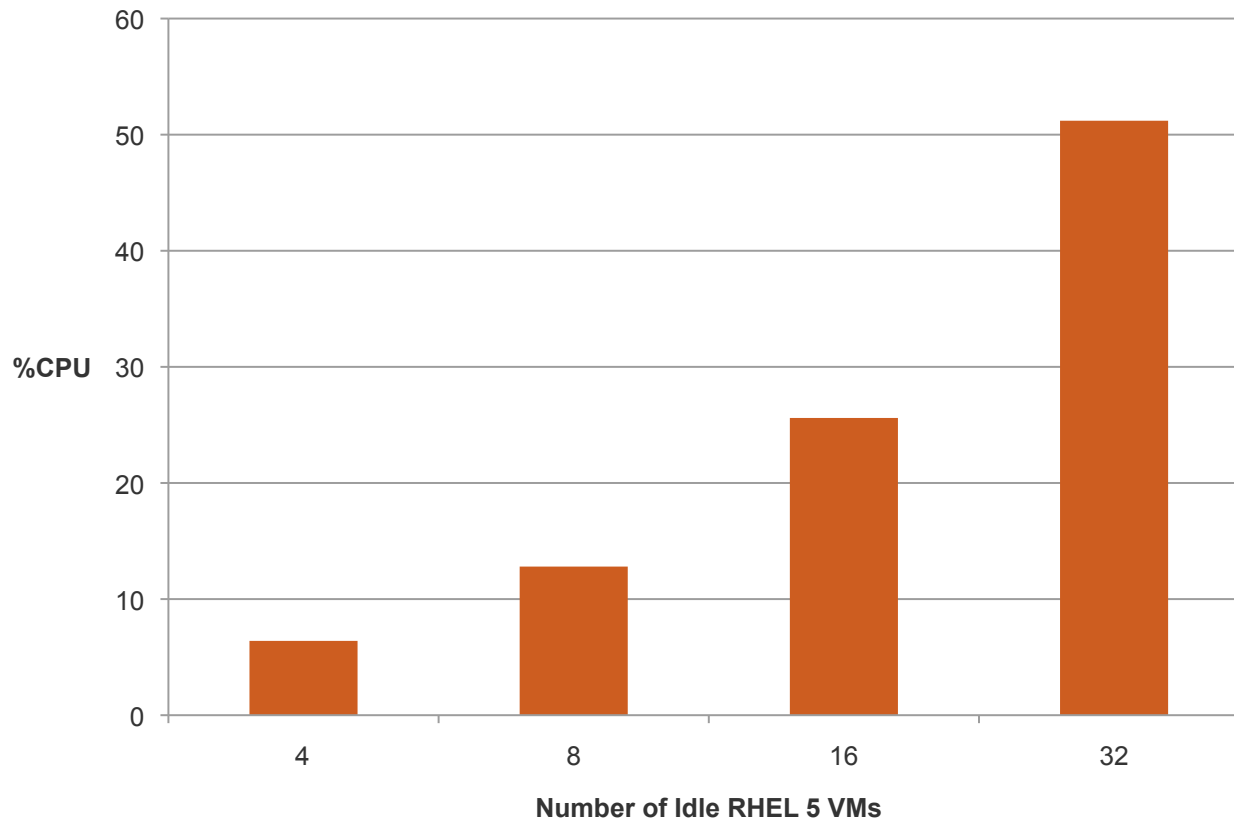
- Apache/PHP software stack scaling



# Timer Interrupt

---

1000 Hz timer Interrupt Overhead on Intel Nehalem Processor



# Timer Interrupt

---

## ■ Evolution of Timer Interrupts in the mainline kernel

- 2.4 kernel and earlier uses 100 Hz
- 2.6.0 – 2.6.13 - uses 1000 Hz
- 2.6.14 onward uses 250 Hz
- 2.6.18 and above has tick-less timer support

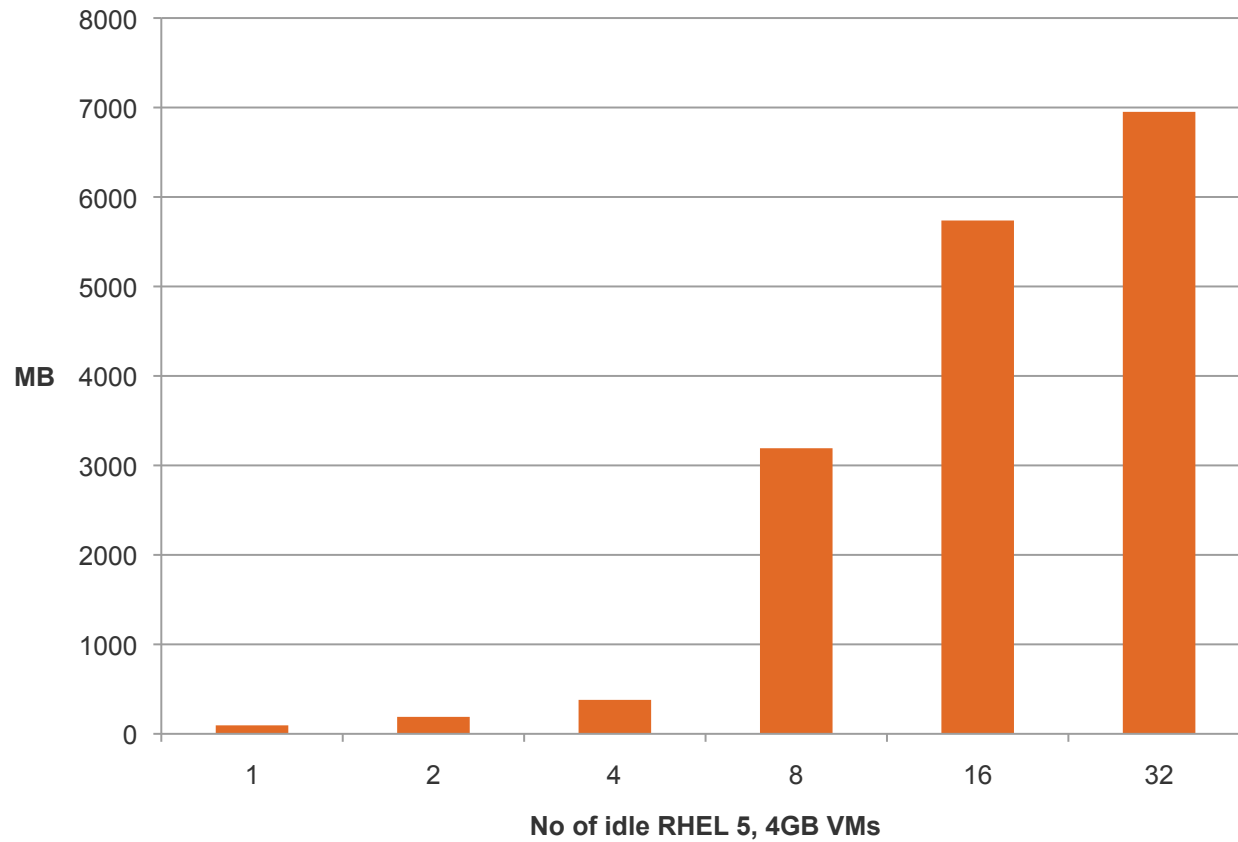
## ■ Timer interrupt rate on Linux Distributions

- RHEL distributions use 1000 Hz
- SLES and Ubuntu distribution uses 250 Hz
- RHEL 5.1 (and above) and RHEL 4.7 supports “divider=” kernel boot option to reduce timer interrupt rate



# Page Sharing Benefits

## Memory Savings with Page Sharing



# Linux VM Cloning issues

---

## ■ Install-time configuration

- Some distributions install different kernel based on CPU type, available memory etc.
- VMs are more agile and dynamic - processor specific features will break when the VMs migrate across machines/datacenters

## ■ Guest reconfiguration – (sysprep equivalent)

- No single standard, each Linux distribution has its own way of storing configuration files
- MAC address based device naming doesn't work well when the VM is cloned

## Summary

---

- **VMware has the broadest support for Linux Guest Operating Systems**
- **VMware tools is open sourced and is getting merged into Linux kernel to further enhance the out-of-the box experience**
- **Virtual Appliance concept is gaining momentum, open Source operating systems play an important role**
- **Efficiency and scale play an important role in cloud computing**
- **Kernel changes and certain OS default settings can affect virtual environments**

## SEE ALSO

- **Installing and configuring Linux Virtual machines**
  - [http://www.vmware.com/files/pdf/linux\\_install\\_config.pdf](http://www.vmware.com/files/pdf/linux_install_config.pdf)
- **Time keeping in VMware virtual machines**
  - [http://www.vmware.com/pdf/vmware\\_timekeeping.pdf](http://www.vmware.com/pdf/vmware_timekeeping.pdf)
- **Numerous other whitepapers**
  - <http://www.vmware.com/technical-resources/>