

CLUE – XEN Discussion

Mike Friesenegger
Linux/Open Source Specialist

October 13, 2006



Novell.[®]

Novell® Offers Solutions in Five Key Markets

- 1 Data Center
- 2 Security and Identity
- 3 Resource Management
- 4 Workgroup
- 5 Desktop

From Community to Enterprise

openSUSE feeds SUSE Linux Enterprise offerings



Various open source projects

Rapidly growing in number of projects and their capabilities

Isolated, uncoordinated efforts



SL10.1

SL10.0

SL9.3

SL9.2

SL9.1

Linux for enthusiasts and individual end users

Technical previews and leading edge applications

Novell

Enterprise Desktop

Enterprise Server

Hardened, Secure, Supported

Certified for leading hardware and software platforms

7-year code maintenance guarantee

Suse Linux Enterprise Server and Desktop 10 shipped July 17, 2006!



JORGE MADE
LINUX FASTER

WE MADE IT
ENTERPRISE-WIDER



Your Linux is ready.

Introducing SUSE, Linux Enterprise 10 from Novell. Built by a global community and secured, supported, tested and proven by Novell. From the desktop to the data center, SUSE Linux Enterprise 10 is the Linux platform that brings discipline to open and innovation to the enterprise. So it's more than cool and secure. It's the Linux you've been waiting for.

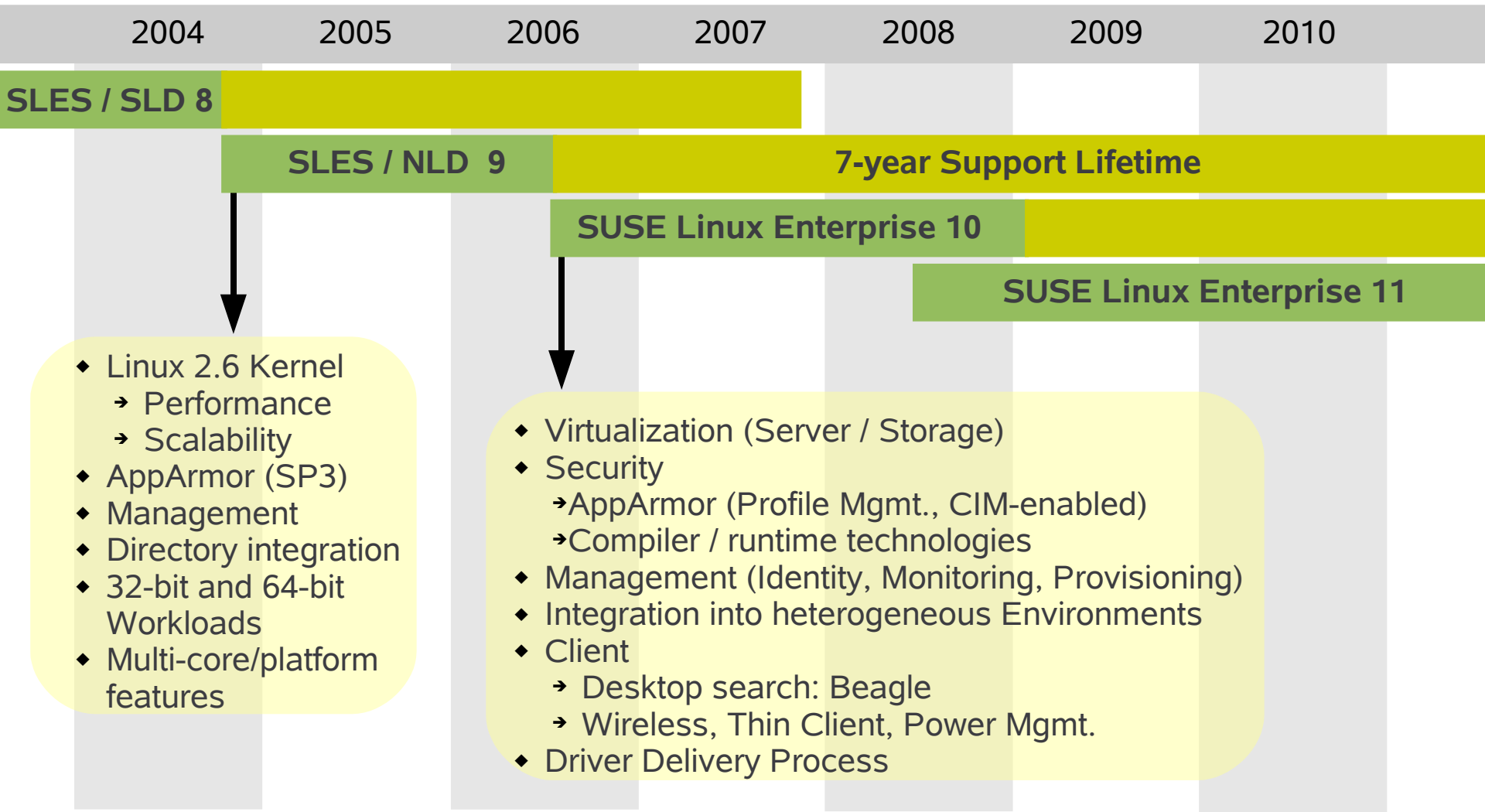
Get it at www.novell.com/linux

Novell
This Is **Your** Open Enterprise.™

Get a copy at

www.novell.com/linux

SUSE Linux Enterprise Roadmap



Novell's SUSE Linux Enterprise 10

Server

SUSE Linux Enterprise Server 10
(SLES 10)

Novell Open Enterprise Server
(OES)

Desktop

SUSE Linux Enterprise Desktop 10
(SLED 10)

SUSE Linux Enterprise
Point of Service 10 (SLEPOS)

SUSE Linux Enterprise 10 (“Code 10”)
Common Code Base

SUSE® Linux Enterprise Desktop 10

- **Exciting new desktop tools**

- Usability with improved menus
- Beagle, f-spot, Banshee, Tomboy

- **OpenOffice.org 2.0+ Novell Edition**

- Improved format interoperability
- VB macro support
- Performance improvements

- **Deployment and management**

- Thin client including diskless
- Central management and deployment
- Desktop lockdown

- **New hardware support**

- iPod, cameras, USB Storage
- IPMI and power management
- Video acceleration
- PlugNPlay!

- **Interoperability**

- Microsoft Active Directory
- Novell eDirectory

- **Multi-media support**

- Support audio/video formats
- Clean intellectual property

Desktop Subscriptions

	Basic	Standard	Priority
1 Year	\$50	--	--
3 Year	\$125	--	--

- Simple
 - Per device
 - Installation assistance
 - Updates
 - Ala carte support

SUSE® Linux Enterprise Server 10

• Virtualization

- XEN 3 for SMP, PAE and 64-bit CPUs
- Server consolidation, compatibility
- Scheduled maintenance

• Deployment and management

- CIM providers and improved YaST
- Pattern deployments

• Security

- Application security, EAL4+
- Network detection, monitoring
- Secure encapsulation with Xen

• Performance and scalability

- 1024 CPUs, 10+TB memory

• New hardware and update drivers

- X86, x86-64, ia64, ppc, ppc64 and s390x
- New storage, network, graphics drivers
- Hotplug improvements

• Storage foundation

- Clustered file system OCFS2, EVMS
- HA with Heartbeat 2 (to 16 nodes)
- NFS v4
- iSCSI target and initiator
- DRDB for disaster recovery

Subscription Pricing

Support Included	Basic	Standard	Priority
1 Year	\$349	\$799	\$1,499
3 Year	\$873	\$1,998	\$3,748
<hr/>			
1 Year	\$11,999	\$15,000	\$18,000
3 Year	\$29,998	\$37,500	\$45,000



Non-mainframe

- Valid for
 - x86 (32-bit), x86-64 (AMD64 & EM64T), Itanium 2 (IA), IBM POWER (PPC)
- Up to 32 processor sockets per server
- No additional cost per virtual image
- No additional cost for AppArmor, Xen, Storage Foundation

Mainframe

- No longer priced by type of mainframe
- Priced per engine (IFL)
- SLES 10 only has 64-bit for mainframe, no 31-bit (as per IBM request)
- Prior version rights include 31-bit for SLES8 & SLES9

Xen Architecture

Xen Background

Small VMM supporting
para-virtualization (modified guests)
full virtualization (shrink wrapped guests)

Excellent performance (~5% degradation w/paravirtualization)

Originally a research project from University of Cambridge

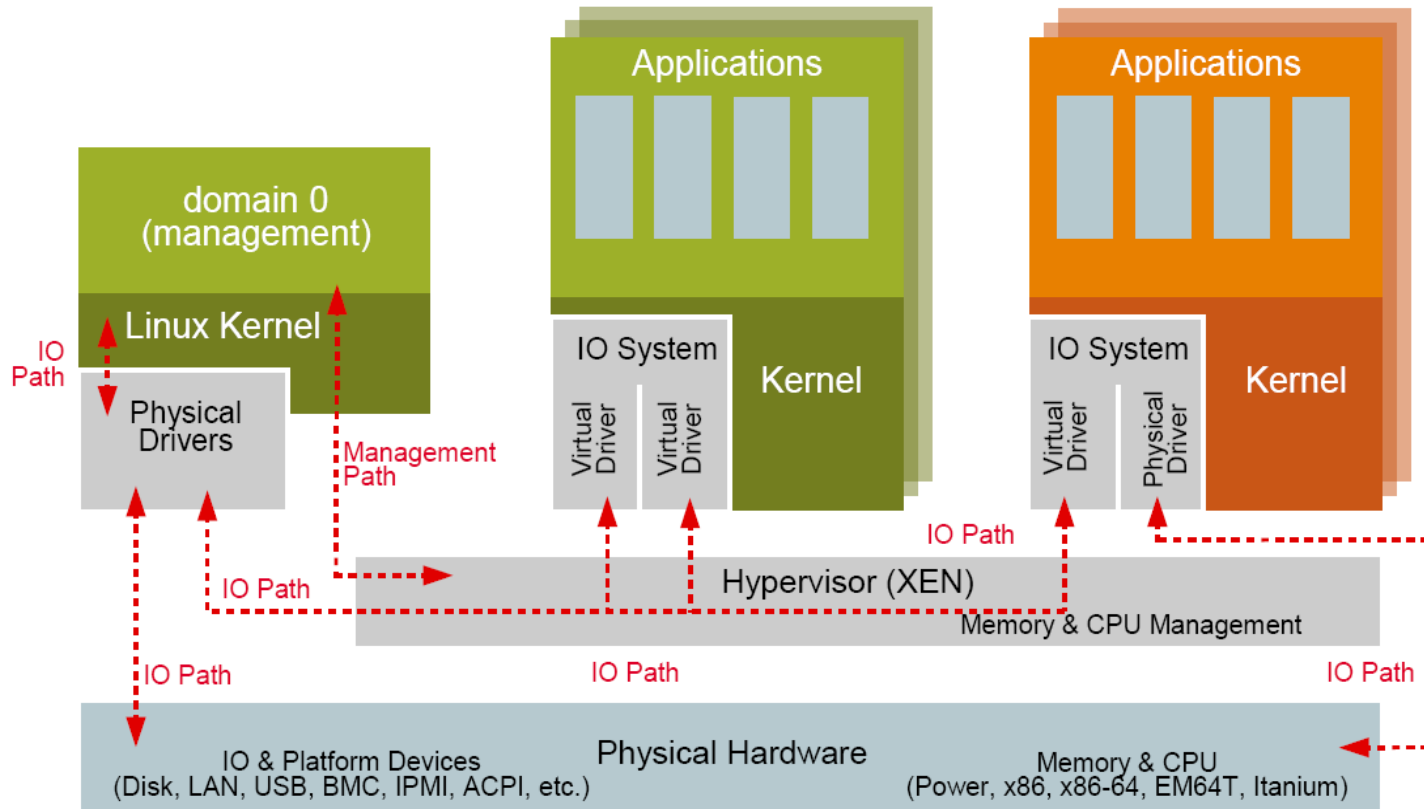
Open source

Xen 2.0 released November 2004

Xen 3.0.0 released December 2005

Xen 3.0.2 release planned mid March 2006 (SLES 10 Target)

Xen Architecture (Showing Para-virtualized Guests)



Processor Support

Xen 2.0:

- IA-32

Xen 3.0.x:

- IA-32
- IA-32 + PAE
- x86-64
- x86 hardware assist 32/64 (Intel VT; AMD Pacifica)

Future:

- Itanium
- Power

Para-virtualized Guest OS Support

Remember: Xen supports para-virtualization (kernels performance enhanced to run on Xen)

Currently ported to Xen:

- Linux 2.4
- Linux 2.6
- NetWare 6.5 +
- NetBSD
- FreeBSD
- Plan9
- OpenSolaris

In the works:

- ReactOS
- Others?

Fully Virtualized Guest OS Support

With appropriate hardware support Xen can host shrink wrap guest operating systems

In SLES 10 Xen supports

- Intel VT

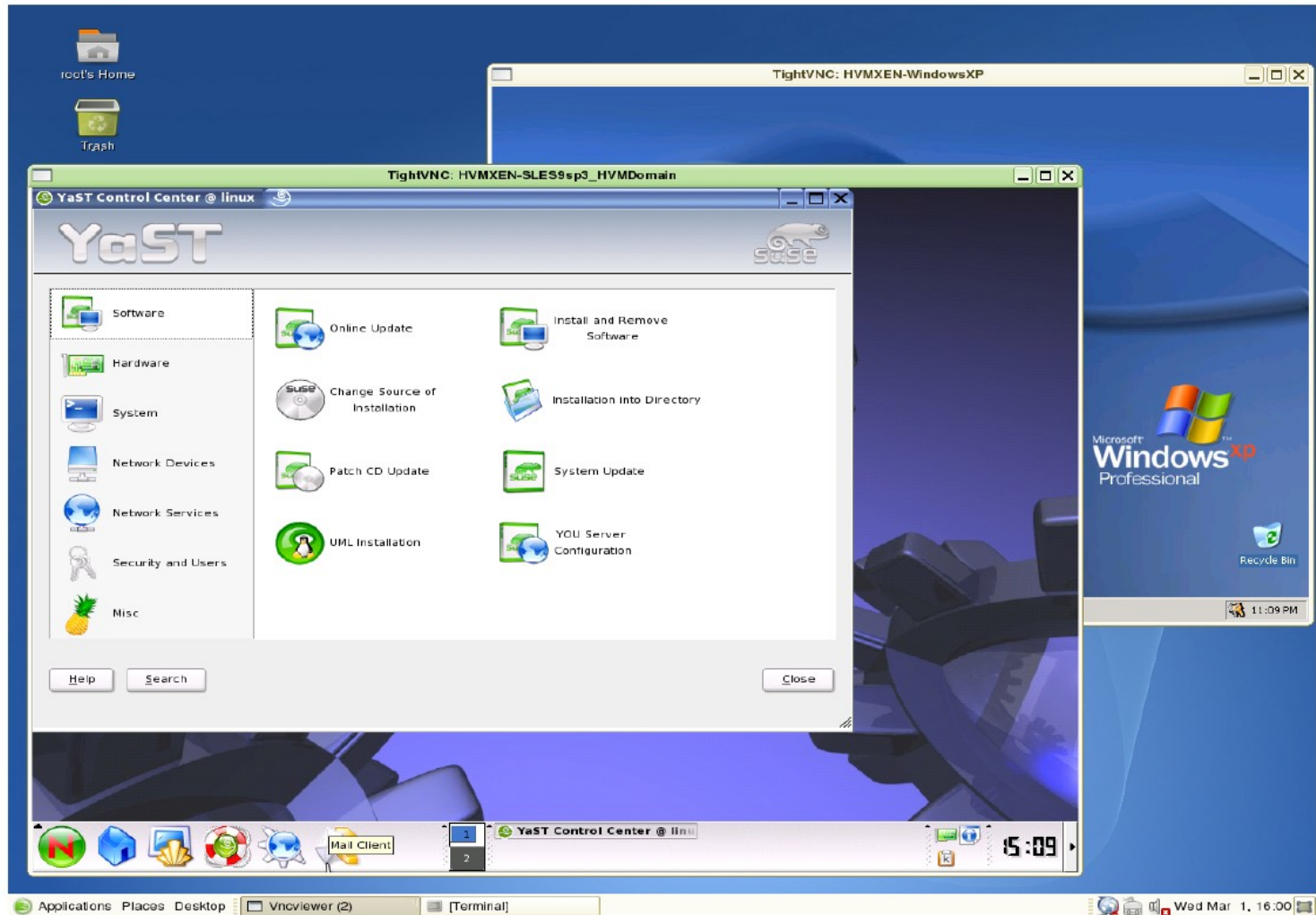
- AMD Virtualization (formerly known as Pacifica)

Fully Virtualized OS Support

A number of shrink wrap operating systems have successfully run on Xen

SLES 10 Host

Fully Virtualized Guests (SLES 9 and Windows XP)



Using Xen

Terminology

Domain: A container for a running virtual machine.

Domain 0: The first domain. Privileged to manage other domains. a.k.a. “dom0”.

Unprivileged domain: Any domain other than domain 0. Cannot manage other domains. a.k.a. “domU”.

Driver domain: A domain that contains physical drivers. Usually this is just domain 0.

Physical driver: A device driver (usually in the driver domain) that talks to the hardware.

Virtual driver: A device driver (usually in a domU) that fulfills requests by going to the physical driver.

Installing Xen

From binary:

- Install Linux distro with support for Xen
 - SLES 10
- During installation, select “xen” and “kernel-xen” packages

From source:

- Install Linux distro
- Get Xen source:
<http://www.cl.cam.ac.uk/Research/SRG/netos/xen/downloads.html>
- make world
- make install

Prerequisites:

- bridge-utils
- python
- python-devel (to build)

VM Installation using YaST

Novell makes VM installation simple with YaST

VM Installation using YaST on SLES 10

- Support for para-virtualized guests

- Support for fully virtualized guests

YaST does not address VM life cycle management

- This will be a feature of Novell's data center management suite

Daemons

xend:

- Required
- Runs in dom0
- Must be running to access management commands
- Manages VM configuration

xenddomains:

- Optional
- Starts VMs when dom0 starts
- Stops VMs when dom0 stops

VM Disks

Dom0 manages a mapping between a “virtual” disk (within the guest domain) and “physical” disk (within a driver domain).

“Physical” disk can be:

- iSCSI
- local physical disk or partition
- file via loopback
- file via a cluster filesystem
- ...

How to create a disk?

- YaST Xen tool
- dd + mkreiserfs + mount + YaST “install to directory”
- dd from existing server's disk to file
- mk-xen-rescue-image.sh

VM Configuration Files

```
cp /etc/xen/xmexample1 /etc/xen/MyVm
```

Fields to change:

if manual kernel:

kernel: The location of the kernel, as visible from dom0.

initrd: (*nix-centric) Extra module for kernel, as per Multiboot specification.

if domUloader:

bootloader: /usr/lib/xen/boot/domUloader.py

bootentry: paths of kernel and initrd to extract from VM's disk.

memory: Memory to give the guest, in megabytes.

name: The name must be unique among running VMs.

vif: List of virtual network interfaces.

disk: This maps the disk device (visible in domain 0) to the device in the guest.

root: (*nix-centric) Device containing root filesystem.

extra: (*nix-centric) Extra kernel arguments, e.g., runlevel.

Xen Command-line

Log in as root

Use “xm” command (Xen Management)

```
xm list
```

```
xm info
```

```
xm create -c -f configfile
```

```
xm shutdown id
```

```
xm destroy id
```

```
xm console id
```

Disconnect from VM's console with Ctrl-]

VM Migration

Migration

- Moving a VM from one physical machine to another, without interrupting the VM's state.
- IP address follows the VM, so networking is not interrupted

Live Migration

- Migration with minimal down-time (10's of ms)
- Streams VM's pages over network while VM is still running

Disk must still be visible on destination! (e.g., use iSCSI)

```
xm migrate id destination-host  
xm migrate -l id destination-host
```

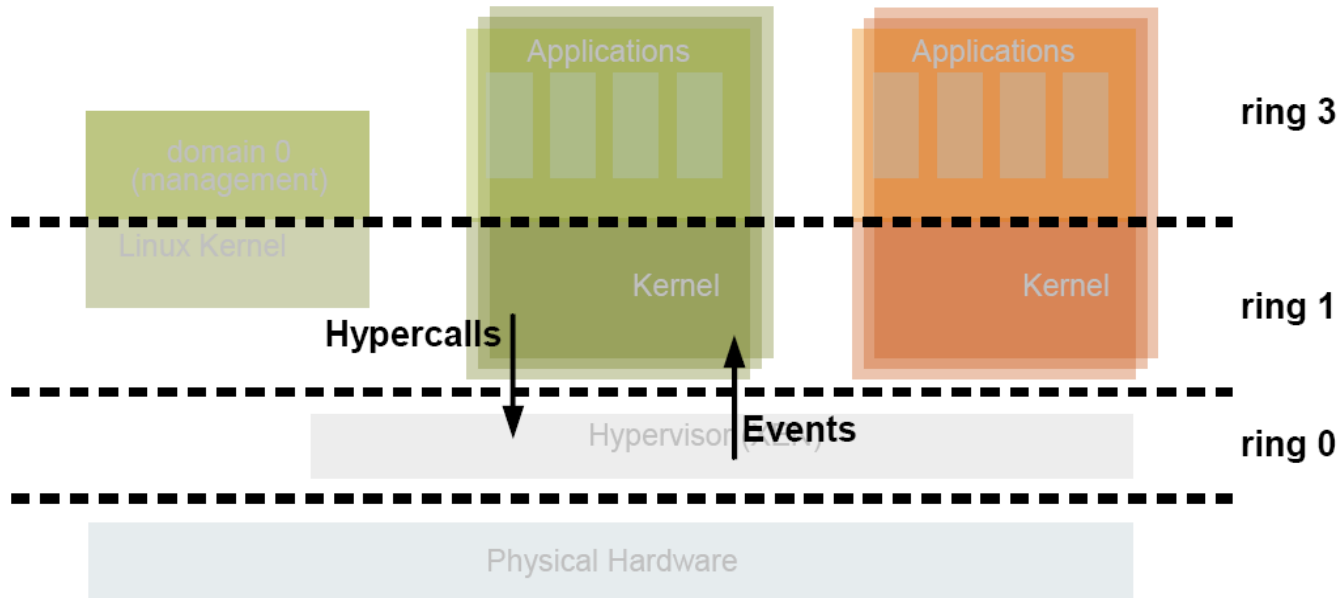
Inside Xen

Privilege Rings

Xen runs at ring 0 (highest privilege)

All domains run at rings 1 - 3.

- Kernel is ring 1
- User-space is ring 3



Linux: Xen sub-architecture

Linux is ported to a new “xen” sub-architecture.

Processor-specific sub-architectures: i386/xen, x86_64/xen

Abstracts privileged instructions:

- Control registers
- Page table operations
- IDT, GDT, LDT changes
- Debug registers
- Timers
- Disabling interrupts
- Segmentation
- etc...

Hypercalls

Para-virtualized kernel calls into Xen via a “hypercall” (software interrupt).

Without Xen:

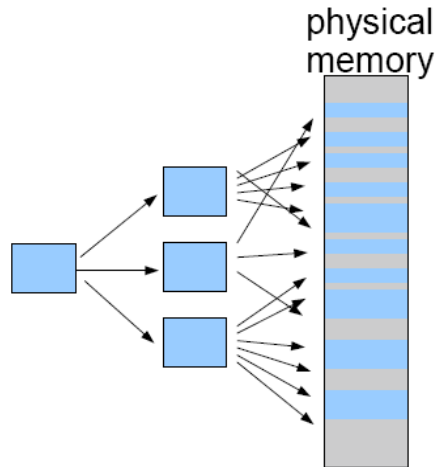
```
idle:
    hlt
```

With Xen:

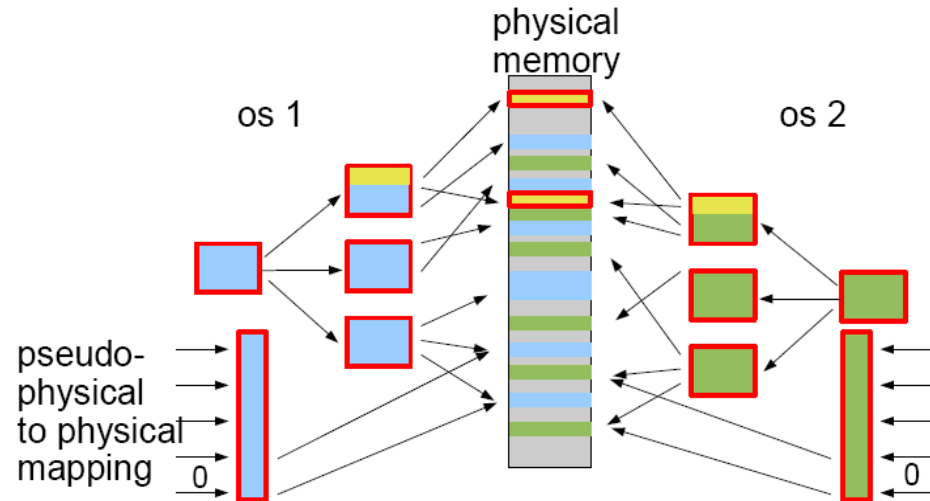
```
idle:
    mov $HYP_schedop, %eax
    mov $SCHEDOP_yield, %ebx
    int $0x82
```

Memory

Without Xen:



With Xen:



Without Xen: Virtual to physical mapping.

With Xen: Virtual to “physical” to machine mapping. Physical memory is discontinuous. Page tables are read-only*.

Optimization: Batch page table updates.

* Can use shadow page tables for lower porting effort

Interrupts and Exceptions

Hardware interrupts are replaced with light-weight events.

Interrupt Descriptor Table is registered through Xen:

- Can't specify ring 0
- Can't specify full 4 gig segment
- Page faults must indirect through Xen, to get cr2

“Fast” system call

Failsafe callback

Drivers

Xen virtualizes PCI config space.

Driver support comes from one or more driver domains (usually dom0).

“Front end” interfaces between Xen and virtual driver in consuming domain

“Back end” interfaces between Xen and physical driver in driver domain.

Xen (and xend) act as mediators.

Xen supports virtualized disk and network adapters. Other types of devices (e.g., sound, video) cannot be shared (yet?)

Debugging with Xen

Debug registers are virtualized, so kernel debuggers should work under Xen.

Xen has a [simple and unmaintained] debugger:

```
debug=y debugger=y make world
```

Modify grub:

```
kernel (hd0,0)/boot/xen.gz com1=115200,8n1 pdb=com1H
```

Serial line needs a “splitter”. Included in SUSE's Xen, or search xen-devel mailing list for serial-split.c. pdb is “high”; printk and console is “low”.

Connect to Xen with gdb:

```
target remote localhost:12011
```

Management

The Problems

Server consolidation means many VMs!

- From managing servers to managing VMs
- Peak usage could be *worse*

Hardware monitoring

- CPU load, CPU temp, memory failures, ...
- Trivial decisions should be automated

Virtual machines depend on virtual storage

- Tie VM config to disk image
- Patch management of (potentially inactive) disk images

Identity and roles

- Who can create a VM? Migrate a VM? Destroy a VM?

Data Center Management for the Open Enterprise

Virtualization is cool...

...but it's just an enabling technology!

The Novell Formula:

Virtual Machines +

Virtual Storage +

Resource Management +

Identity Management

Novell has core business strengths in each of these areas

Stay tuned as Novell neatly ties all these technologies together with management tools based on open standards

Novell®

Unpublished Work of Novell, Inc. All Rights Reserved.

This work is an unpublished work and contains confidential, proprietary, and trade secret information of Novell, Inc. Access to this work is restricted to Novell employees who have a need to know to perform tasks within the scope of their assignments. No part of this work may be practiced, performed, copied, distributed, revised, modified, translated, abridged, condensed, expanded, collected, or adapted without the prior written consent of Novell, Inc. Any use or exploitation of this work without authorization could subject the perpetrator to criminal and civil liability.

General Disclaimer

This document is not to be construed as a promise by any participating company to develop, deliver, or market a product. Novell, Inc., makes no representations or warranties with respect to the contents of this document, and specifically disclaims any express or implied warranties of merchantability or fitness for any particular purpose. Further, Novell, Inc., reserves the right to revise this document and to make changes to its content, at any time, without obligation to notify any person or entity of such revisions or changes. All Novell marks referenced in this presentation are trademarks or registered trademarks of Novell, Inc. in the United States and other countries. All third-party trademarks are the property of their respective owners.

