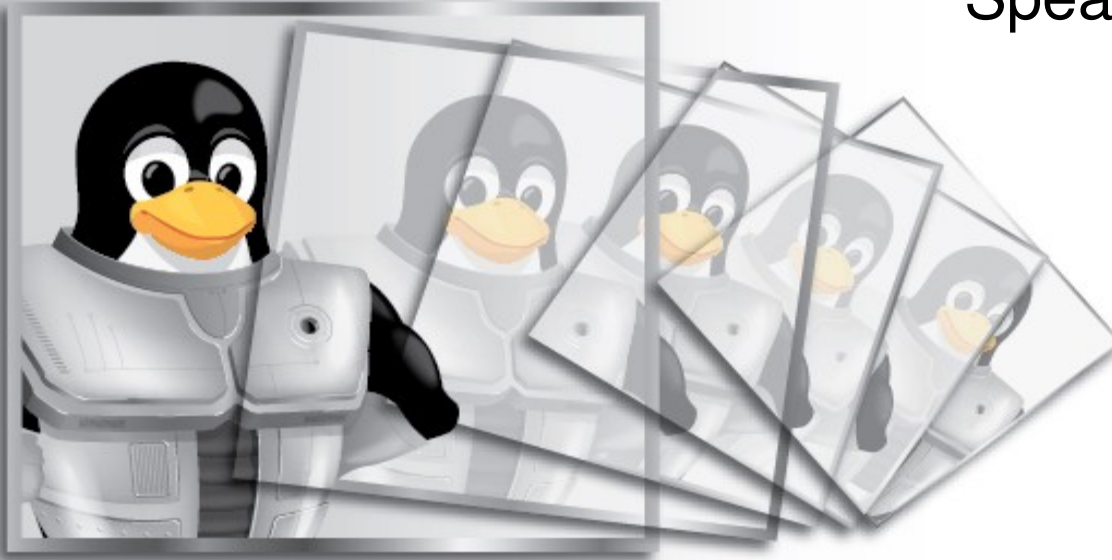Speaker: Dan Magenheimer

ORACLE®

VM

ONE COMPLETE SOFTWARE STACK.
ONE SOURCE FOR SERVER VIRTUALIZATION AND LINUX.
ONE CALL FOR SUPPORT.

Topic : The Xen of Virtualization on Linux

13th May 2008 @ **CLUE** **Colorado Linux Users and Enthusiasts**

What is virtualization?

A brief history of computing…

Three immutable(?) laws of computing

1. Wirth's law: Software gets slower faster than hardware gets faster

2. …

3. ..

Three immutable(?) laws of computing

1. Wirth's law: Software gets slower faster than hardware gets faster: "**What Grove giveth, Gates taketh away**"
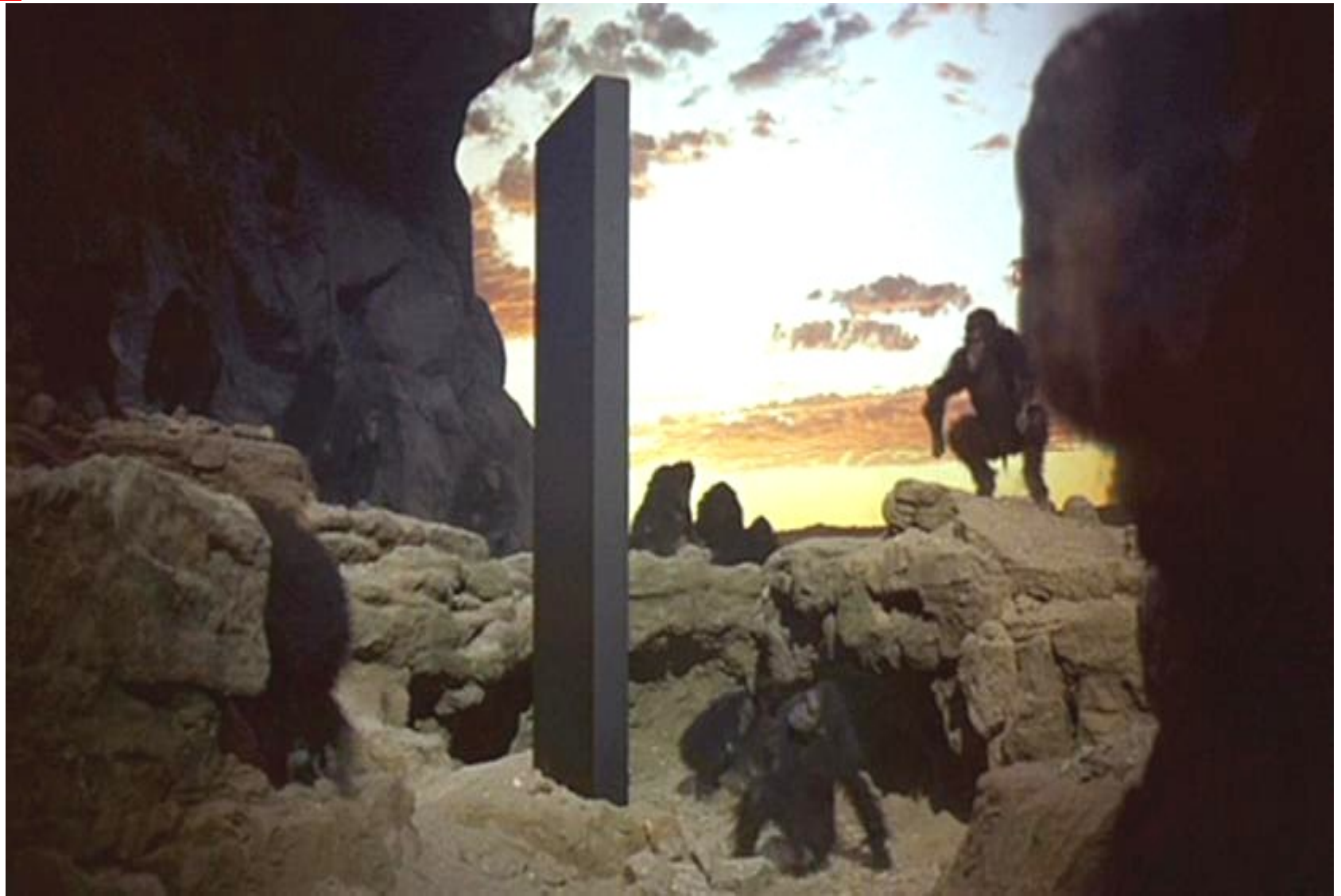
2. …

3. …

Three immutable(?) laws of computing

1. …

2. Everything in computing can be solved by *adding another layer of indirection* (/ **abstraction**)

3. …

ORACLE®

Three immutable(?) laws of computing

1. …

2. …

3. There's nothing new in computing that IBM didn't already invent 30 (40, 50) years ago

Three immutable laws of computing -- or are they (immutable?)

1. Software gets slower faster than hardware gets faster, aka "*What Grove giveth, Gates taketh away*"

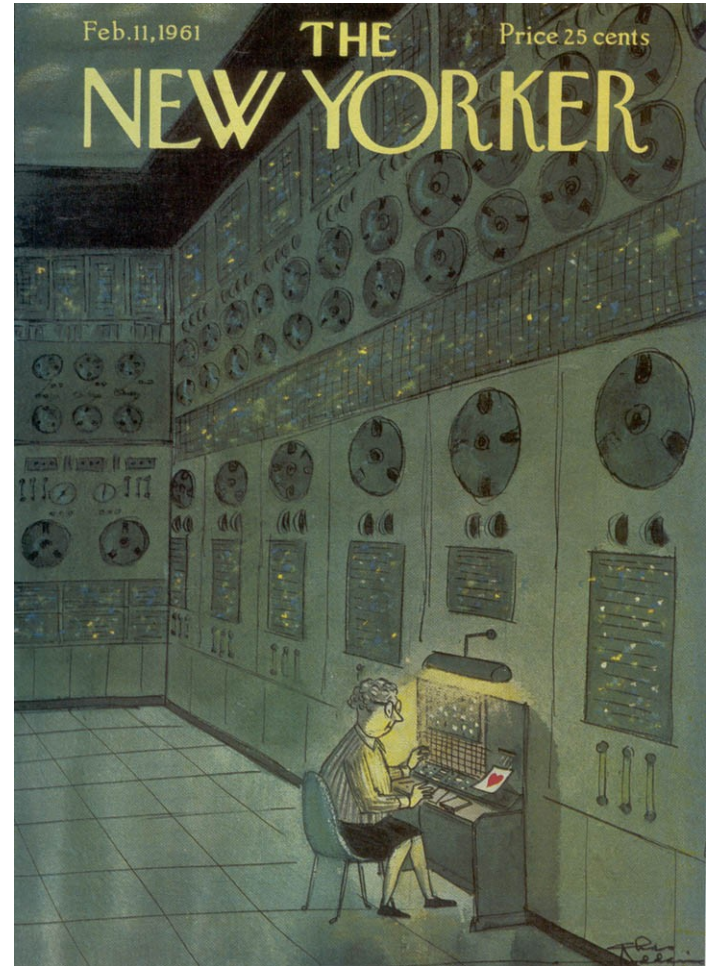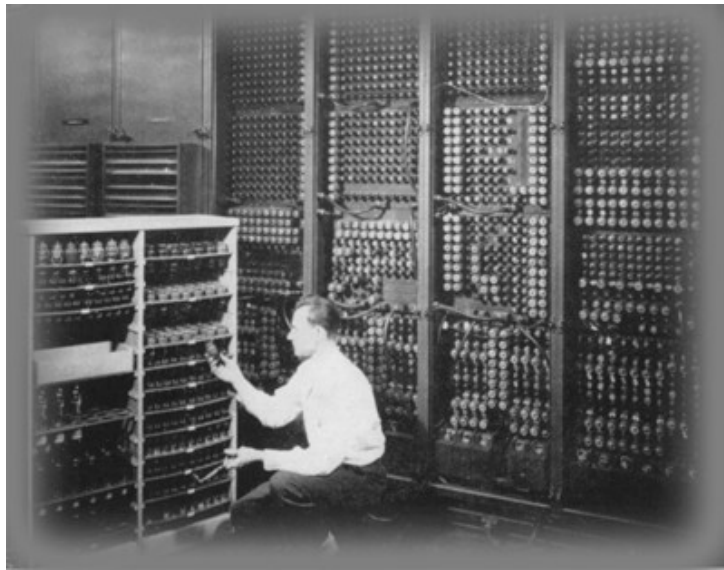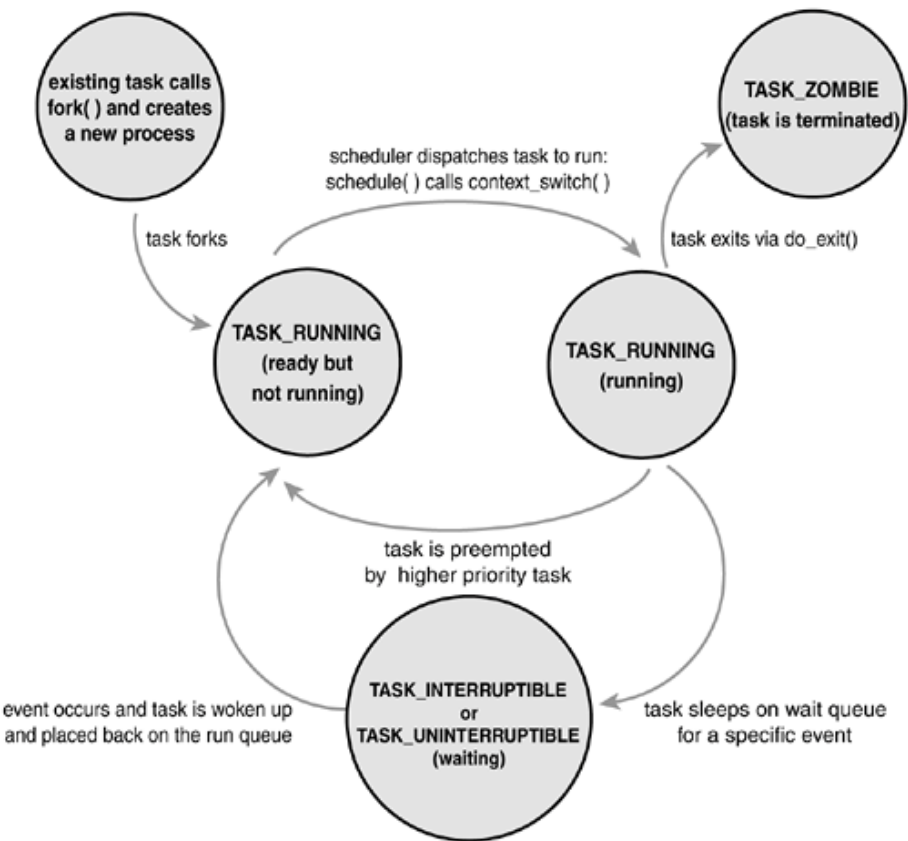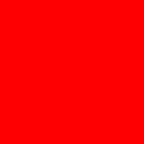2. Everything in computing can be solved by adding *another layer of indirection* (/ abstraction)

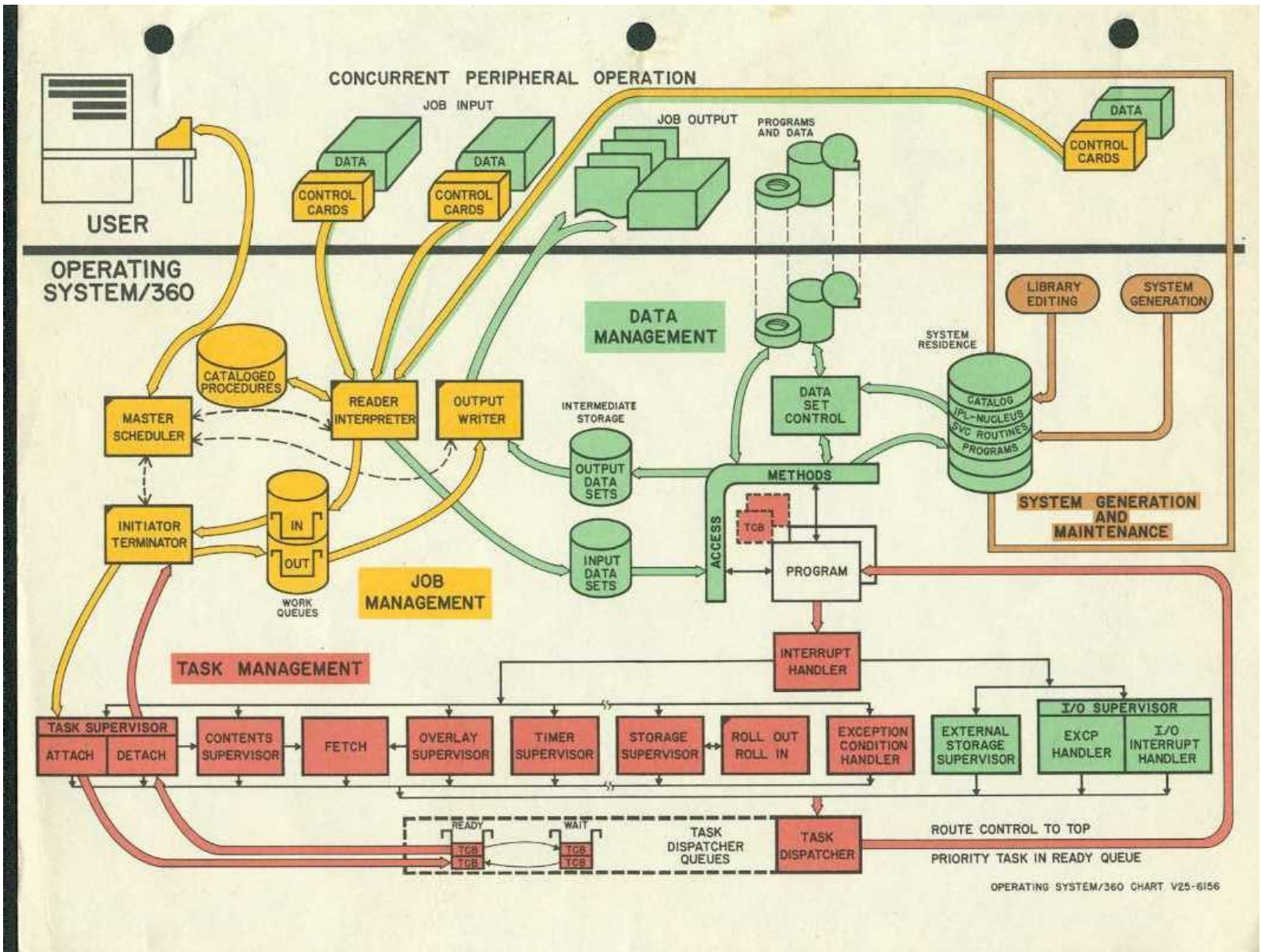3. There's nothing new in computing that IBM didn't already invent 30 (40, 50) years ago

THE NEW YORKER

ORACLE

Process state transition diagram:

- existing task calls fork( ) and creates a new process
- task forks → TASK_RUNNING (ready but not running)
- scheduler dispatches task to run: schedule( ) calls context_switch( ) → TASK_RUNNING (running)
- task exits via do_exit() → TASK_ZOMBIE (task is terminated)
- task is preempted by higher priority task
- task sleeps on wait queue for a specific event → TASK_INTERRUPTIBLE or TASK_UNINTERRUPTIBLE (waiting)
- event occurs and task is woken up and placed back on the run queue


```
QF d:

        COMMUNICATIONS DIVISION DISPATCH QUERY SCREEN




▶QVR NAM/                    ▶QA ENO/
▶QTOM NAM/                   ▶DQ.NAM/                        .DOB/
▶QHAR NAM/                   ▶DQ.SSN/
▶QCHIS NAM/                  ▶DQ.OLN/              .OLS/
▶QE NAM/                     ▶QV.LIC/        .LIS/
▶QPT NAM/                    ▶RQ.LIC/        .LIS/
▶QPT ENO/                    ▶QVM VIN/
                            ▶QVM LIC/
```

ORACLE®

OPERATING SYSTEM/360 CHART V25-6156

| User Program 1 |
| Oerating System 1 |
| Virtual 370 |

| User Program n |
| Operating System n |
| Virtual 370 |

| VM370 |
| Hardware |

VIRTUAL MACHINE ARCHITECTURE (VM370)

```
                        Windows

A fatal exception 0E has occurred at 0028:C0011E36 in VXD VMM(01) +
00010E36. The current application will be terminated.

*    Press any key to terminate the current application.
*    Press CTRL+ALT+DEL again to restart your computer. You will
     lose any unsaved information in all applications.

                    Press any key to continue _
```
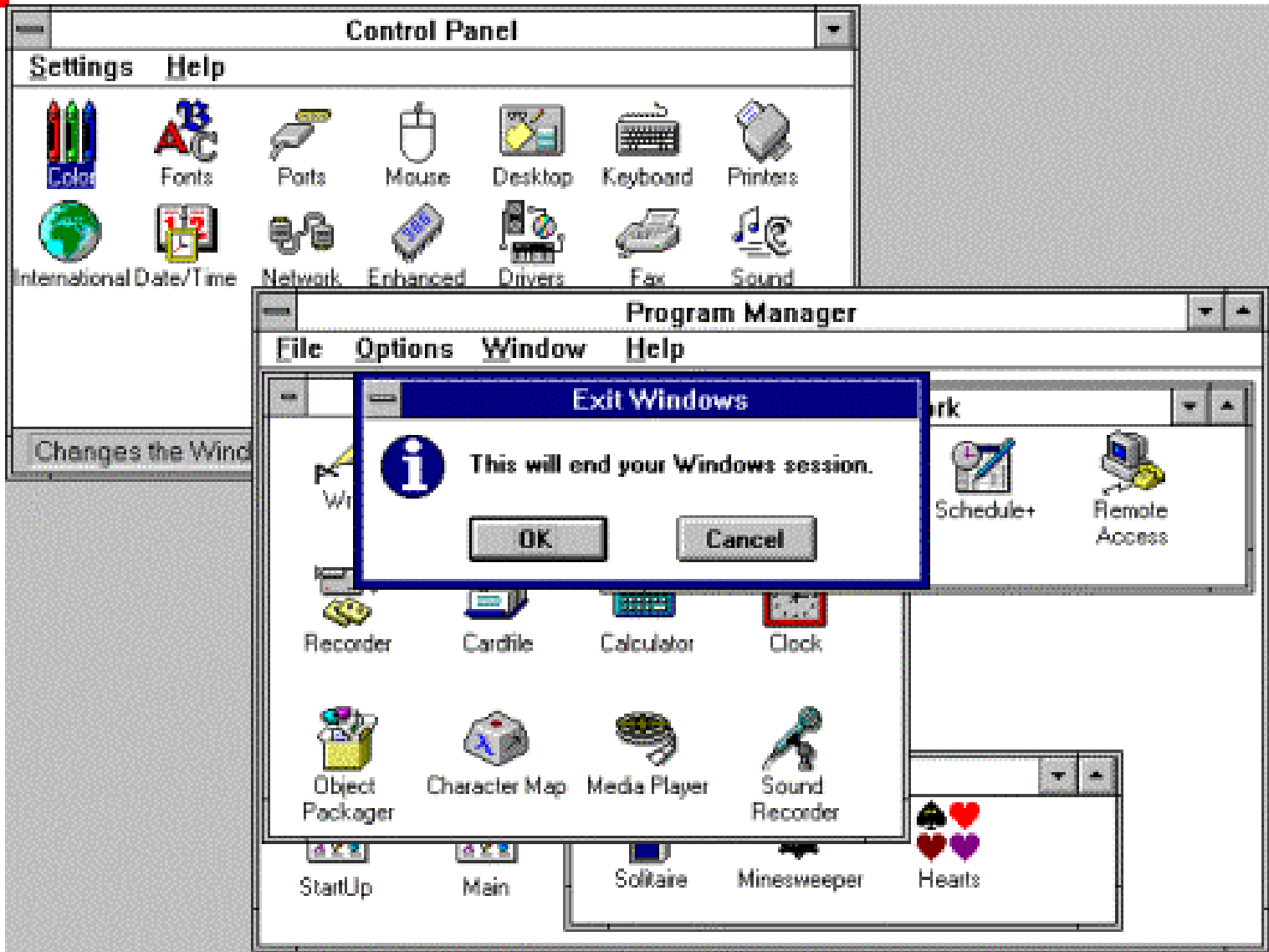
```
                         Windows

A fatal exception 0E has occurred at 0028:C0011E36 in UXD UMM(01) +
00010E36. The current application will be terminated.

*   Press any key to terminate the current application.
*   Press CTRL+ALT+DEL again to restart your computer. You will
    lose any unsaved information in all applications.

                Press any key to continue _
```
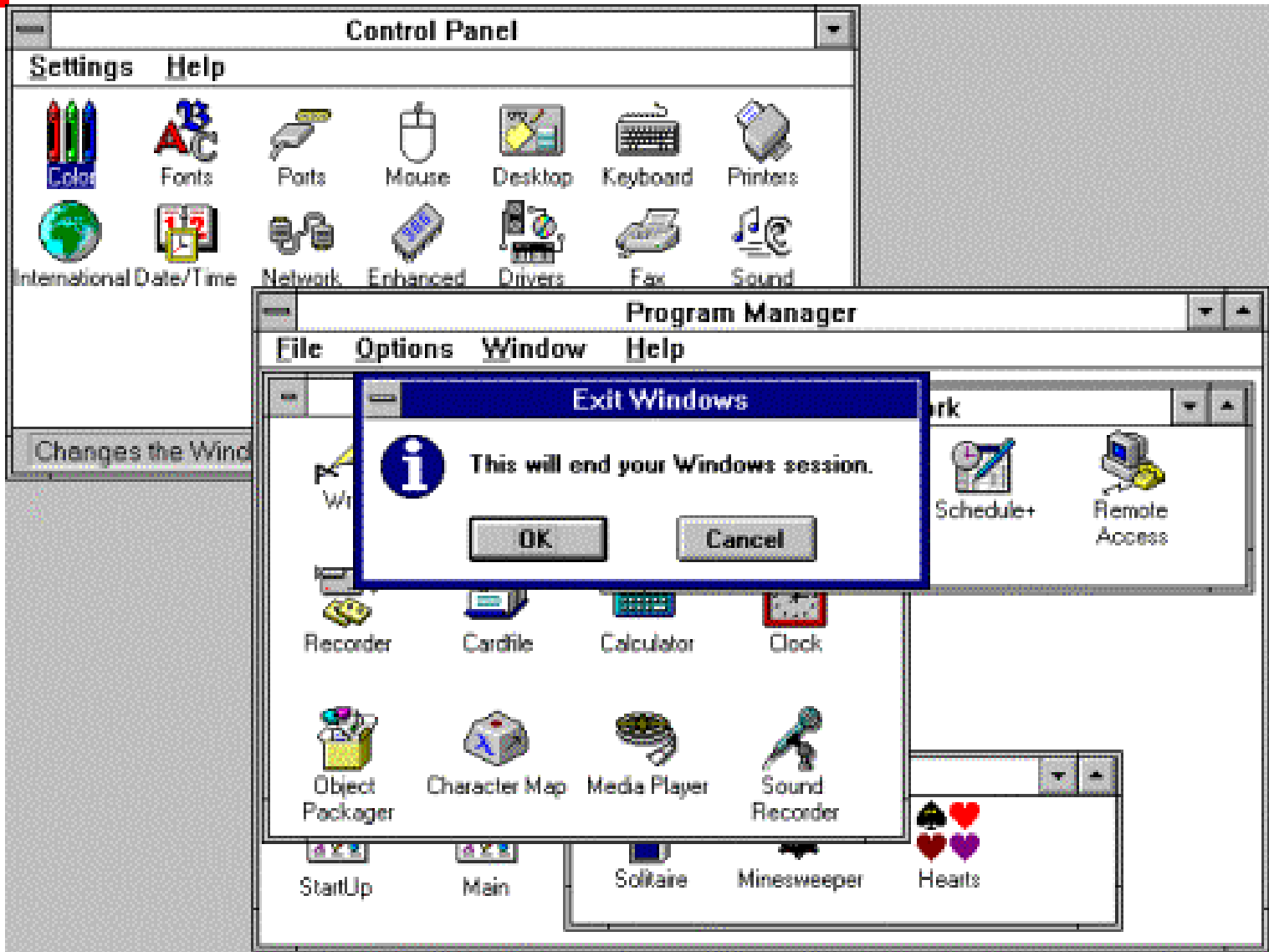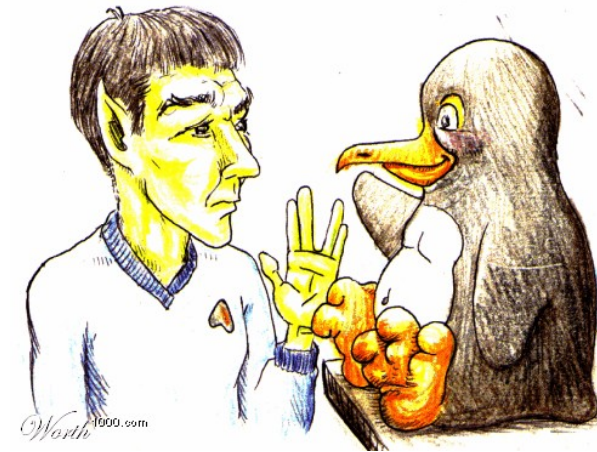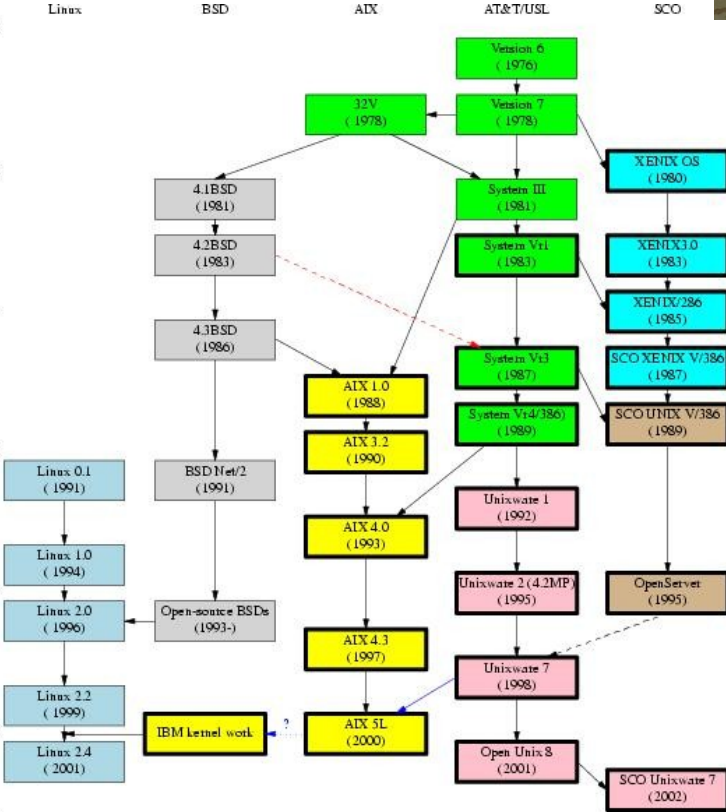
Linux | BSD | AIX | AT&T/USL | SCO

1975

Version 6
(1976)

32V
(1978) — Version 7
(1978)

1980

4.1BSD
(1981) — System III
(1981)

XENIX OS
(1980)

4.2BSD
(1983) — System Vr1
(1983)

XENIX 3.0
(1983)

1985

4.3BSD
(1986)

XENIX/286
(1985)

System Vr3
(1987) — SCO XENIX V/386
(1987)

AIX 1.0
(1988)

System Vr4/386
(1989) — SCO UNIX V/386
(1989)

AIX 3.2
(1990)

1990

Linux 0.1
(1991)

BSD Net/2
(1991)

Unixware 1
(1992)

AIX 4.0
(1993)

Linux 1.0
(1994)

1995

Linux 2.0
(1996)

Open-source BSDs
(1993-)

Unixware 2 (4.2MP)
(1995)

OpenServer
(1995)

AIX 4.3
(1997)

Linux 2.2
(1999)

Unixware 7
(1998)

2000

Linux 2.4
(2001)

IBM kernel work

AIX 5L
(2000)

Open Unix 8
(2001)

SCO Unixware 7
(2002)

2003

Distribution of OS patches and/or disk images to multiple servers at once

Network

Server Blades

Storage

Management Server and Console

Centralized system management

Compute elements—processors, drives, memory, and network connections—move to the blade

Blades plug into the chassis, sharing redundant power and cooling, as well as Ethernet switches

Typical "integrated" server

Infrastructure elements—power, management, networking, and cooling components—move to the chassis

ORACLE®

APPLICATIONS

OPERATING SYSTEM

VIRTUAL LAYER

Virtual Infrastructure
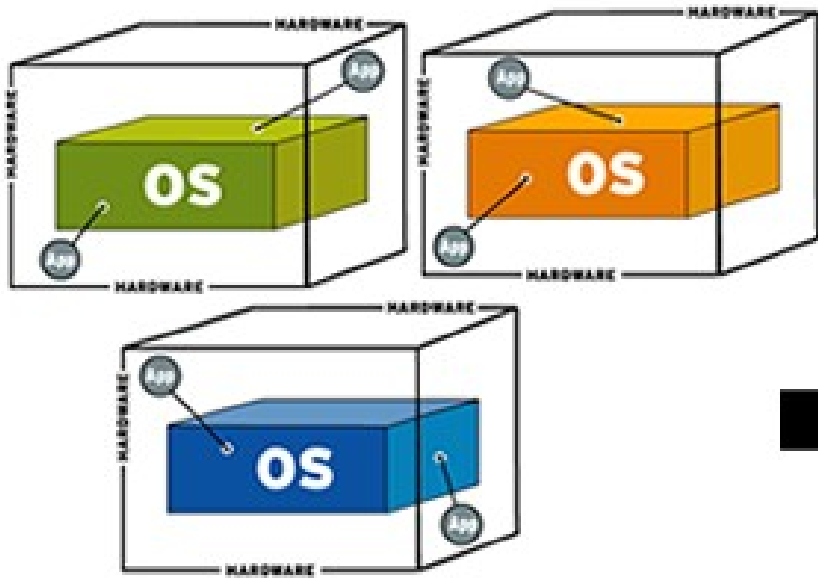
STORAGE
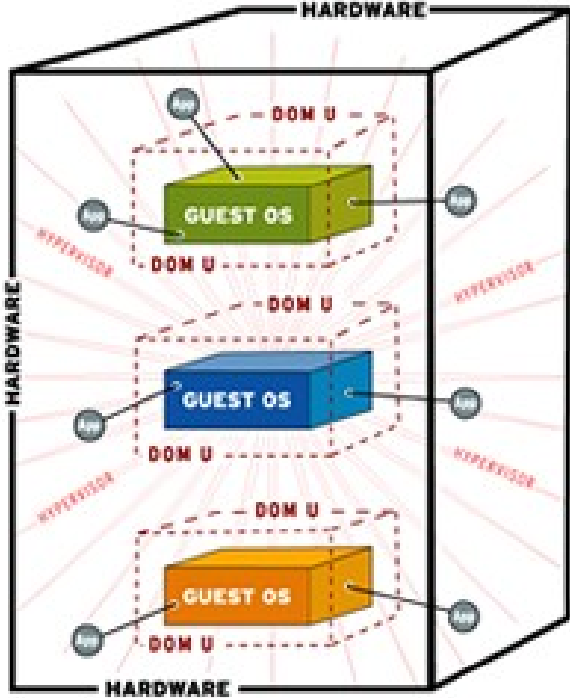
NETWORK
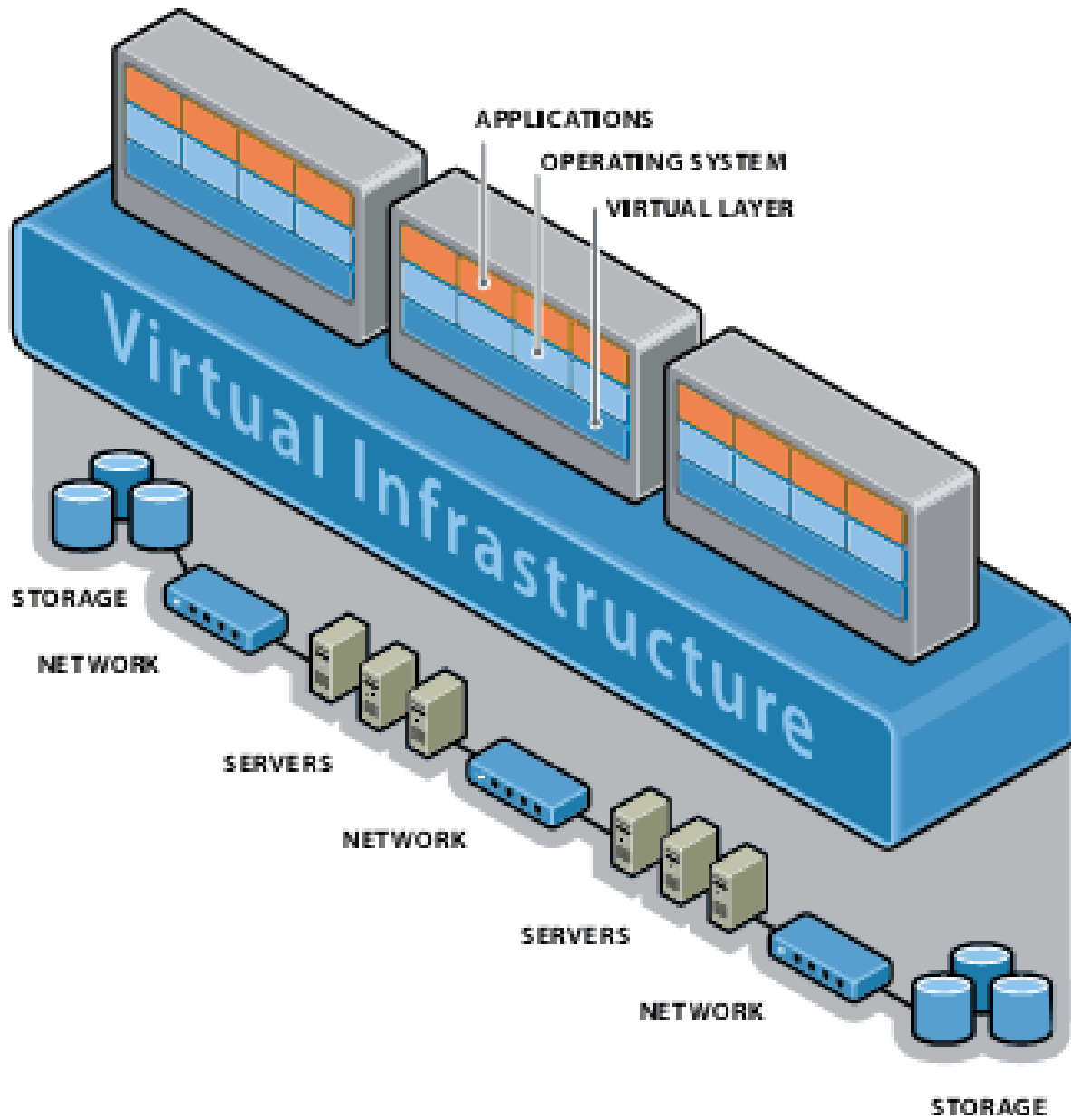
SERVERS

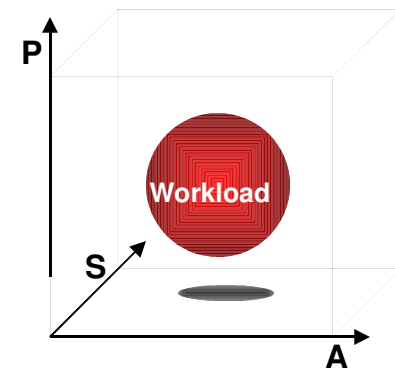NETWORK

SERVERS

NETWORK

STORAGE

ORACLE

Overview:
Virtualization Technology
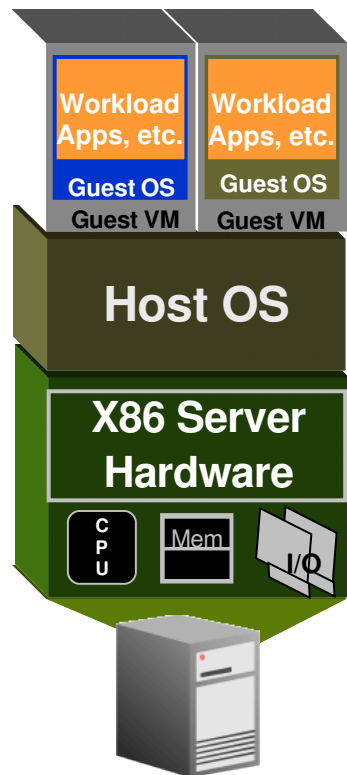
ORACLE®

Defining Virtualization

- Abstraction of underlying technology
  - Decouple workloads from resource boundaries
- Benefit:
  - Increased agility in matching resources to the changing needs of your workload:
    - Performance
    - Availability
    - Security
    - Stability

ORACLE

Server Virtualization Technologies
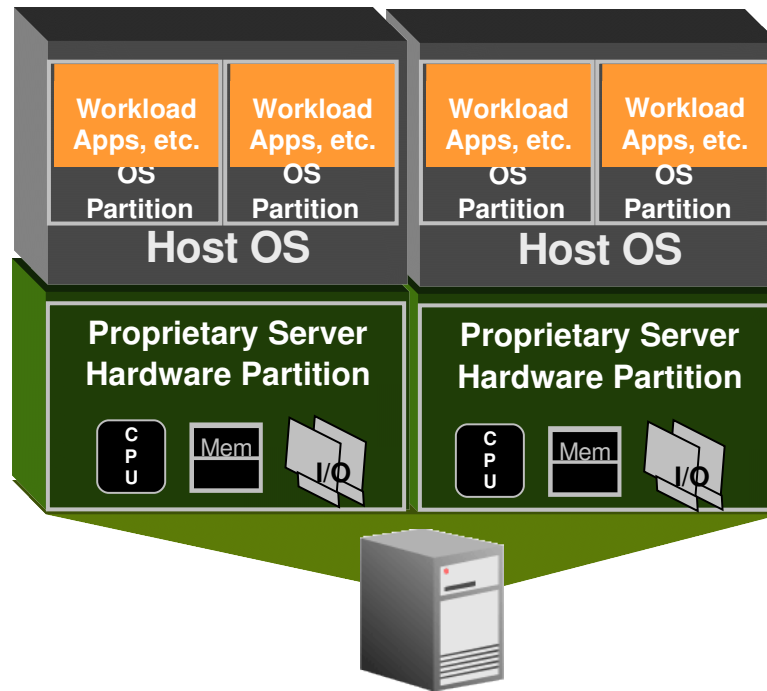
Host OS-based, e.g
- VMware Workstation
- Microsoft Virtual Server

Hardware Partitioning, e.g.
- Sun SPARC Domains
- IBM PPARs
- HP nPARs

OS Partitioning, e.g.
- Solaris Containers
- AIX dLPARs
- HP vPARs

Hypervisor-based, e.g.
- **Oracle VM**
- VMware ESX Server
- Citrix XenServer
- Windows Hyper-V

Workload Apps, etc.
Guest OS
Guest VM

Workload Apps, etc.
Guest OS
Guest VM

Host OS

X86 Server Hardware

CPU | Mem | I/O

Workload Apps, etc.
OS Partition

Workload Apps, etc.
OS Partition

Host OS

Proprietary Server Hardware Partition

CPU | Mem | I/O

Workload Apps, etc.
OS Partition

Workload Apps, etc.
OS Partition

Host OS

Proprietary Server Hardware Partition

CPU | Mem | I/O

Workload Apps, etc.
Guest OS
Guest VM

Workload Apps, etc.
Guest OS
Guest VM

Hypervisor

X86/64 Bare-Metal Server

CPU | Mem | I/O | I/O | Mem | CPU

- **Primarily desktop**
- **Easy to use**
- **Very slow (2 OSes)**

- **Excellent isolation**
- **Expensive, proprietary hw**
- **Coarse grain resources**
- **Mix OSes / versions**

- **Only moderate isolation**
- **Potentially good scalability**
- **Fine-grained resources**
- **Cannot mix OS/patch levels**

- **Excellent isolation**
- **Affordable, multi-source hw**
- **Fine-grained resources**
- **Mix OSes / versions**

ORACLE

"Hypervisor" Vs. "Virtualization Server"?

- Terms commonly used interchangeably but actually technically distinct:

  - A *hypervisor* is a very low-level, very efficient platform code layer that allows multiple operating systems to run on a host computer at the same time
    - Fundamental abstraction of physical *compute* resources
      - *I/O resources* may- (VMware) or may not (Xen) be handled by the hypervisor itself
    - Allows resource (re)allocation between VMs without disruption

  - A *virtualization server* includes a hypervisor function but also includes additional functionality, e.g.
    - I/O resource/traffic management (Xen)
    - Server configuration, access, and security management, etc.

ORACLE®

# Hardware Support for Virtualization (HVM) Where does it fit-in?

- Goal: efficiently insure VM OSes are not "hurting" each other
  - E.g. Trying to unsafely modify common resources/state, etc.

Techniques:	Purpose:	Benefits:	Comments:
Paravirtualization (PV) e.g. Oracle VM & Xen	Modify the OS and/or drivers so they knows how to behave in a virtual environment	• Good-to-excellent performance vs. bare-metal	• PV OS kernel req'd (but rapidly not an issue)
Emulation / **translation**, e.g. VMware ESX	Design the virtualization server to intercept or "trap" harmful requests and/or translate requests into appropriate forms	• Use unmodified OS • HVM hardware not req'd (but rapidly not an issue)	• Poor scalability esp. under I/O load
Hardware virtualization (HVM); e.g. Oracle VM & Xen	Design the hardware so it knows how to handle inappropriate requests itself. (Note: PV drivers can be used with an otherwise unmodified OS, e.g. Windows, to improve performance)	• Use unmodified OS	• Poor perf. today • HVM hardware req'd (but rapidly not an issue)

Anatomy of a Virtualization Server
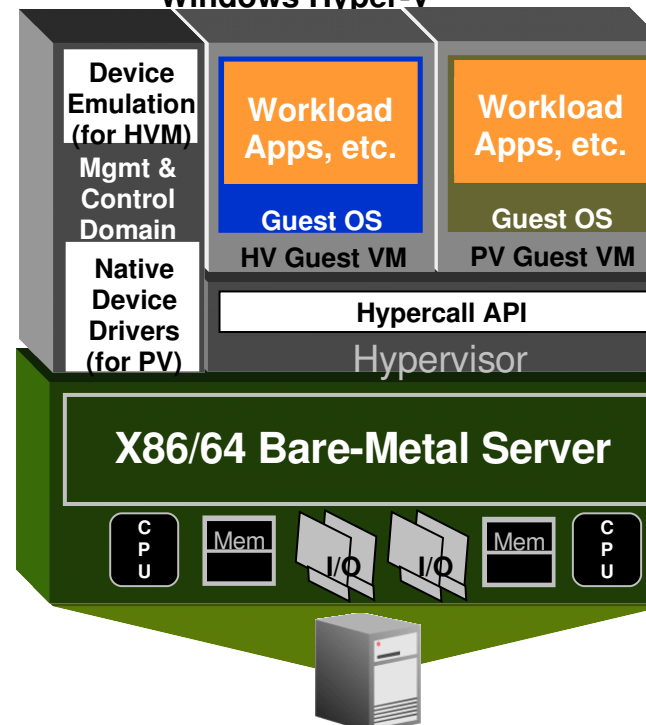
Emulation-based, e.g.

- VMware ESX Server

Paravirtualization (PV)-based, e.g.

- **Oracle VM**
- Citrix XenServer
- Windows Hyper-V
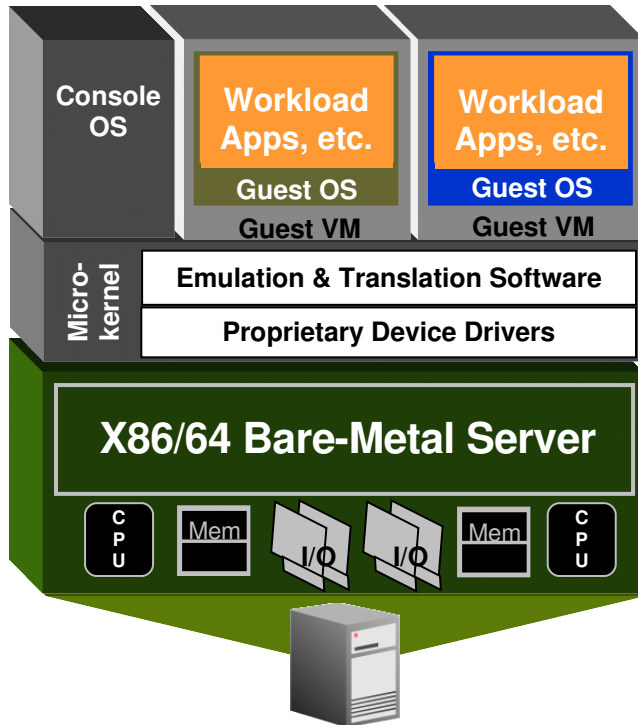
- **Machine images run unmodified: broad compatibility**
- **Does not leverage or require HV hardware**
- **Poor I/O scalability due to emulation architecture**
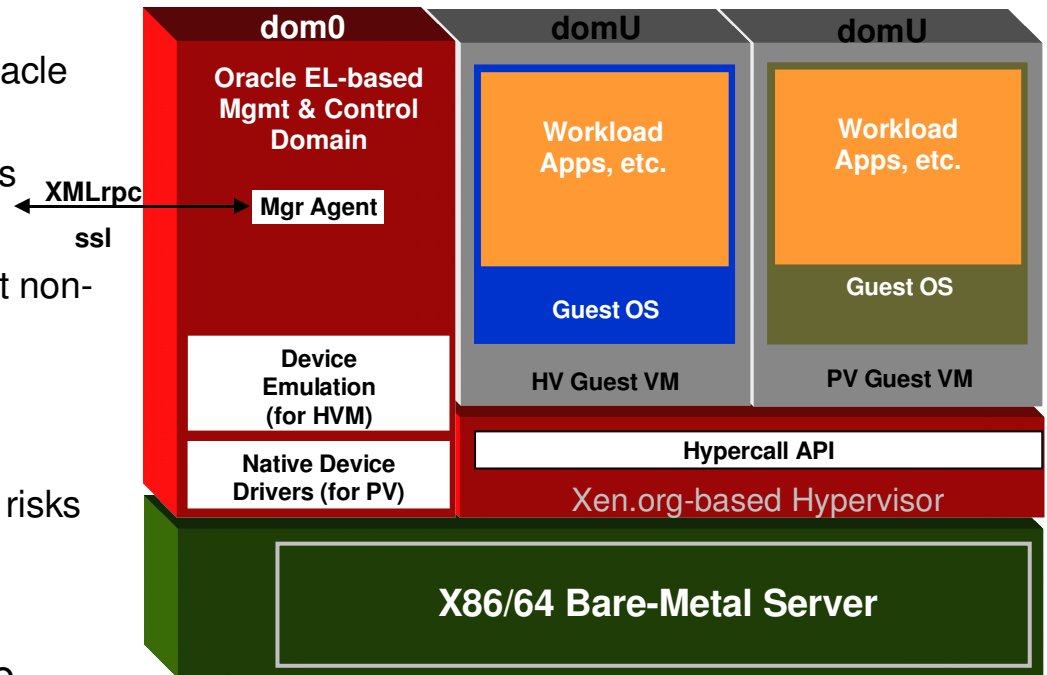- **Not open: dependent on virt. vendor for drivers**

- **Excellent scalability, esp. I/O with PV guests**
- **Requires PV OS kernel for best perf.**
- **Requires HV hardware for unmodified images**
- **Open: wide device support: uses native (e.g. Linux) device drivers**

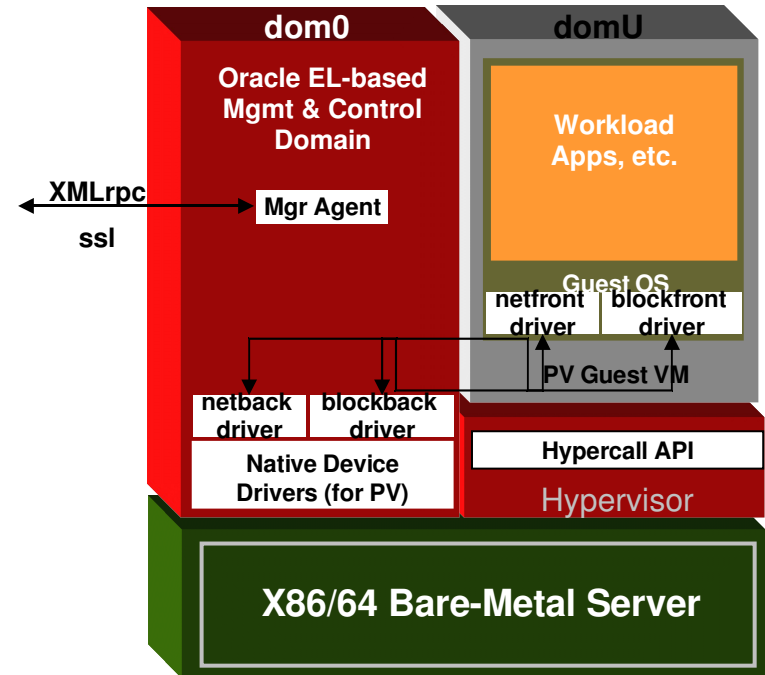# Anatomy of an Oracle VM Server
Key Concepts: Dom0 & DomU

- **"Dom0"** is a privileged management and control domain typically containing…
  - A thin control kernel based on Oracle Enterprise Linux
  - Open / native Linux device drivers
  - Oracle VM Manager agent
  - Device emulation code to support non-PV guests (e.g. Windows)
- Dom0 should not contain 'user' applications as a best practice
  - Minimize performance & security risks
  - Minimize code size
- **"DomU"** is an unprivileged/user domain that is a guest VM on the server
  - Run any normal server workload
  - One domU is not aware of another

dom0

Oracle EL-based Mgmt & Control Domain

XMLrpc

ssl

Mgr Agent

Device Emulation (for HVM)

Native Device Drivers (for PV)

domU

Workload Apps, etc.

Guest OS

HV Guest VM

domU

Workload Apps, etc.

Guest OS

PV Guest VM

Hypercall API

Xen.org-based Hypervisor

X86/64 Bare-Metal Server

Anatomy of an Oracle VM Server
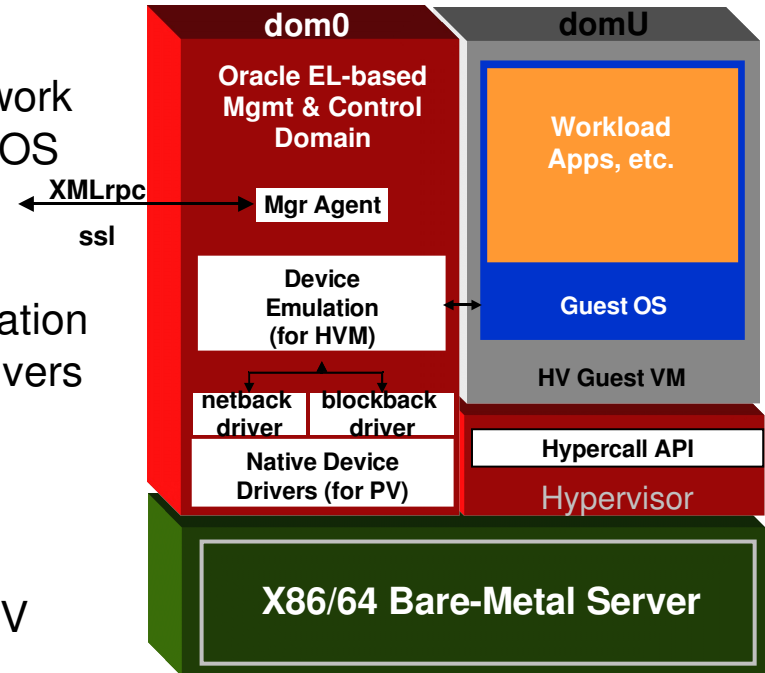## Key Concepts: Driver Architecture: PV

- For PV guests, driver abstraction:
  - Replaces hardware-specific drivers
    - One network driver
    - One block driver
  - Very stable / rarely changes
    - Excellent guest stability
  - **Front-end drivers (net & block)**
    - Inside the VM / domU OS
  - **Back-end drivers (net & block)**
    - In dom0 / shared
  - **Open, native hardware vendor drivers**
    - Uses open Linux drivers
- Virtualization server manages high-performance communication front <-> back
  - Can leverage dom0 kernel security features, e.g. packet sniffing, firewalling, & rate control, etc.

dom0

domU

Oracle EL-based Mgmt & Control Domain

Workload Apps, etc.

XMLrpc

ssl

Mgr Agent

Guest OS

netfront driver

blockfront driver

PV Guest VM

netback driver

blockback driver

Native Device Drivers (for PV)

Hypercall API

Hypervisor

X86/64 Bare-Metal Server

Anatomy of an Oracle VM Server
Key Concepts: Driver Architecture: HVM

- HVM guests driver choices:
  - **Unmodified native driver(s)**
    - OS typically installs basic native network and block drivers that come with the OS (not with the virtualization server)
    - Device support provided via device emulation / translation in the virtualization server on top of net- & block-back drivers in dom0
    - Slower than PV due to overhead
  - **PV drivers** (front- & back)
    - HVM (unmodified) kernels can use PV drivers
    - Leverages PV driver stack same as previously described
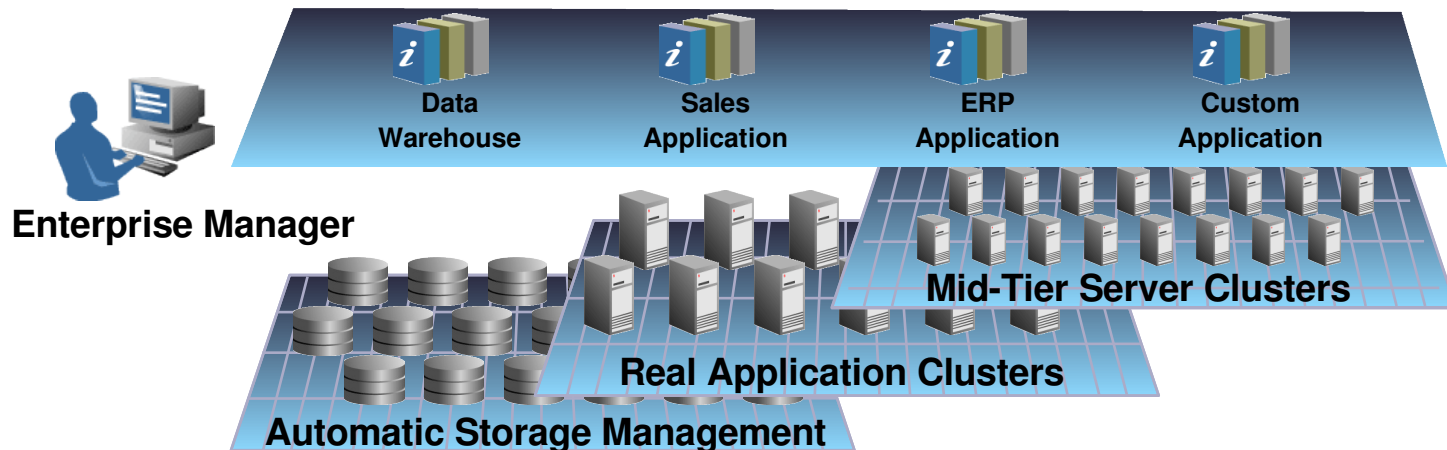   - Excellent performance

dom0 — domU

Oracle EL-based Mgmt & Control Domain

Workload Apps, etc.

XMLrpc
ssl

Mgr Agent

Device Emulation (for HVM)

Guest OS

netback driver — blockback driver

HV Guest VM

Native Device Drivers (for PV)

Hypercall API

Hypervisor

X86/64 Bare-Metal Server

ORACLE

Oracle's Virtualization Platform:
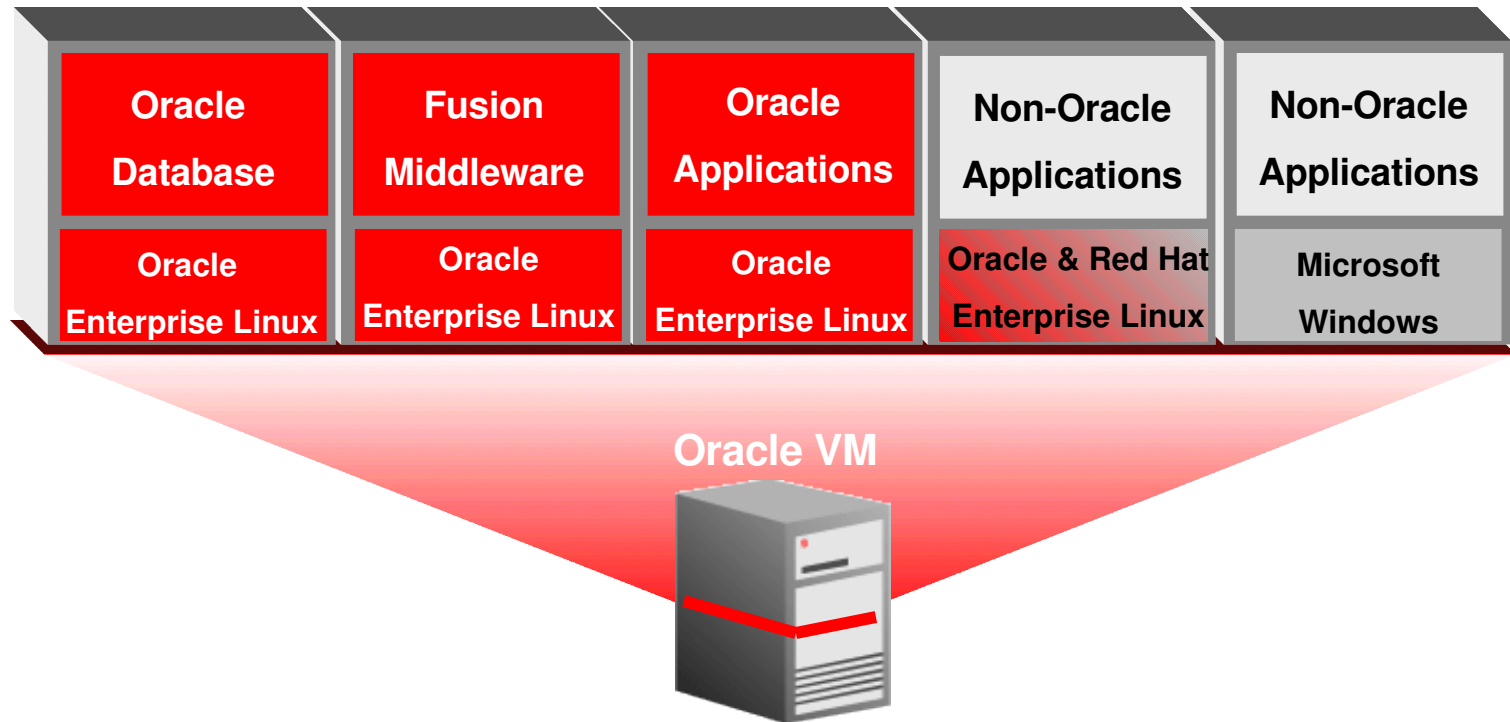# Oracle VM

Oracle Grid-Based Computing
A History in Virtualization

- Grid-Based Storage
  - Oracle Automatic Storage Management
- Grid-Based Server Clusters
  - Oracle Database
  - Oracle Middleware
- Grid-Based Management
  - Oracle Enterprise Manager Grid Control

Enterprise Manager

Data Warehouse   **Sales Application**   **ERP Application**   **Custom Application**

Mid-Tier Server Clusters

Real Application Clusters

Automatic Storage Management

ORACLE®

Oracle VM

| Oracle Database	Fusion Middleware	Oracle Applications	Non-Oracle Applications	Non-Oracle Applications
Oracle Enterprise Linux	Oracle Enterprise Linux	Oracle Enterprise Linux	Oracle & Red Hat Enterprise Linux	Microsoft Windows

Oracle VM

- Oracle tested and supported server virtualization
- Maximizes consolidation of Linux and Windows servers
- Saves on power, cooling and space
- Improves operational agility through the lifecycle

ORACLE®

Oracle VM

- Server virtualization software and support
  - Enterprise-class hypervisor
  - Free product download; low-cost support

- Runs both Linux and Windows Guests
  - Paravirtualized and hardware virtualized guests
    - HV on x86 hardware with HVM support
  - 64-bit and 32-bit guests
  - Up to 64-way SMP hardware
  - Up to 32 virtual processors per guest
  - Includes Live Migration at no additional cost
  - Integrated, browser-based management console
  - Free downloadable VM images

ORACLE®
VM

- Enterprise-quality support worldwide

ORACLE®

Oracle VM
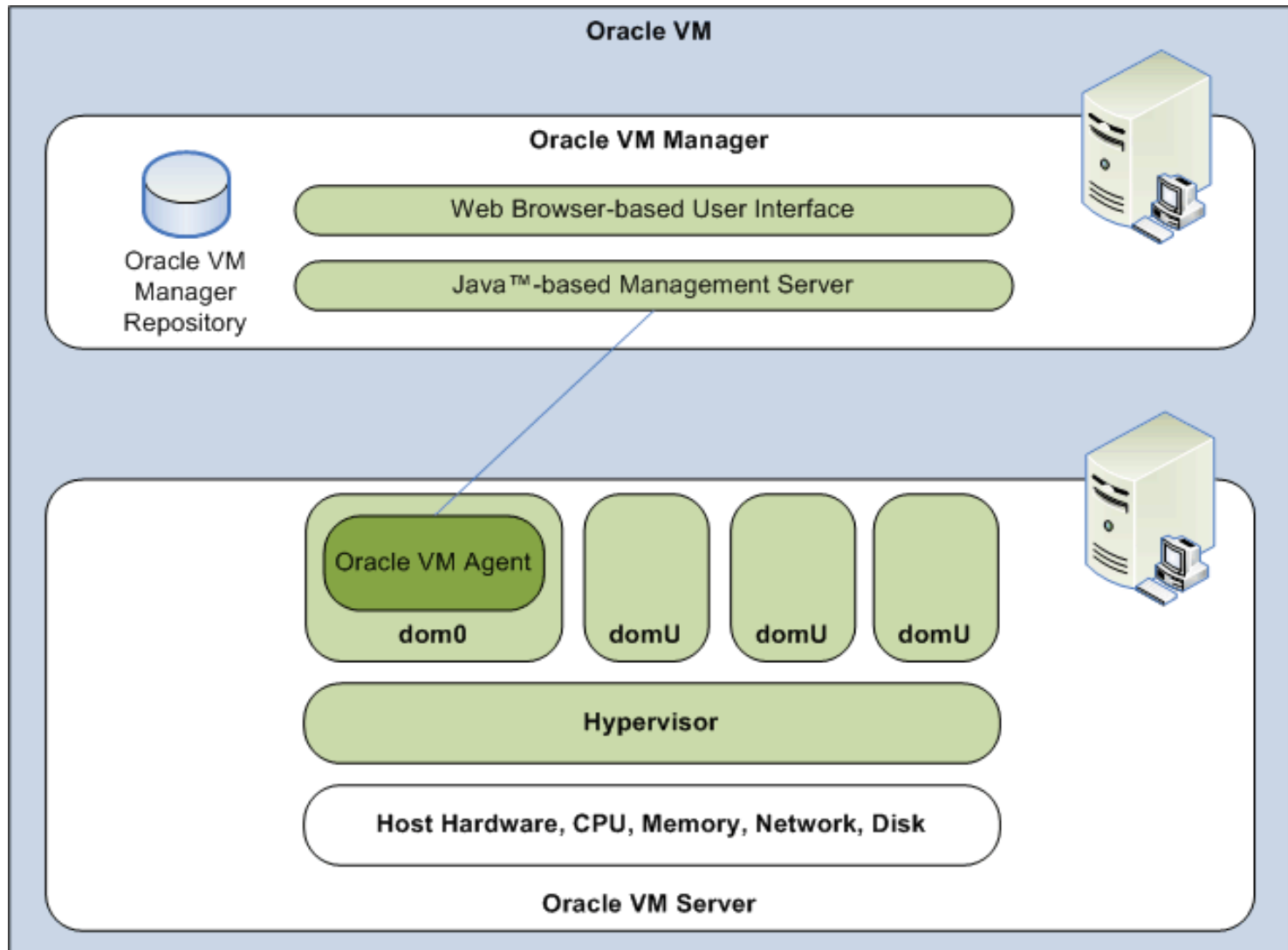
Oracle VM Features

- Simplified installation
  - Single CD server installation
  - Installs and configures in a minute
- Faster stack deployment
  - Pre-configured Virtual Machine images of Oracle Database and Oracle Enterprise Linux
- Live VM migration included
  - Better uptime, agility, and HA
- Linux and Windows guest support
  - Oracle Enterprise Linux 4 and 5;
  - RHEL3, RHEL4 and RHEL5
  - On HV capable hardware: Windows 2000, Windows 2003, Windows Server 2003 (32- and 64bit) and Windows XP

Oracle VM Manager

- Browser-based management solution
- Included with Oracle VM
- Full VM lifecycle management:
  - Create
  - Configure
  - Clone
  - Share
  - Boot
  - Migrate

Oracle Product Certification with Oracle VM

- Oracle Database
- Oracle Application Server
- Oracle Enterprise Manager
- Oracle Berkeley DB
- Oracle TimesTen
- Oracle E-Business Suite
- Oracle PeopleSoft
- Oracle Siebel
- Oracle Hyperion
- More information on Metalink Note 464754.1

ORACLE®

VM

ORACLE®

File   Edit   View   Favorites   Tools   Help

Back   •   ✖   🔄   🏠   Search   ⭐ Favorites   Links   🔗 Custoi   Address   🔗 http://linux.oracle.com/pls/apex/f?p=102:2:5419553457747:   Go

Oracle VM 2.1

Oracle Validated Configurations

For

Oracle VM

Validated Configuration Details

Publication Date	26-DEC-07
Version	1
Server Platform	HP ServerBlade BL480c
Storage Model	SmartArray P400 StorageBlade 40c,StorageBlade, SB600c
Guest VMs	Oracle Database Single Instance on OEL 4 Update 5 x86_64 , OEL 5 x86_64

Guest VM Configurations

Oracle recommends Paravirtualized guests

Oracle VM Server 2.1 , Guest VM - Oracle Enterprise Linux 4 Update 5 x86_64 ,Hardware Virtualized with Oracle 10.2.0.3

Oracle VM Server 2.1 , Guest VM - Oracle Enterprise Linux 5 x86_64 ,Hardware Virtualized with Oracle 11.1.0.6

Oracle VM Server 2.1 , Guest VM - Oracle Enterprise Linux 4 Update 5 x86_64 ,Hardware Virtualized with Oracle 11.1.0.6

Oracle VM Server 2.1 , Guest VM - Oracle Enterprise Linux 5 x86_64 ,Hardware Virtualized with Oracle 10.2.0.3

Oracle VM Server 2.1 , Guest VM - Oracle Enterprise Linux 5 x86_64 ,Paravirtualized with Oracle 10.2.0.3

Oracle VM Server 2.1 , Guest VM - Oracle Enterprise Linux 5 x86_64 ,Paravirtualized with Oracle 11.1.0.6

Oracle VM Server 2.1 , Guest VM - Oracle Enterprise Linux 4 Update 5 x86_64 ,Paravirtualized with Oracle 11.1.0.6

Oracle VM Server 2.1 , Guest VM - Oracle Enterprise Linux 4 Update 5 x86_64 ,Paravirtualized with Oracle 10.2.0.3

Server and Storage Platform Details

Oracle & Alliance Partner Collaboration:

Making Linux & Virtualization Easy to Deploy

Server Model	2 X HP ServerBlade BL480c
Processors	2 X Quad Xeon
Hardware Virtualization Support	Yes

🔗   🖳 Local intranet
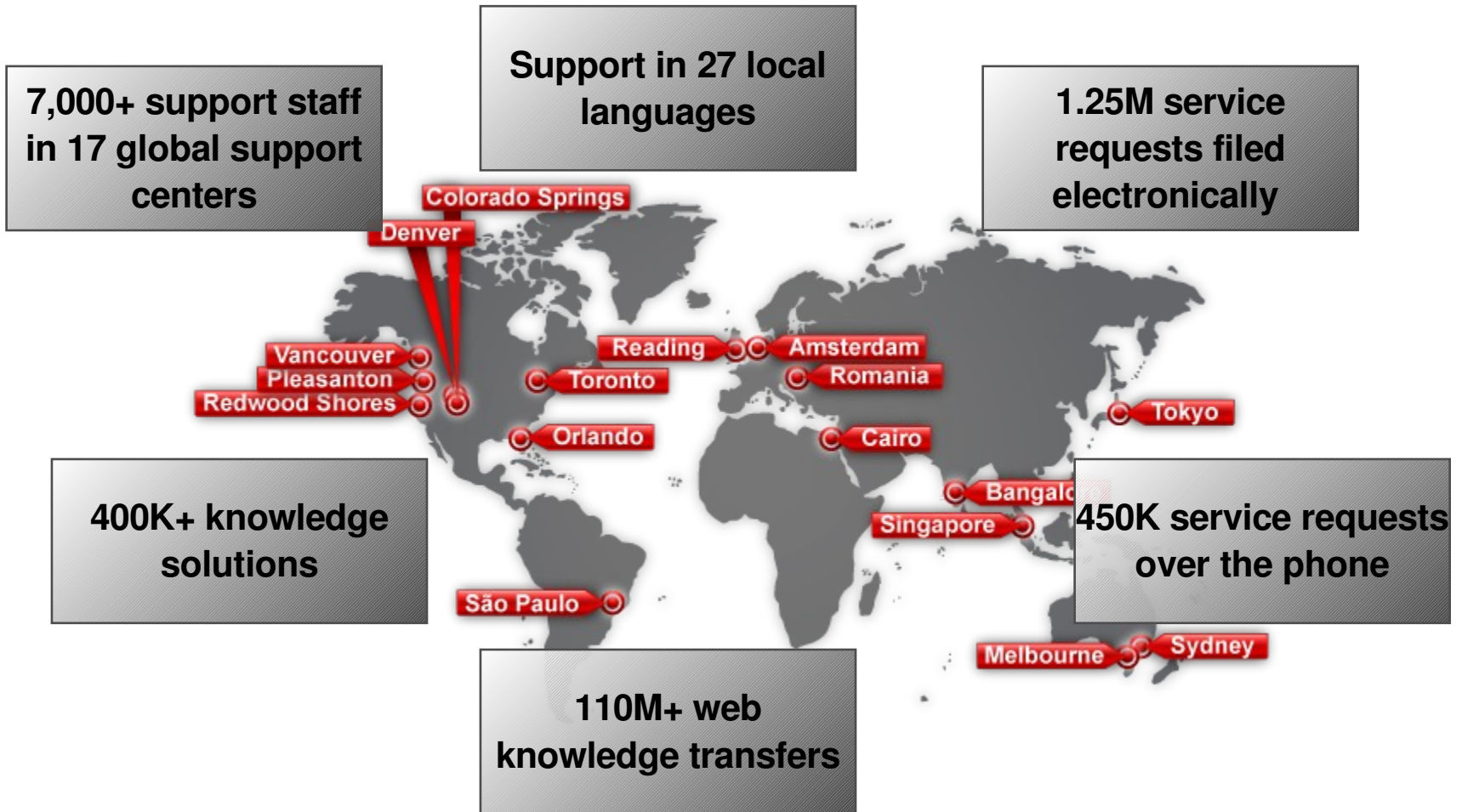
Partner Support

Oracle's Contributions

- Enhanced and optimized Xen technology
  - I/O overhead
  - Memory overhead
  - Process Scheduling
- Community involvement
  - Dedicated Xen development team
  - Code and bug fix contributions to Xen community
  - Members of Xen community at Oracle
  - Member of XenSummit committee
- Significant testing
  - Real-world testing with Oracle On Demand workloads
  - Testing with Oracle Validated configuration workloads

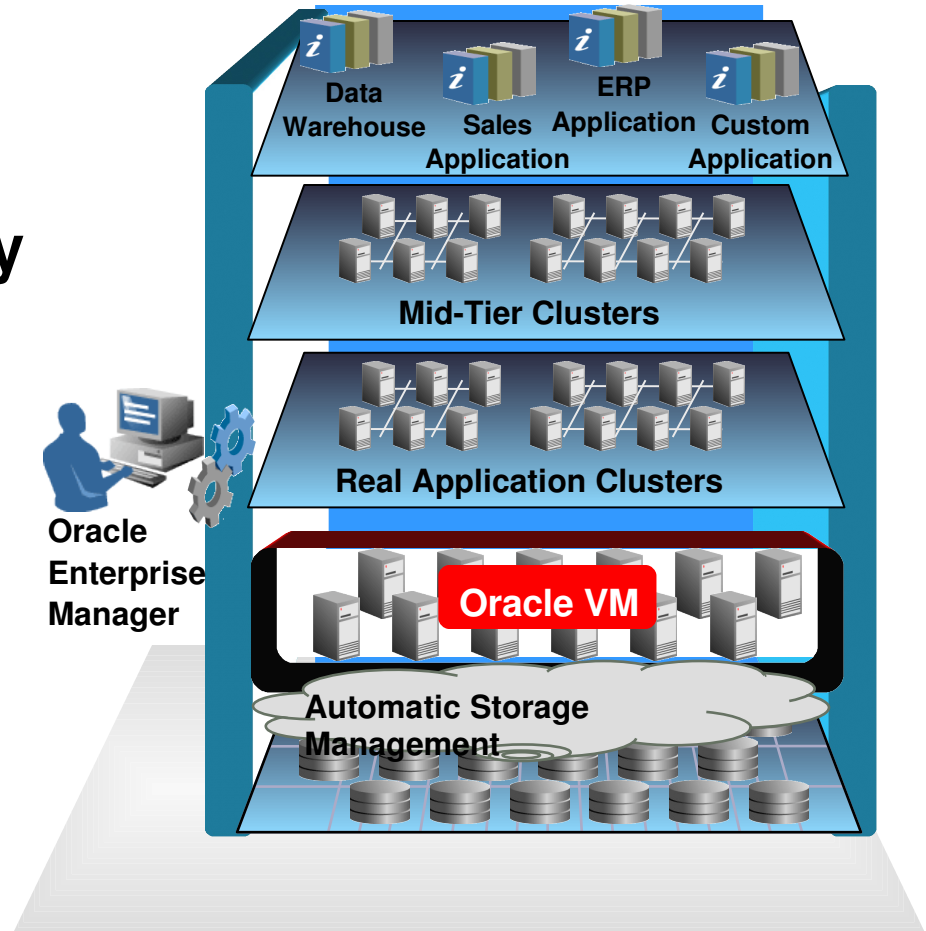ORACLE®

Performance Testing

- Extensive real-world testing
  - LMbench: microbenchmarking
    - Context switching, Null calls, TCP Selects, etc
  - Swingbench: DB workload benchmarking
    - Typical OLTP workloads
    - Varying SGA size, # of users, # of vCPUs
  - Paravirtualized domains with Enterprise Linux 4, 64bit
- Results:
  - On average **three times less overhead** compared server virtualization products from other vendors.
  - Virtual SMP scalability of Oracle VM is at 90+%
- Small workloads are at same performance as real hardware

ORACLE®

One Support Call for the Complete Stack

7,000+ support staff in 17 global support centers

Support in 27 local languages

1.25M service requests filed electronically

400K+ knowledge solutions

450K service requests over the phone

110M+ web knowledge transfers

Colorado Springs

Denver

Vancouver

Pleasanton

Redwood Shores

Toronto

Reading

Amsterdam

Romania

Orlando

Cairo

Tokyo

Bangalore

Singapore

São Paulo

Melbourne

Sydney

ORACLE®

Summary

- **Oracle Uniquely Combines**
  - Proven Grid capabilities
  - Server virtualization

- **Full Virtualization Strategy**
  - Applications
  - Data
  - Servers
  - Storage

- **Real-World Benefits**
  - Maximized consolidation
  - Optimized agility
  - High availability
  - Affordable enterprise-class support

ORACLE

Resources

- **Virtualization Center on OTN**
  otn.oracle.com/goto/virtualization
  - wiki, downloads, discussion forum, howtos

- **Oracle VM Home Page**
  oracle.com/virtualization

- **Free Download: Oracle VM**
  edelivery.oracle.com/oraclevm