Red Hat Enterprise Linux 5 Technology Overview

Development Summary

- Kernel base: 2.6.18
- Core deliverable: virtualization
- Additional development areas:
  - Network storage (Autofs, CacheFS / NFS persistent local cache, iSCSI)
  - Desktop (GNOME, X.Org 7.1, Laptop)
  - Stateless Linux foundation (Desktop / Server / Virtualized)
  - New Driver Model (better support for 3rd party drivers)
  - Development Tools (SystemTap, Frysk)
  - GFS2 (Single Node GFS / Clustering)
  - Single Sign On and Security
  - Integrated directory and security integration
  - RHN support for virtualization
  - Yum/Pup based RHN updater
- Broad range of new HW support
- Better USB support
- Large SMP support
- IPv6 support and conformance enhancements
- IPSEC enhancements
- SELinux and auditing enhancements
- Kexec/Kdump improved serviceability (replace diskdump/netdump)
- Performance improvements through finer grained locking
- Multi-Core beyond Dual
- I/O-AT – Intel's network accelerators
- Improved ACPI support, suspend to disk
- Installer improvements

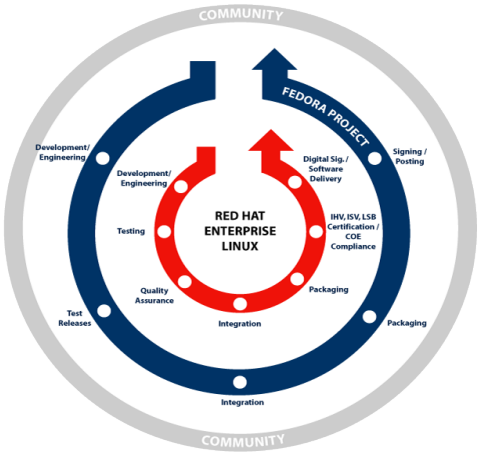
Red Hat Enterprise Linux 5: Feature Summary

- Red Hat Enterprise Linux 5 comprises more than 1200 components
- Over two years of development
- Technology created by:
  - Red Hat
  - Partners
  - Community

Virtualization
Server virtualization is provided in the base server product & is available for the client product
Storage and extended server virtualization is provided with the Virtualization Platform option
Red Hat Network support for virtualization
Virt-Manager, libvirt/virsh management tools
Packaging
Replacement of previous AS, ES and WS products with a single server and a single client
New Options provide additional server and client product capabilities
Kernel & Performance
Red Hat Enterprise Linux is based on the Linux 2.6.18 kernel
Support for multi-core processors
Broad range of new hardware support
Updated crash dump capability provided by Kexec/Kdump
Support for Intel network accelerator technology (IOAT)
Numerous enhancements for large SMP systems
Enhanced pipe buffering
IPv4/IPv6 fragmentation offload & buffer management
Dynamically switchable per-queue I/O schedulers
Kernel buffer splice capability for improved I/O buffer operations
Security
SELinux enhancements include Multi-Level Security and targeted policies for all services
SEtroubleshooter GUI simplifies SELinux management
Integrated directory & security capabilities
IPSEC enhancements improve security and performance
ExecShield enhancements, such as a call frame Canary word, strengthen hacker defenses
New Audit features provide powerful new search/reporting and realtime monitoring

Red Hat Enterprise Linux 5: Feature Summary

- Features exposed to extensive testing with Fedora Core 4/5/6
  - Ensures high quality



- Application interfaces held stable for life of product

Networking & Interoperability
Network storage enhancements include Autofs, FS-Cache and iSCSI support
IPv6 support and conformance enhancements
Improved Microsoft file/print and Active Directory integration
Desktop
Desktop enhancements provide updated configuration tools, applications and laptop support
Foundational Stateless Linux features (X autoconfigure, NetworkManager, etc)
Improved ACPI and laptop support
Smart card login - with PKI/Kerberos authentication
Integrated multi-media support
Enhanced plug and play hardware support (cameras, printers, scanners, etc)
Network Manager provides automatic wired and wireless network configuration
Enhanced graphics using AIGLX/Compiz (with fading, transparency, etc)
Development Environment
Enhanced application development tools including SystemTap profiler and Frysk debugger
GCC 4.1 and glibc 2.4 toolchain
Storage
Support for root device multipath IO (MPIO) improves availability
Single system/guest version of Red Hat Global File System included in the base product
Block device data encryption support
Management
Numerous installer improvements make system configuration simpler
Yum/Pup-based updater for Red Hat Network

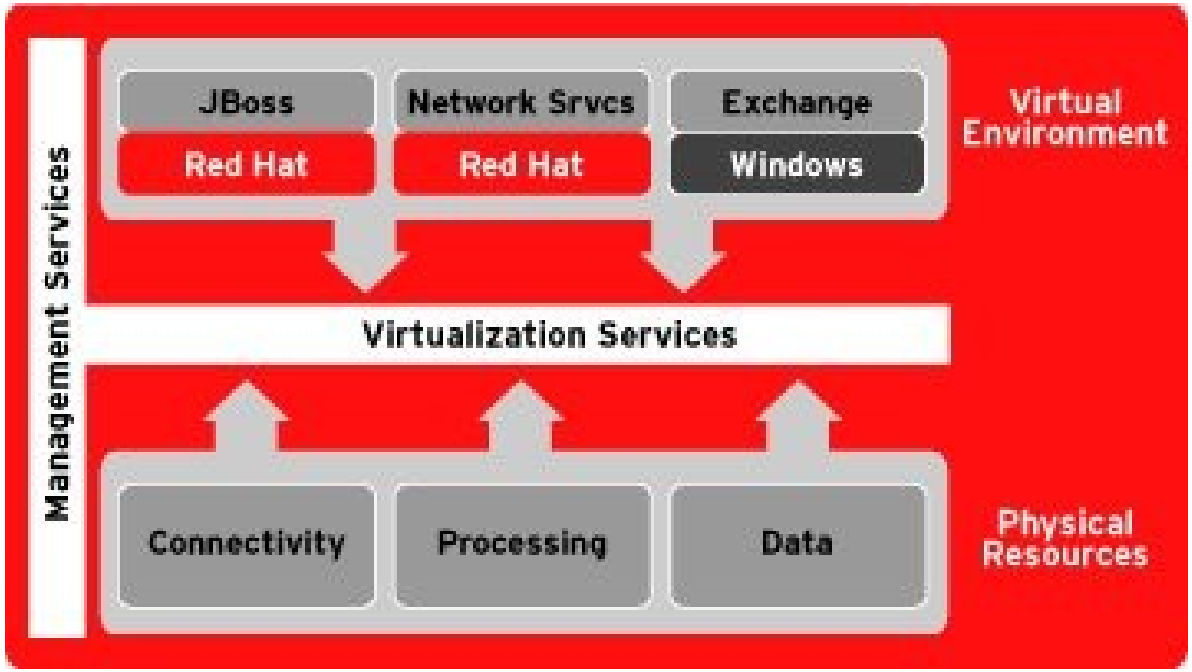
Virtualization

Virtualization Use Cases

- Datacenter Consolidation
  - Primary driver: increase utilization (primarily CPU, memory)
  - Thin host, workload in guests
  - Special case: security isolation (E.g. hosting providers)
  - Server-oriented
- Development and Testing
  - Driver: multiple development and testing environments, isolation from main workspace in host. Fault injection.
  - Typically main workload in host, testing in guests
  - Client and Server oriented
- Hardware Abstraction
  - Driver: new hardware with older guest OS
  - Support via virtual device drivers
  - Typically only few guests (maybe only 1)
  - Server-oriented

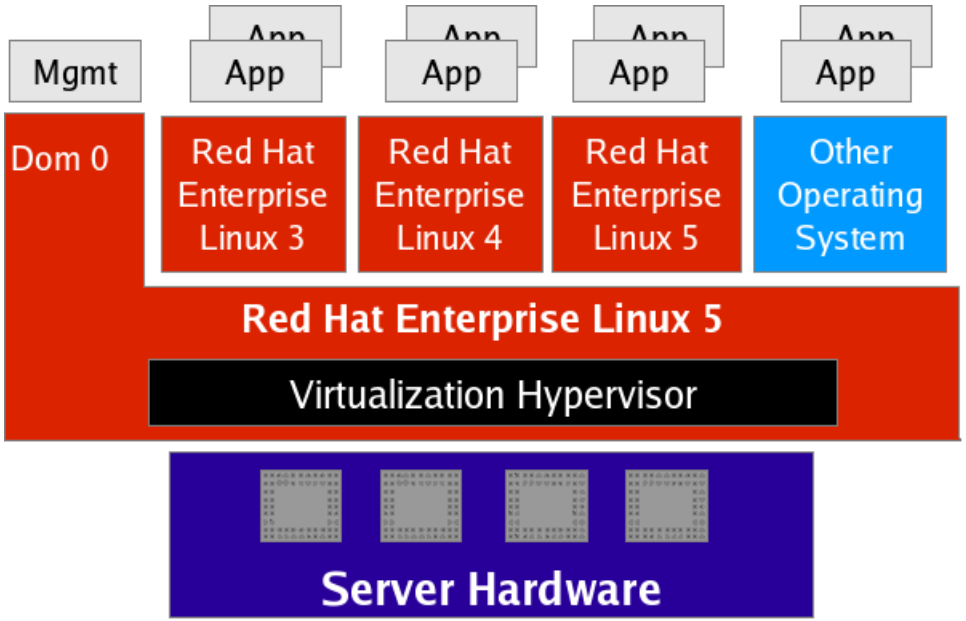
Red Hat virtualization architecture

- Physical computing and storage components provide a central pool of resources
- Virtual systems with appropriate compute, memory, and data can be dynamically allocated, provisioned, and managed



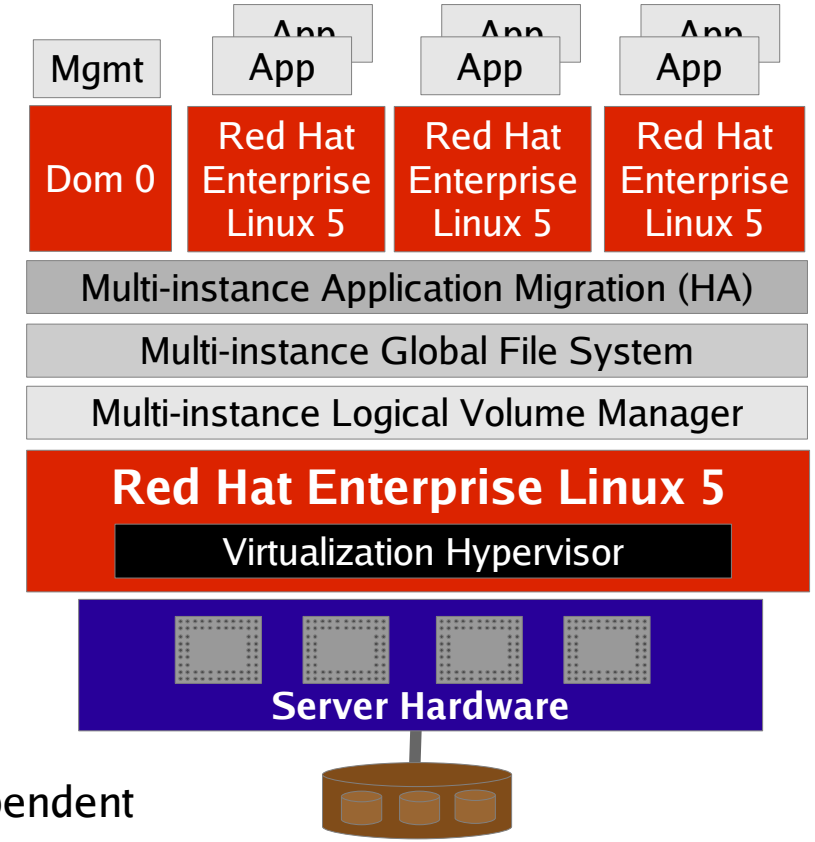
Full (hardware assisted) & Para-Virtualization

- Red Hat Enterprise Linux 5 will support hardware and software virtualization scenarios:
  - Fully virtualized on Intel VT & AMD-V systems
    - Allows guest to be Red Hat Enterprise Linux 3, 4 as well as other operating systems
  - Para-virtualized all systems
    - Red Hat Enterprise Linux 5.0
    - Red Hat Enterprise Linux 4.5
  - Support for x86, x86_64 at product release
    - Support for IA64 at 5.1



Red Hat Enterprise Linux Advanced Platform

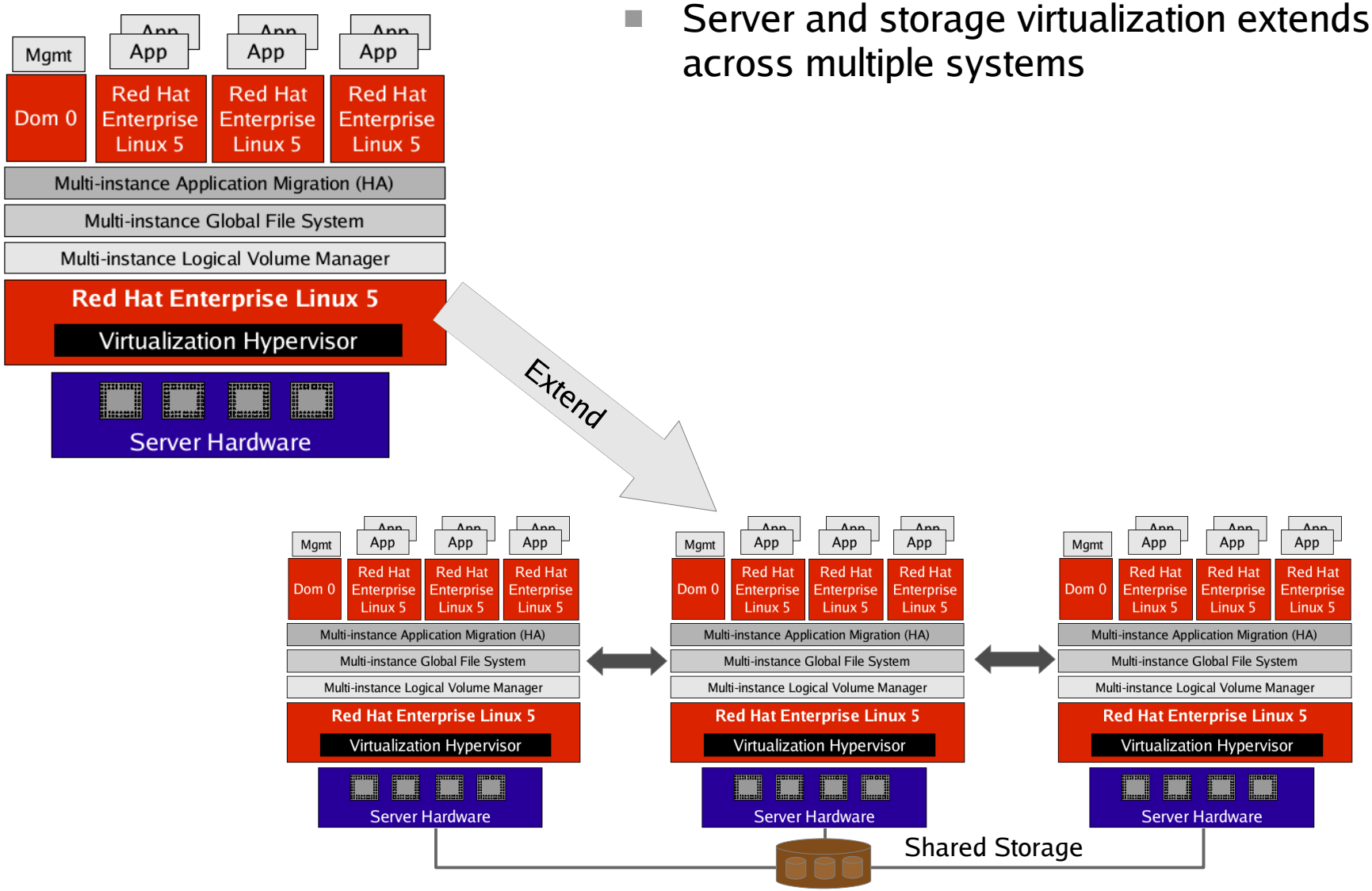
- A fully integrated server and storage virtualization environment
  - Multi Host/Instance Logical Volume Management
  - Multi Host/Instance Global File System
  - Multi Host/Instance Application Migration (with Cluster Suite failover)



- **Provides a complete virtualization platform**
  - **Server : Storage : Management**
  - Simplifies deployment & manageability
  - Increases flexibility & scalability
  - Integrates server & storage virtualization with no special hardware
  - Server & storage resources may be shared or independent

Red Hat Enterprise Linux Advanced Platform

- Server and storage virtualization extends across multiple systems



Virtualization Support

Host – Guest Virtualization Support		HOST	
		64 bit	32 bit
GUEST	64 bit	Yes	No
	32 bit	Para-virt: No Full-virt: Yes	Yes
For para-virtualization support the host & guest environments must match			

Notes:
- 64 bit = x86-64 compliant architecture
- 32 bit = x86 with PAE compliant architecture
- Host = Red Hat Enterprise Linux 5 only
- Para-virt guests = Red Hat Enterprise Linux 4.5 & 5 only
- Full virt guests = Red Hat Enterprise Linux 2.1, 3, 4, & 5; Microsoft Windows 2000, 2003, XP
- Full virt requires AMD-V or Intel VT technology processors
- Tables do not apply to Itanium2 virtualization, in which host and guest are always compatible. For RHEL 5.0, Itanium2 virtualization is a technology preview

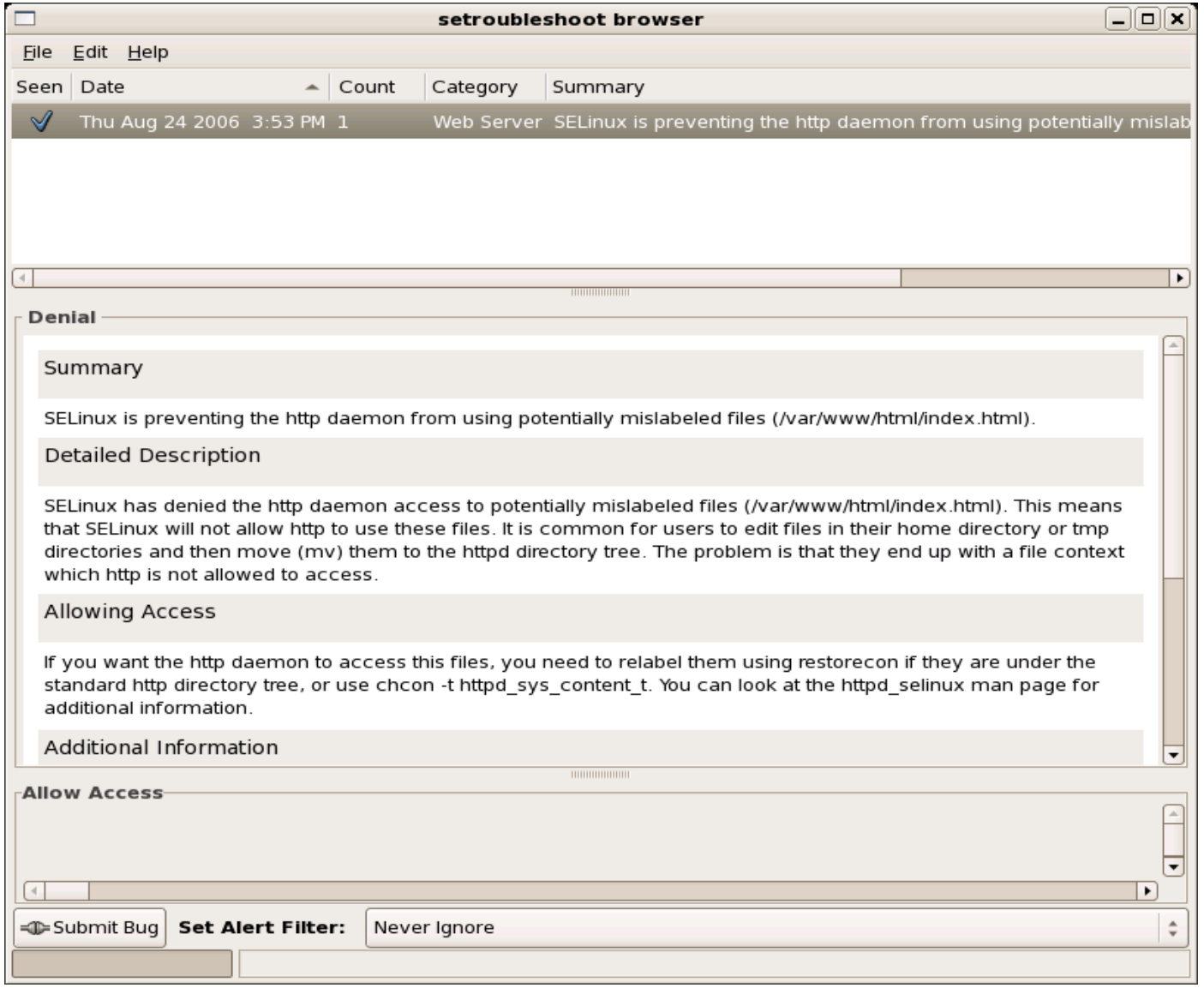
Security

Security: SELinux Enhancements

- Expanded SELinux targeted policy coverage
  - Provides coverage for all core system services, versus 11 in Red Hat Enterprise Linux 4
  - Includes support for Multi Level Security (MLS) enforcement model
    - In addition to existing RBAC and TE models
- An additional level of protection against security exploits
  - Fine-grained policies via kernel-enforced mandatory access controls
  - Limits the scope of security vulnerabilities
  - Beyond what any other general-purpose OS can deliver

Security: SELinux Enhancements

- New Setroubleshooter provides clear, easy-to-understand, GUI-based, security violation notifications
- Over 60 events defined today



Security: SELinux Enhancements

- Greatly improved logging, with easy-to-decipher information

OLD: Red Hat Enterprise Linux 4 /var/log/messages entry

```
time->Thu Aug 24 15:50:58 2006
type=AVC_PATH msg=audit(1156449058.917:552):
path="/var/www/html/index.html"
type=SYSCALL msg=audit(1156449058.917:552): arch=40000003 syscall=196
success=no exi
t=-13 a0=8d4d4d0 a1=bfb5e97c a2=434ff4 a3=2008171 items=0 ppid=23799
pid=23805 auid=3267 uid=48 gid=48 euid=48 suid=48 fsuid=48 egid=48 sgid=48 fsgid=48
tty=(none) comm="httpd" exe="/usr/sbin/httpd" subj=user_u:system_r:httpd_t:s0 key=(null)
type=AVC msg=audit(1156449058.917:552): avc: denied { getattr } for pid=23805 com
m="httpd" name="index.html" dev=dm-0 ino=6260297
scontext=user_u:system_r:httpd_t:s0
 tcontext=system_u:object_r:user_home_t:s0 tclass=file
```

NEW: Red Hat Enterprise Linux 5 /var/log/messages entry

```
Aug 24 15:53:10 localhost /usr/sbin/setroubleshootd:      SELinux is
preventing /usr/sbin/httpd "getattr" access to /var/www/html/index.html.
See audit.log for complete SELinux messages.
```

Security: Binary Code Protection

- Execshield enhancements provide additional armoring against most common kinds of security exploits
- Introducing stack “canary” word feature to detect overflow exploits
- Core packages built with new FORTIFY_SOURCE GCC option which implements run-time bounds checking to prevent buffer overflow exploits

Security: Identity Management

- Native support for Identity management in conjunction with Red Hat Directory Server and Red Hat Certificate System
- Integration of Identity & Certificate Management capabilities with Red Hat Enterprise Linux and community applications
  - Clear and secure architecture
  - Addition of Enterprise Security Client (smartcard, physical token, support)
  - Centralized key management for core desktop applications
    - system login, web browser, email, SSH
- Integration of certificate-based security and Kerberos infrastructure via PKInit
- Enables centralized management of users and rights
- Enables “Single Sign-On” user experience

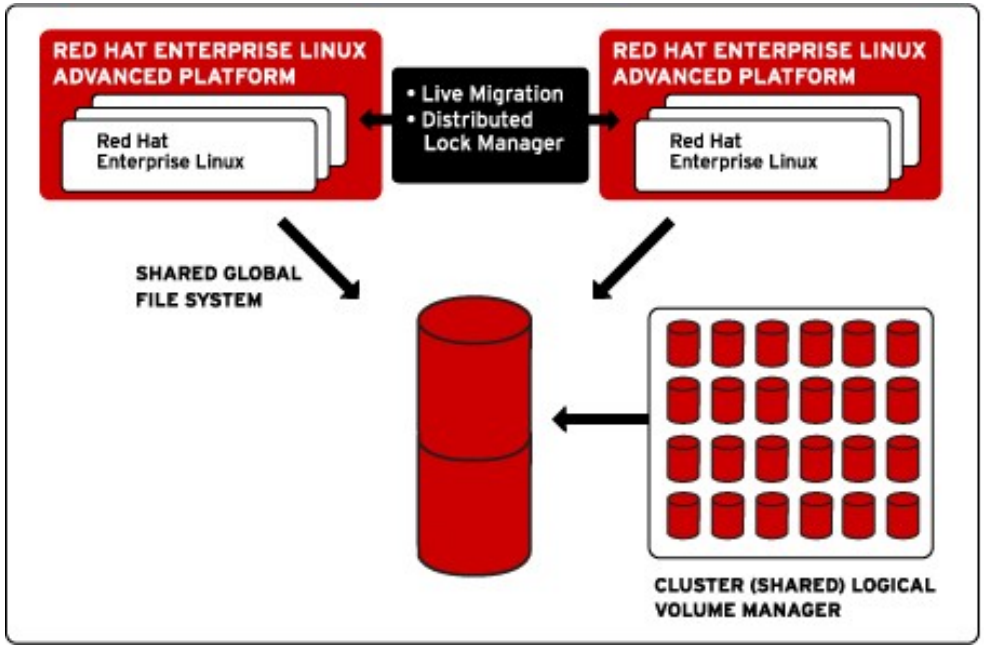
Storage
Desktop
Development Environment

Storage Improvements

- NFSv4 Improvements
  - More complete implementation of the specification
    - Delegation (aka lease), increased client caching
    - Server migration (failover)
  - Improved security integration
    - Kerberos authentication
    - 2 different encryption options, header-only & payload
  - Performance improvements – FS Cache integration
- iSCSI Software Target under consideration for Red Hat Enterprise Linux 5.1
- Ext3 enhancements for speed and scalability
- More complete automounter with Autofs5
- Volume Management, Multipathing and SAN integration improvements.
- Single node GFS2 in base OS
  - Scales beyond the current Ext3 8TB limit and provides performance enhancements.

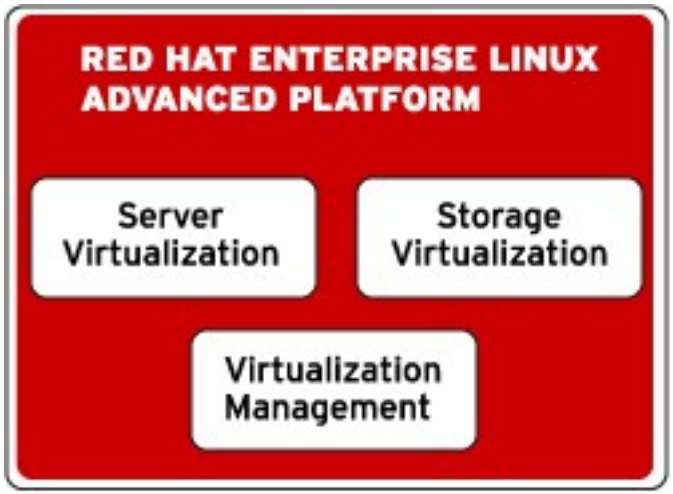
Advanced Platform benefits

- Designed for mainstream customers who seek the advantages of virtualization without complexity or risk
- Provides a full virtualization infrastructure in one complete package
- Eliminates the need to create a solution using multiple products from different vendors
- Reduced solution cost
- Reduced deployment time
- Integrated installation and management
- Extensive documentation
- Training services
- Installation and maintenance services



Advanced Platform technical benefits

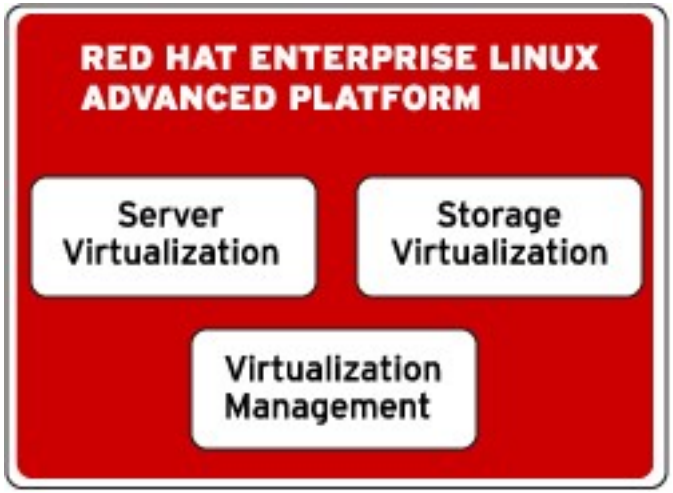
1) Enhanced application availability – multi-system and multi-site (Disaster tolerance)
   - Red Hat Cluster Suite
     - Provides application failover
     - Deployable with wide range of hardware
     - Highly scalable (up to 125 nodes)
     - Multiple application control scripts (MySQL, Apache, OpenLDAP, NFS, Oracle, SAP [some in 5.1])
     - Flexible (fine application granularity)
     - Commercial strength
       - Novell offers heartbeat – Inferior, underfunded, small-scale project



2) Virtualized storage – Easy to manage, easy to grow
   - Cluster Logical Volume Manager
     - Provides consistent storage across entire virtualized environment
     - SAN (Fibre Channel and SAN)
     - Concatenation, striping, mirroring, dynamic volume growth
     - Not available from Novell – Requires proprietary solution (Veritas)

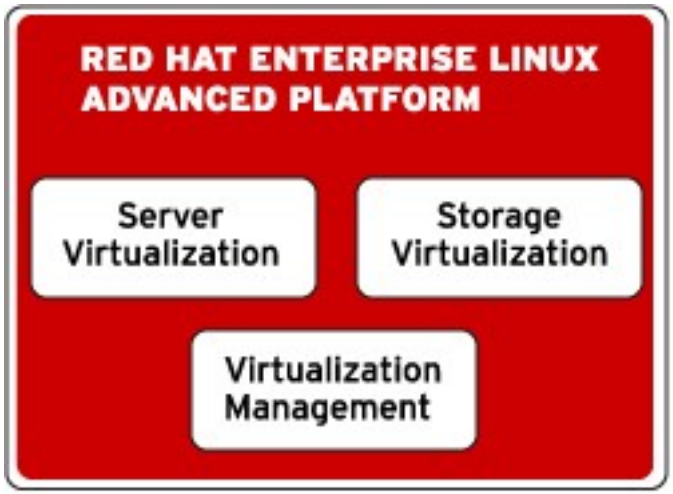
Advanced Platform technical benefits

4) Distributed application synchronization
   - Distributed Lock Manager
     - Enables applications to synchronize across guests and hosts
     - Scalable, rich API, efficient messaging
     - Included in the upstream kernel
     - Red Hat project
5) Intuitive multi-system/guest storage and cluster management
   - Conga system management GUI
     - Web-based
     - Can manage multiple clusters and large storage infrastructures
     - Logical, easy to use

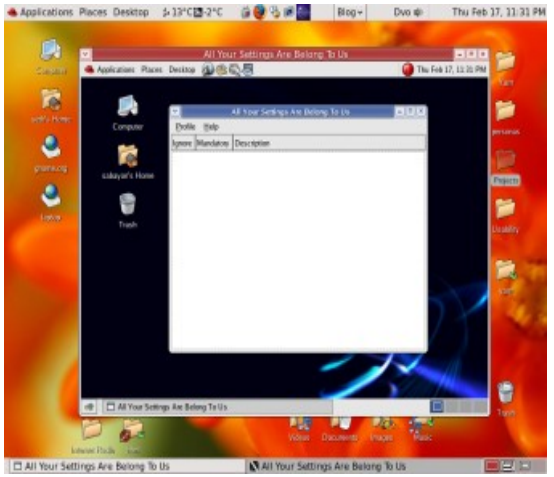


Advanced Platform benefits: summary

- Red Hat Enterprise Linux Advanced Platform enables customers to
  - Increase operational flexibility
  - Reduce costs
  - Fully enjoy the power and benefits of a virtualized environment
  - Deploy a complete, integrated solution quickly and cost effectively
- Includes all the necessary technologies
  - Virtualized storage
  - Enhanced application availability
  - Shared, scalable, high-performance data
  - Distributed application synchronization
  - Intuitive multi-system/guest storage and cluster management



Desktop Environment

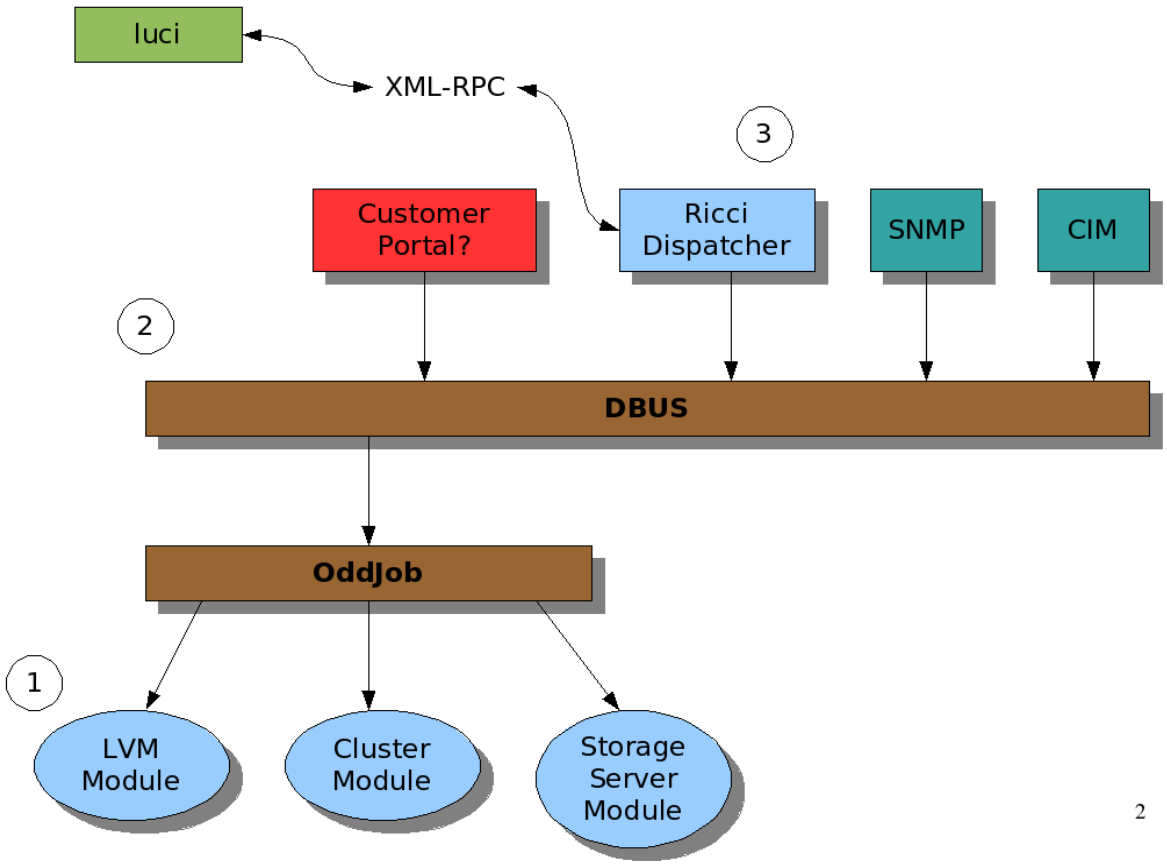


- Foundation for Stateless Linux project
- Updated desktop environment and applications
- Sabayon planned for inclusion
  - New tool enables central management of desktop settings
- X.Org - Modularization of Xorg into multiple packages
  - Improves maintainability – groundwork for new acceleration architecture, Look & Feel improvements
- ACPI enhancements.
- Internationalization and Localization
  - Additional languages and wider font support
  - Improved input method integration with desktop
- Network Manager
  - Automatic management of wired/wireless
  - Network environments, secure network access and VPN support

Management

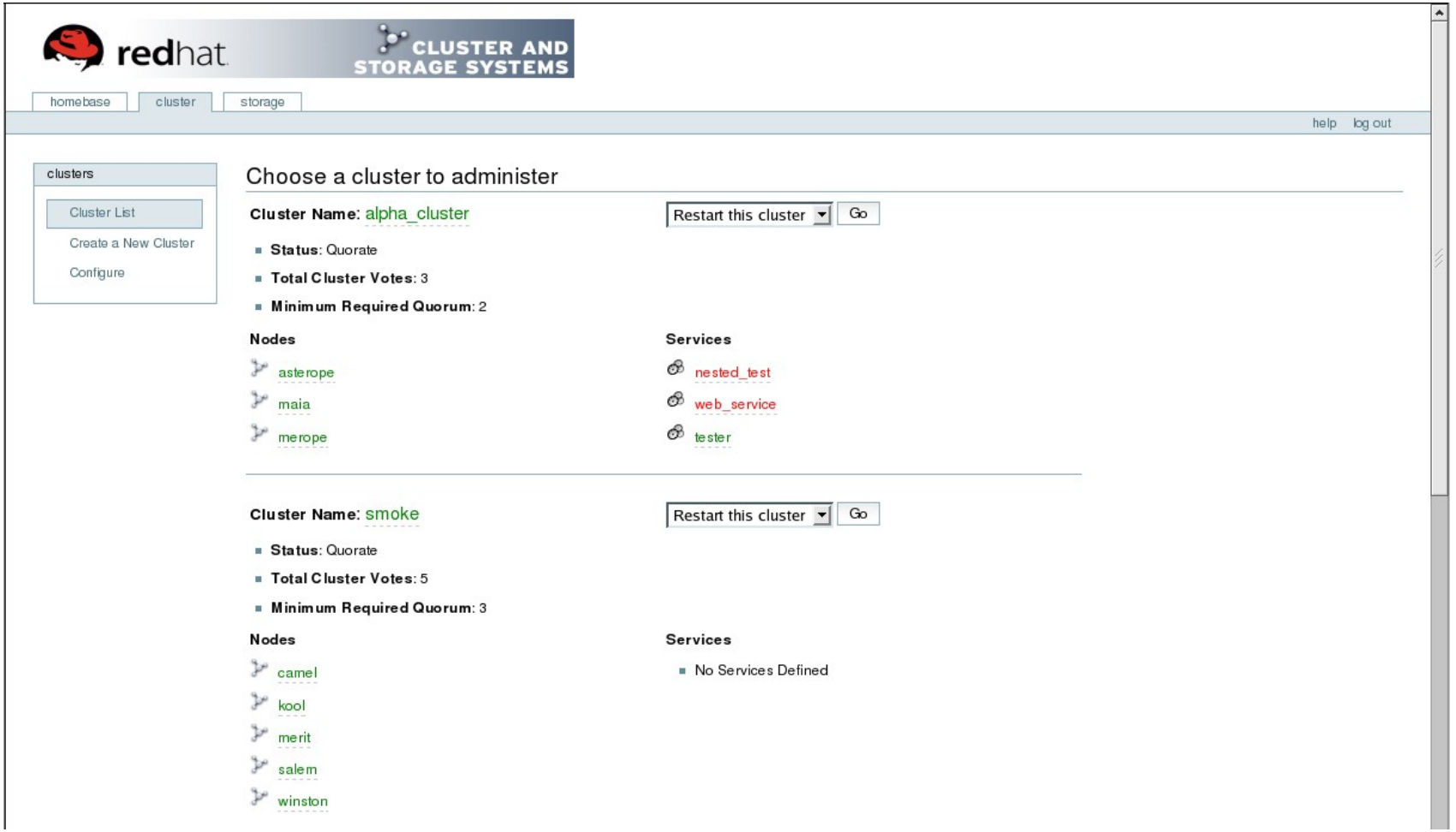
Cluster & Storage Management: Conga

- New web/XML based management tool for cluster and storage managment
- Provides multi-system and multi-guest management
  - Installation, configuration, control
  - Flexible, extensible
- Included with Advanced Platform



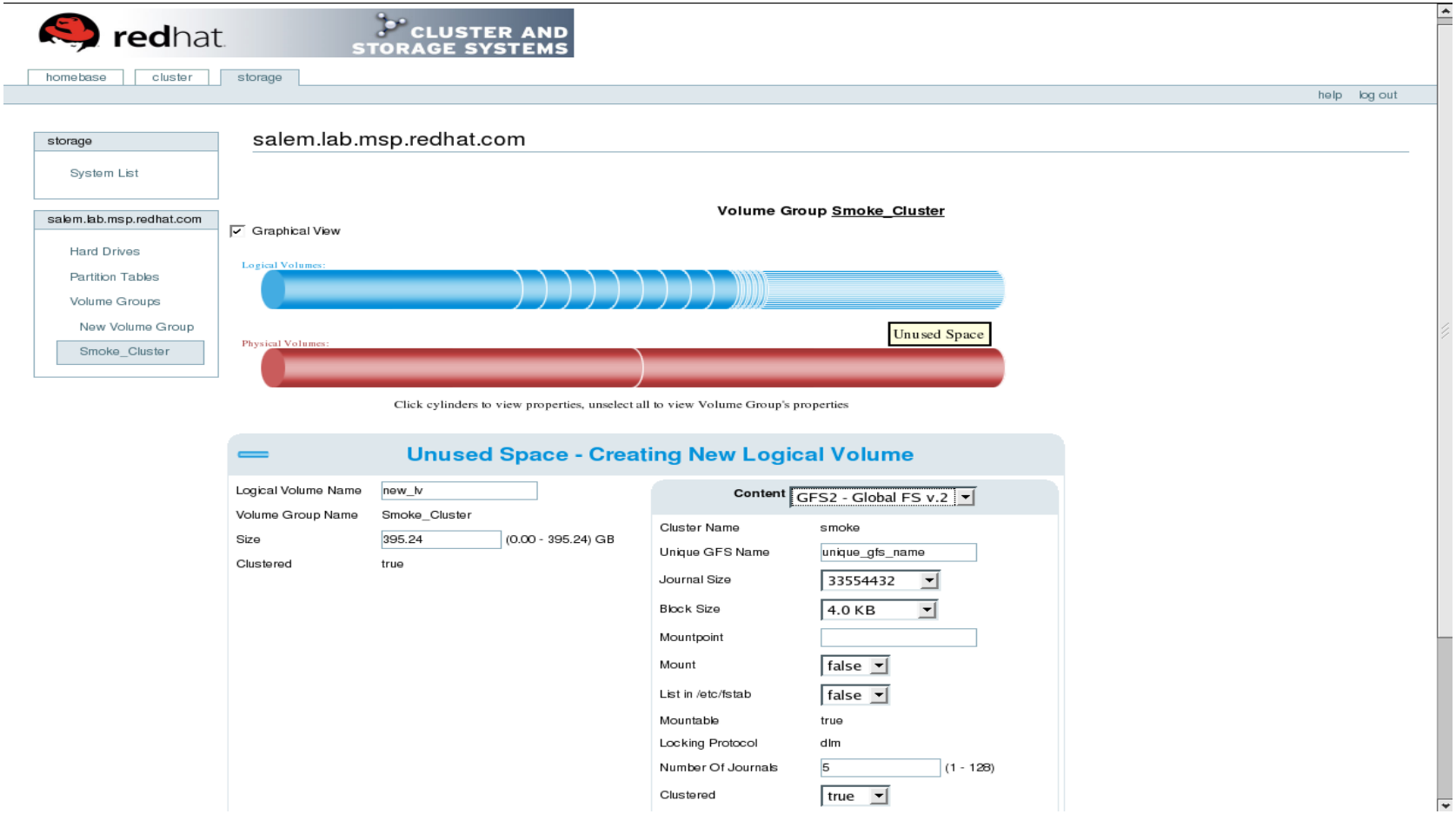
Conga: Cluster Management

- Controls multiple clusters
- Comprehensive features



Conga: Storage Management

- Controls multi-system and multi-guest logical volumes
- Comprehensive features



Red Hat Enterprise Linux 5

General New Features

RED HAT
ENTERPRISE LINUX ADVANCED PLATFORM

Delivers: Lower Operating Costs

- Lifecycle management: application encapsulation
- Administration overhead: sysadm efficiency ratio, server consolidation, reduced time to deploy servers
- Lower power consumption: dynamic resource allocation, speed scaling
- Reduced facilities cost: dynamic resource allocation, better performance
- Lower cooling cost: dynamic resource allocation, speed scaling
- Reduced overtime cost: elimination of outage windows, availability
- Faster time to solution: dynamic resource allocation, virtual sandbox

RED HAT
ENTERPRISE LINUX ADVANCED PLATFORM

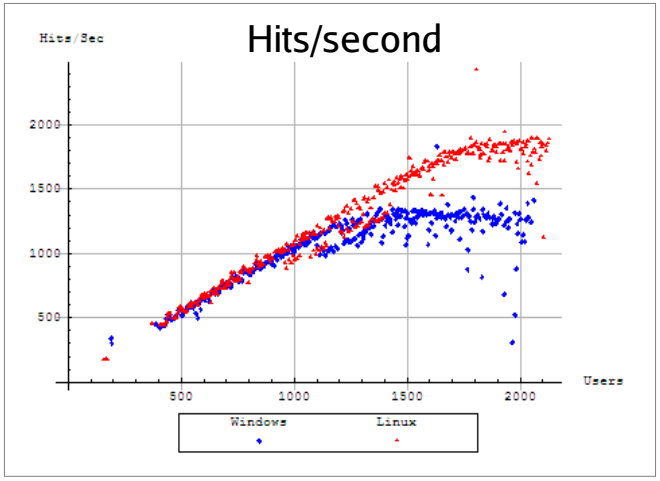
Delivers: Better Service Delivery

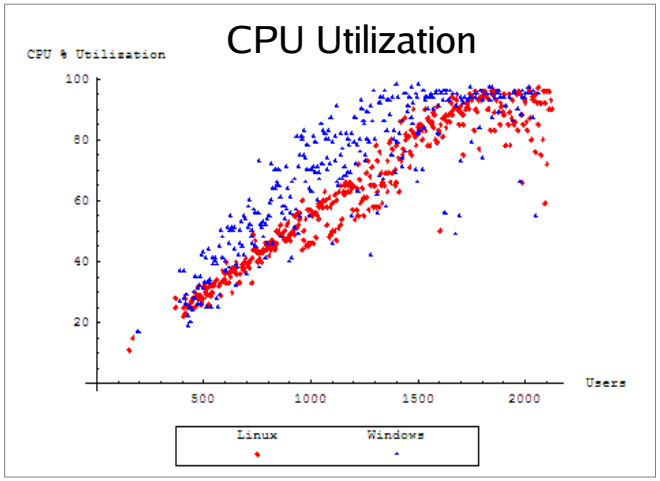
- Increased availability: redundancy, failover, migration, elimination of planned downtime
- Better response times: dynamic infrastructure, load balancing
- Faster application and service delivery: eliminate purchasing cycle, dynamic provisioning, live migration
- ... and much more with JBoss

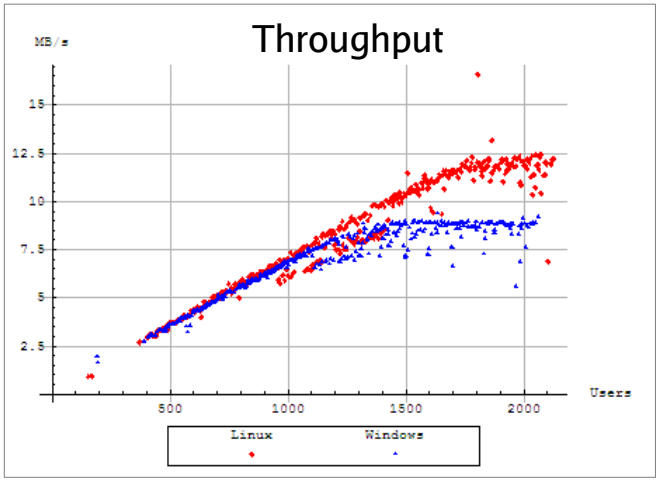
Performance

Red Hat Enterprise Linux performance is proven across multiple workloads

- Many world record performance results
- Example: Apache Tomcat running on Red Hat Enterprise Linux and Windows 2003, tested by www.webperformanceinc.com (2006)



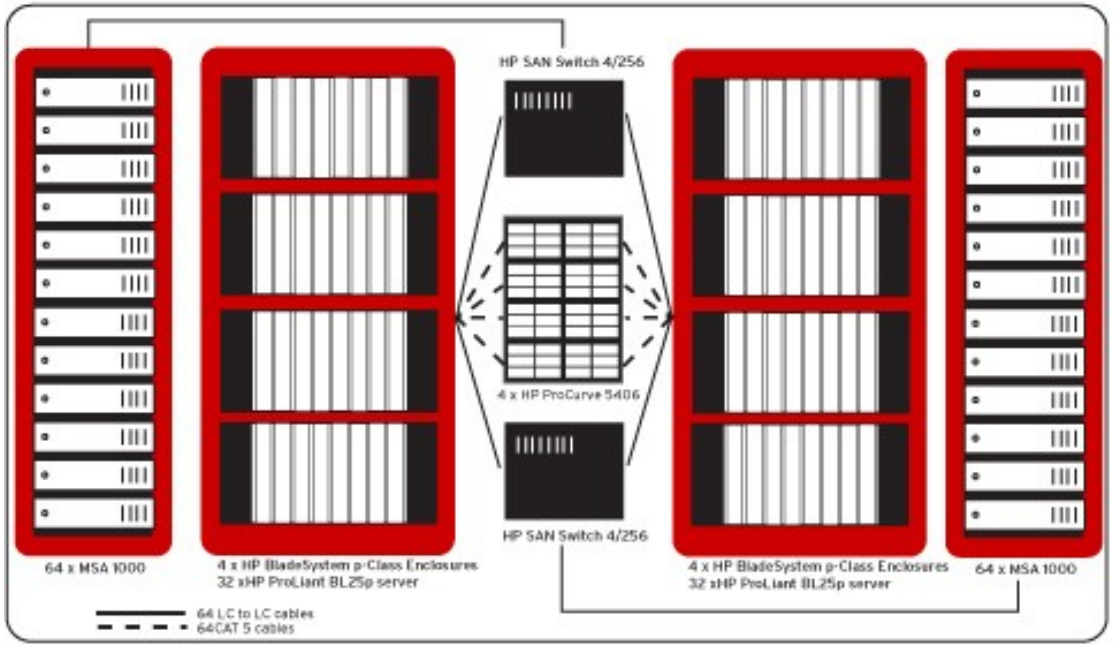




Red Hat Enterprise Linux can handle more web requests with lower CPU utilization and higher throughput than Windows.

Recent benchmark results

- World record TPC-H performance with 3000GB database
  - HP : Oracle : Red Hat
  - 5% faster and 30% cheaper than #2 : Sun Solaris 10 on E25K SPARC Server
- Red Hat Enterprise Linux also holds six of the top ten results at 300GB database size



System: HP BladeSystem ProLiant BL25p cluster 64P DC Spec.

Performance: 110,576.5 QphH@3000GB

Price/Performance: $37.80 USD/QphH@3000GB

Database total system cost: $4,179,238 USD

Database software: Oracle Database 10g Release 2, Enterprise Edition with Oracle Real Application Clusters and Partitioning

Operating system: Red Hat Enterprise Linux 4ES

Total # Nodes/Processors/Cores/Threads: 64/64/128/128

Processors: Dual-Core AMD Opteron™ 285, 2.6GHz/1MB

Availability: June 8, 2006 Submitted: June 8, 2006

Source: www.tpc.org 20-Jan-2007

SPECweb2005: 1st, 2nd, 3rd, 4th ,5th, 6th & 7th!

- Red Hat Enterprise Linux dominates web serving performance testing

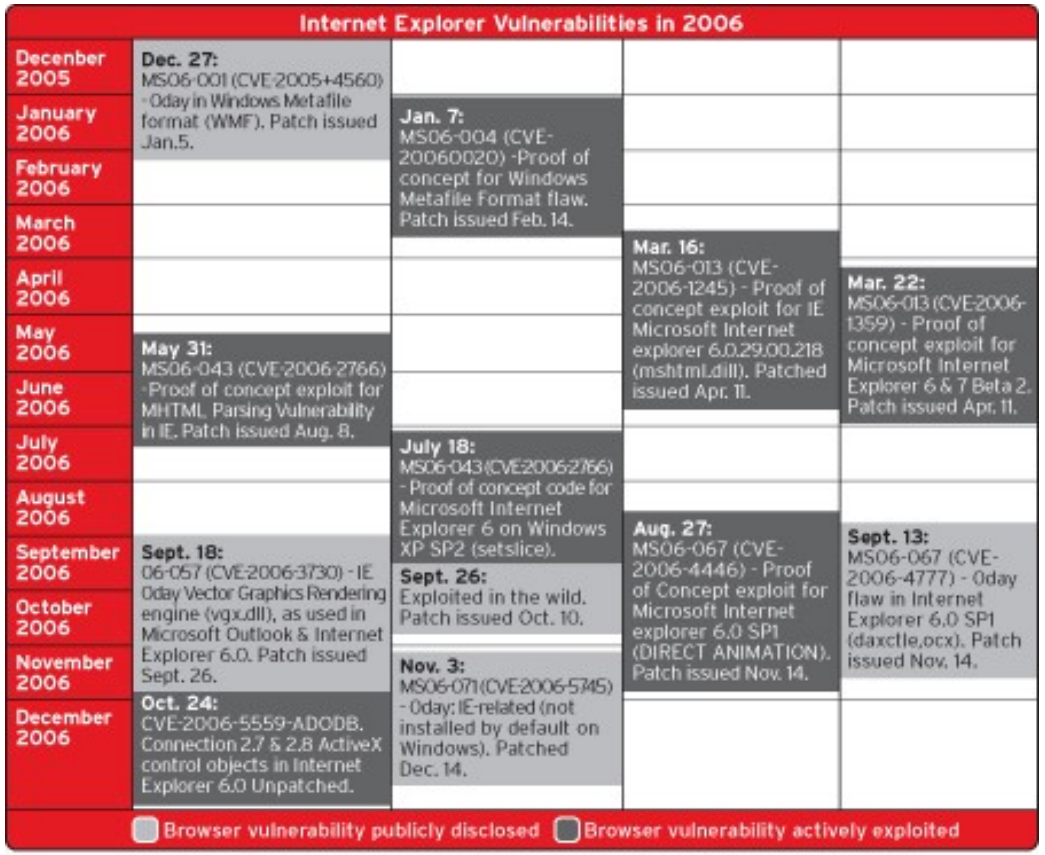
HP ProLiant DL585 G2 (AMD Opteron 8220, 8 cores, 4 chips, 2 cores/chip) Result (simultaneous users): SPECweb2005 = 20,235	hp
Fujitsu Siemens PRIMERGY TX300 S3, (Intel Xeon, 8 cores, 2 chips, 4 cores/chip) Result (simultaneous users): SPECweb2005 = 18,160	FUJITSU COMPUTERS SIEMENS
Dell PowerEdge 2950 (Intel Xeon, 8 cores, 2 chips, 4 cores/chip) Result (simultaneous users): SPECweb2005 = 16,830	DELL
HP ProLiant DL585 (AMD Opteron 885, 8 cores, 4 chips, 2 cores/chip) Result (simultaneous users): SPECweb2005 = 15,850	hp
HP ProLiant DL585 G2 (AMD Opteron 8212, 8 cores, 4 chips, 2 cores/chip) Result (simultaneous users): SPECweb2005 = 14,964	hp
Fujitsu Siemens PRIMERGY RX600 S3, (Intel Xeon, 8 cores, 4 chips, 2 cores/chip) Result (simultaneous users): SPECweb2005 = 14,896	FUJITSU COMPUTERS SIEMENS
Dell PowerEdge 2950 (Intel Xeon, 4 cores, 2 chips, 2 cores/chip) Result (simultaneous users): SPECweb2005 = 14,495	DELL

Source: www.spec.org 20-Jan-2007

Open source security

- Red Hat no longer does Windows vulnerability comparisons “because Windows deliberately does not disclose all their vulnerabilities, making vulnerability count comparisons worthless“ - *Mark Cox, Red Hat Security Response Team Leader*
- One 2006 comparison:
  - Internet Explorer was unsafe for 284 days
  - Firefox was unsafe for 9 days

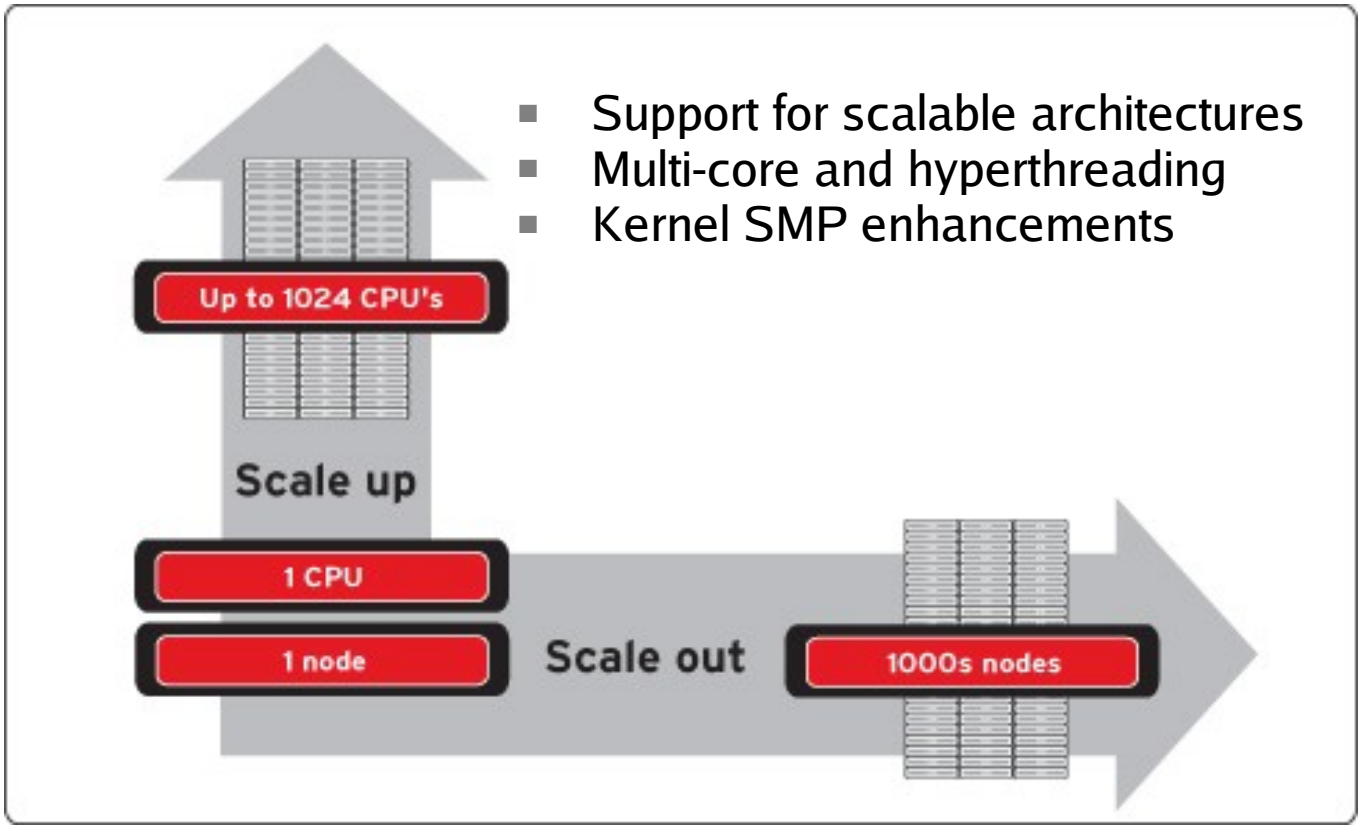
Source: Washington Post



Source: http://www.washingtonpost.com/wp-srv/technology/daily/graphics/index20070104.html

Red Hat Enterprise Linux: Scale up and out

- Traditional scale-out capabilities have been complemented over the past two years with scale-up capabilities
  - Red Hat brings open source value and flexibility to the traditional large UNIX market



Kernel Performance/Scalability

- The Red Hat Enterprise Linux 5 kernel includes many new features implemented between the 2.6.9 & 2.6.17/8 kernels (2 years of development). Including:
- Voluntary preemption patches (2.6.13) (subset in RHEL4)
  - Reduces scheduling latency (<1ms) enabling improved performance for applications such as audio/video
- Lightweight userspace priority inheritance (PI) support for futexes, useful for realtime applications (2.6.18)
  - Assists priority inversion handling. Ref: http://lwn.net/Articles/178253/
- New 'mutex' locking primitive (2.6.16)
  - Similar to spinlocks, but permitted to block
- High resolution timers (2.6.16)
  - Provide fine resolution and accuracy depending on system configuration and capabilities - used for precise in-kernel timing

Kernel Performance/Scalability Features

- Modular, on-the-fly switchable I/O schedulers (2.6.10)
  - Only provided as a boot option in RHEL4
  - Improved algorithms (esp. for CFQ)
  - Per-Queue selectable (previously system-wide)
- Conversion to 4-level page tables (2.6.11) (Architecture specific)
  - Allows x86-64 to increase from 512G to 128TB of virtual address space
- New Pipe implementation (2.6.11)
  - 30-90% perf improvement in pipe bandwidth
  - Circular buffer allow more buffering rather than blocking writers
- "Big Kernel Semaphore": Turns the Big Kernel Lock into a semaphore
  - Latency reduction, by breaking up long lock hold times and adds voluntary preemption
- X86 "SMP alternatives"
  - Optimizes a single kernel image at runtime for UP or SMP operation.
  - Ref: http://lwn.net/Articles/164121/

Monitoring Features

- Inotify (2.6.13)
  - New file system event monitoring mechanism (replaces dnotify)
  - Ideal for security and performance monitoring
- Process Events Connector (2.6.15)
  - Reports fork, exec, id change, and exit events for all processes to userspace
  - Useful for accounting/auditing (e.g. ELSA), system activity monitoring, security, and resource management
- Blktrace
  - Block queue IO tracing – monitor block device queue traffic (2.6.17)

Generic Features

- kexec & kdump (2.6.13)
  - Provide new crash-dumping capability with reserved, memory-resident kernel
- Generic RTC (Real Time Clock) subsystem (2.6.17)
- Splice()
  - Application file connector for high performance kernel-based data transfer (2.6.17)
  - Ref: http://lwn.net/Articles/178199/

File System Features

- EXT3
  - Ext3 block reservation & on-line growth (2.6.10 & RHEL4)
  - Extended Attributes in the body of large inode
    - Saves space and improves performance (2.6.11)
  - Increases maximum ext3 file-system size from 8TB to 16TB (2.6.18)
- Device mapper multipath support (RHEL4)
- ACL support for NFSv3 and NFSv4 (2.6.13)
- NFS
  - Support large reads and writes on the wire (2.6.16)
  - Linux NFS client supports transfer sizes up to 1MB

File System Features

- VFS changes
  - "Shared subtree" enhancement (2.6.15). Ref: http://lwn.net/Articles/159077/
  - Read Copy Update used for improved performance & locking
- Big CIFS update: (2.6.15)
  - Lots of performance improvements
  - Kerberos and CIFS ACL support
- Autofs4
  - Updated to provide direct mount support for userspace autofs (2.6.18)
- FS-Cache
  - Core enablers (2.6.18)
  - Tech preview until 5.1

Security Features

- Address space randomization:
  - Address randomization of multiple entities – including stack & mmap() region (used by shared libraries) (2.6.12; more complete implementation than in RHEL4)
  - Greatly complicates and slows down hacker attacks
- Multilevel security (MLS) implementation for SELinux (2.6.12)
  - Third policy scheme for SELinux, with RBAC & TE
  - TBS
- Audit subsytem
  - Support for process-context based filtering (2.6.17)
  - More filter rule comparators (2.6.17)
- TCP/UDP getpeersec
  - Enable a security-aware application to retrieve the security context of an IPSec security association a particular TCP or UDP socket in using (2.6.17)

Networking

- Provision of several selectable TCP congestion modules (2.6.13)
  - Ref: http://lwn.net/Articles/128681/
- IPV6 - Support several new sockopt / ancillary data in Advanced API (2.6.14)
  - TBS
- IPv4/IPv6: UFO (UDP Fragmentation Offload) (2.6.15)
  - Offloads IP fragmentation functionality of large UDP datagram to hardware
  - Improves performance
- Add nf_conntrack subsystem: (2.6.15)
  - Common IPv4/IPv6 generic connection tracking subsystem
  - Allows IPv6 to have a stateful firewall capability (not previously possible)
    - Increased security
    - Enables analysis of whole streams of packets, rather than only checking the headers of individual packets

Networking

- IPv6
  - RFC 3484 compliant source address selection (2.6.15)
  - Add support for Router Preference (RFC4191) (2.6.17)
  - Add Router Reachability Probing (RFC4191) (2.6.17)
- Generic segmentation offload (GSO) (2.6.18)
  - Available in place of TSO (TCP Segmentation Offload)
  - Performance improvements for large packet transfers without hardware assistance
- SELinux per-packet access controls
  - Replaces old packet controls
  - Add Secmark support to core networking
    - Allows security subsystems to place security markings on network packets (2.6.18)
- Inclusion of DCCPv6 – Datagram Congestion Control Protocol (2.6.16)

Networking: Wireless Support

- Numerous updates, including:
- Performance:
  - IEEE 802.11 hardware crypto and fragmentation offload support (2.6.15)
  - QoS (WME – Wireless Multimedia Enhancements) support
- Management:
  - Wireless spy support – simplifies network management
  - WE-20, version 20 of the Wireless Extensions (2.6.17)
- Development:
  - Add the hardware independent software MAC layer (2.6.17) – simplifies wireless software development
- Security:
  - Mixed PTK/GTK (Pairwise Transient Key/Group Temporal Key) support
  - CCMP/TKIP (Counter Mode with Cipher Block Chaining Message Authentication Code Protocol/Temporal Key Integrity Protocol) support
  - Add LEAP (CISCO's Lightweight Extensible Authentication Protocol) authentication type
- Hardware support:
  - BCM43xx wireless driver (open source Broadcom-compatible driver)
  - ZyDAS ZD1211 wireless driver

New Hardware

- Huge number of new hardware support features/platforms/drivers/etc
- General features:
  - x86-64 clustered APIC support (2.6.10)
  - Infiniband support (2.6.11) (mostly in RHEL4)
  - Hot plug
    - Generic memory add/remove & supporting functions (2.6.15)
    - (i386) hot plug CPU support of physical add of new processors (hotplug disable/enable of already existing CPUs was already supported)
- SATA/libata enhancements, additional hardware support (in RHEL4)
  - Increased SATA subsystem reliability
  - Increased performance with Native Command Queuing (NCQ) (2.6.18)
  - Hotplug support (2.6.18)
- EDAC (Error Detection & Correction) support (2.6.16) (in RHEL4)
  - Detects and reports errors that occur within the system (ECC memory, PCI bus parity, cache, etc)
- New ioatdma driver for the Intel(R) I/OAT DMA engine (2.6.18)

NUMA & Multi Core

- Cpusets (2.6.12)
  - Enable CPU & Memory assignment to sets of tasks
  - Allow dynamic job placement on large systems
- Numa-aware slab allocator (2.6.14)
  - Optimized locality & management of slab creation
- Swap migration. (2.6.16)
  - Swap migration relocates physical pages between nodes in a NUMA system while the process is running – improves performance
- Huge page support for NUMA (2.6.16)
- Netfilter ip_tables: NUMA-aware allocation (2.6.16)
- Multi-core
  - Scheduler improvements for shared-cache multi-core systems (2.6.17)
  - Scheduler power saving policy
    - Power consumption improvements through optimized task spreading