Today

PowerKVM Deep Dive
Starting at 10:00 am UK time by Dr. Michael Perzl

Smart Meeting → Put questions into the Chat box
or AT&T Toll Free phone for better audio
- 0800-368-0638 = UK Toll Free
- 0203-059-6451 = UK but you pay for the call
- Then 6403785# Participant Code
- Other countries see chat box for the website
- Please Mute with *6

Previous Sessions:
More Tricks Power Masters
Power8 from hands-on
Power up your Linux
PowerVC
PowerVP
SSP4
Best Practices
Tricks of Power Masters
IBMi and External Storage
Monitoring with ITM
Whole Machine Monitoring
Electric Server Agent
And more…..

Future Sessions →

- To be resumed in September
- Suggestions Welcome

Twitter:
Gareth Coates @power_gaz   Nigel Griffiths @mr_nmon
Jyoti Dodhia @JyotiDodhia   Mandie Quartly @mandieq
Website: http://tinyurl.com/PowerSystemsTechnicalWebinars
Youtube Channel: http://tinyurl.com/IBMPowerVUGYoutubeChannel

Power Virtualization Options

PowerKVM

Q2 2014
Initial Offering

PowerKVM provides an open source choice for Power Virtualization for Linux workloads. Best for clients that aren't familiar with Power and Linux centric admins.

PowerVM

2004
Initial Offering

PowerVM is Power Virtualization that will continue to be enhanced to support AIX, IBM i Workloads as well as Linux Workloads

IBM PowerKVM webinar, July 16, 2014

© 2014 IBM Corporation

New IBM Power Systems based on POWER8 – April 2014

- ❑ POWER8 roll-out is leading with scale-out (1-2S) systems
- ❑ Expanded Linux focus: Ubuntu, KVM, and Open Stack
- ❑ Scale-up POWER8 (>2S) systems will be rolled out over time
- ❑ PCI Gen3 right out of POWER8 processor
- ❑ OpenPOWER Innovations

Power Systems S812L
- 1-socket, 2U
- Linux Only

Power Systems S822L
- 2-socket, 2U
- Linux Only

Power Systems S822
- 2-socket, 2U
- All Operating Systems

Power Systems S814
- 1-socket, 4U
- All Operating Systems

Power Systems S824L
- 2-socket, 4U
- Linux Only
- SOD

Power Systems S824
- 2-socket, 4U
- All Operating Systems

1 & 2 Sockets

Linux

ubuntu®

openstack
CLOUD SOFTWARE

SUSE

redhat.
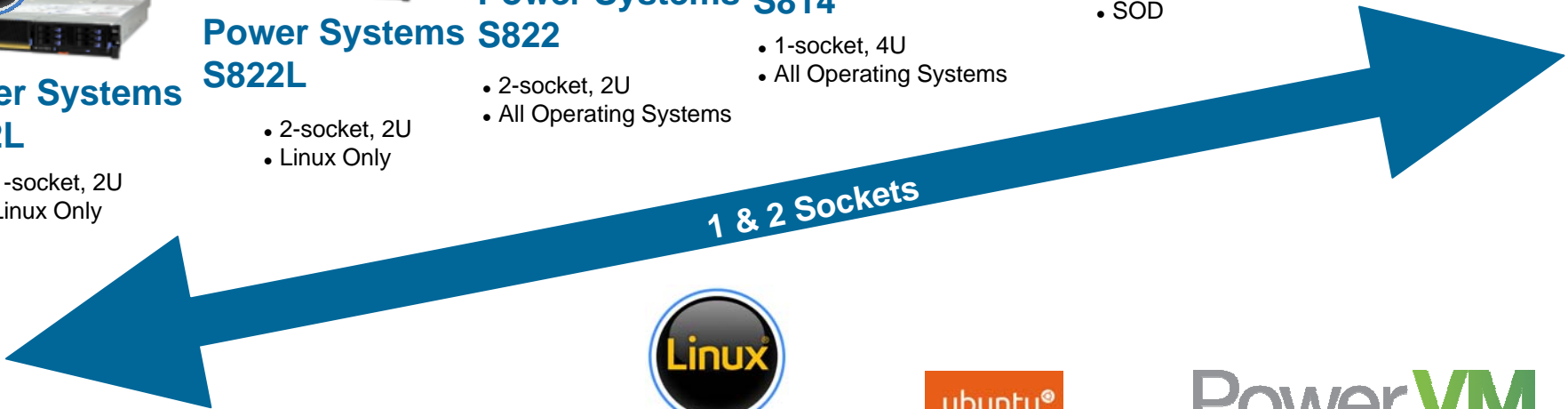
PowerKVM

Power VM

Power SC   **Power VC**

Power HA   **Power VP**

IBM PowerKVM webinar, July 16, 2014

Endianness

Why do I care about endianness?

- Linux on Power has chosen to exploit little endian (LE) processor mode based on OpenPOWER partner feedback.

  – Eases the migration of applications from Linux on x86.

  – Enables simple data migration from Linux on x86.

  – Simplifies data sharing (interoperability) with Linux on x86.

  – Improves Power I/O offerings with modern I/O adapters and devices, e.g. GPUs.

- Creation of an LE operating system for Linux on Power means creating a whole new software "platform" (**ppc64le**) (in addition to BE **ppc** (32-bit) and BE **ppc64** (64-bit)).

- LE distributions for Linux on Power **does NOT mean** x86 applications magically run: applications must still be compiled for Power.

- Power8 CPU can be either big or little endian
  ➔ mixed endianness (big and little) on same system will be possible.
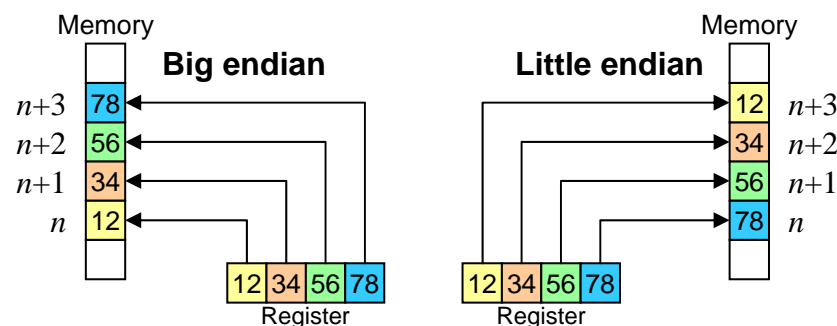
Linux distributions for Power Systems

Linux support for POWER

> ➢ Built from the same source as x86
> ➢ Delivered on the same schedule as x86
> ➢ Supported at the same time as x86

redhat **fedora**

SUSE **openSUSE**

ubuntu®

- RHEL 7
  - POWER8 (native mode) and POWER 7/7+ at GA

- RHEL 6
  - POWER8 supported with U5 (P7-compatibility mode)
  - Full support of POWER6 and POWER7 (native mode)

- Fedora
  - Fedora 16 was first release to re-launch POWER
  - Fedora 20 has POWER8 support

- Supported add-ons
  - JBoss
  - High Performance Network Add-on

- SLES 11
  - POWER8 with SP3 (P7-compatibility mode)
  - POWER7+ encryption, RNG accelerators with SP3
  - Full support of POWER7 (native mode)

- SLES 10
  - POWER7 supported with SP3 (P6-compatibility mode)
  - Full support of POWER6 (native mode)

- openSUSE
  - openSUSE 12.2 re-launched for IBM POWER
  - openSUSE 13.2 includes POWER8 support

- Supported add-ons
  - SUSE Linux Enterprise High Availability Extension

- Ubuntu 14.04
  - POWER8 enabled (native mode)
  - No official support for POWER7+ and older systems
  - No support for 32-bit applications. 64-bit only.
  - Supported in KVM only at this time

- Supported add-ons
  - JuJu Charms
  - MaaS (Metal as a Service)
  - Landscape

- Debian
  - Community enablement in process

What Linux Distributions in various Power Environments?

Linux	Release	Endian	Dedicated LPAR	PowerVM Guest	PowerKVM Guest**
RHEL	5.10	Big	✔	✔	✘
RHEL	6.4	Big	✔	✔	✘
RHEL	6.5	Big	✔	✔	✔
RHEL*	7	Big	✔	✔	✔
SLES	10 SP4	Big	✔	✔	✘
SLES	11 SP3	Big	✔	✔	✔
SLES*	12	Little	✘	✘	✔
Ubuntu*	14.04	Little	✘	✘	✔

Notes:
*Exploits P8
**Requires all guests be LE or all BE. Mixing of BE and LE guests not yet supported.

1. Select the applications you want to run on Linux on Power.
2. Then look at the Linux distributions that are available for those apps.
3. Pick your Linux distribution of choice.

Power8 endianness and OS choices

- Power8 CPU can be either big or little endian
  - determined by value stored in MSR register
- Mixed endianness (big and little) on same system possible
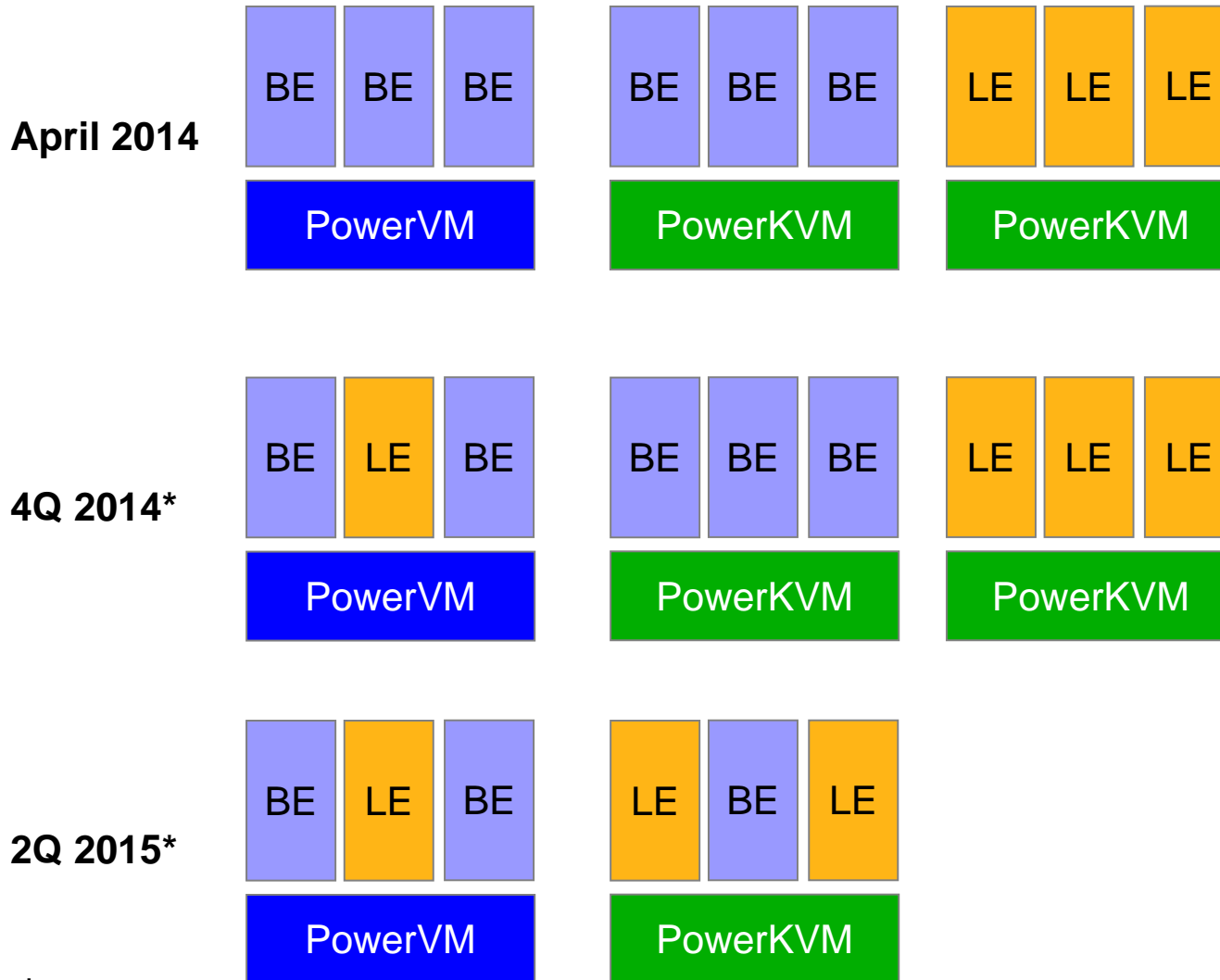 - not yet supported though (see next slide)

Big endian (ppc, ppc64)

- AIX / VIOS
- RHEL6 / RHEL 7
- SLES 11

- -

- Fedora <= 20
- Debian 7.5.0
- OpenSUSE <= 13.1

Little endian 64-bit (ppc64le, ppc64el)

- Ubuntu 14.04
- RHEL 8 ??
- SLES 12

- Fedora 21 ??
- Debian 7.X.0 ??
- OpenSUSE 13.X ??

Hypervisor support for endianness

April 2014

BE	BE	BE

PowerVM

BE	BE	BE

PowerKVM

LE	LE	LE

PowerKVM

4Q 2014*

BE	LE	BE

PowerVM

BE	BE	BE

PowerKVM

LE	LE	LE

PowerKVM

2Q 2015*

BE	LE	BE

PowerVM

LE	BE	LE

PowerKVM

* Planned

IBM PowerKVM webinar, July 16, 2014   © 2014 IBM Corporation

PowerKVM Details

KVM Architecture Overview
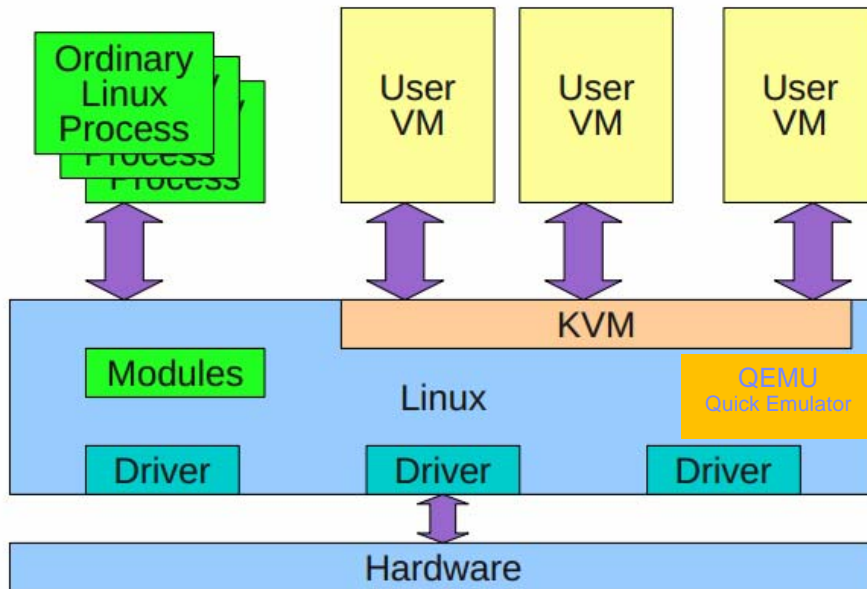
KVM – Kernel-Based Virtual Machine Loadable Kernel Module that provides server virtualization for Memory and CPU
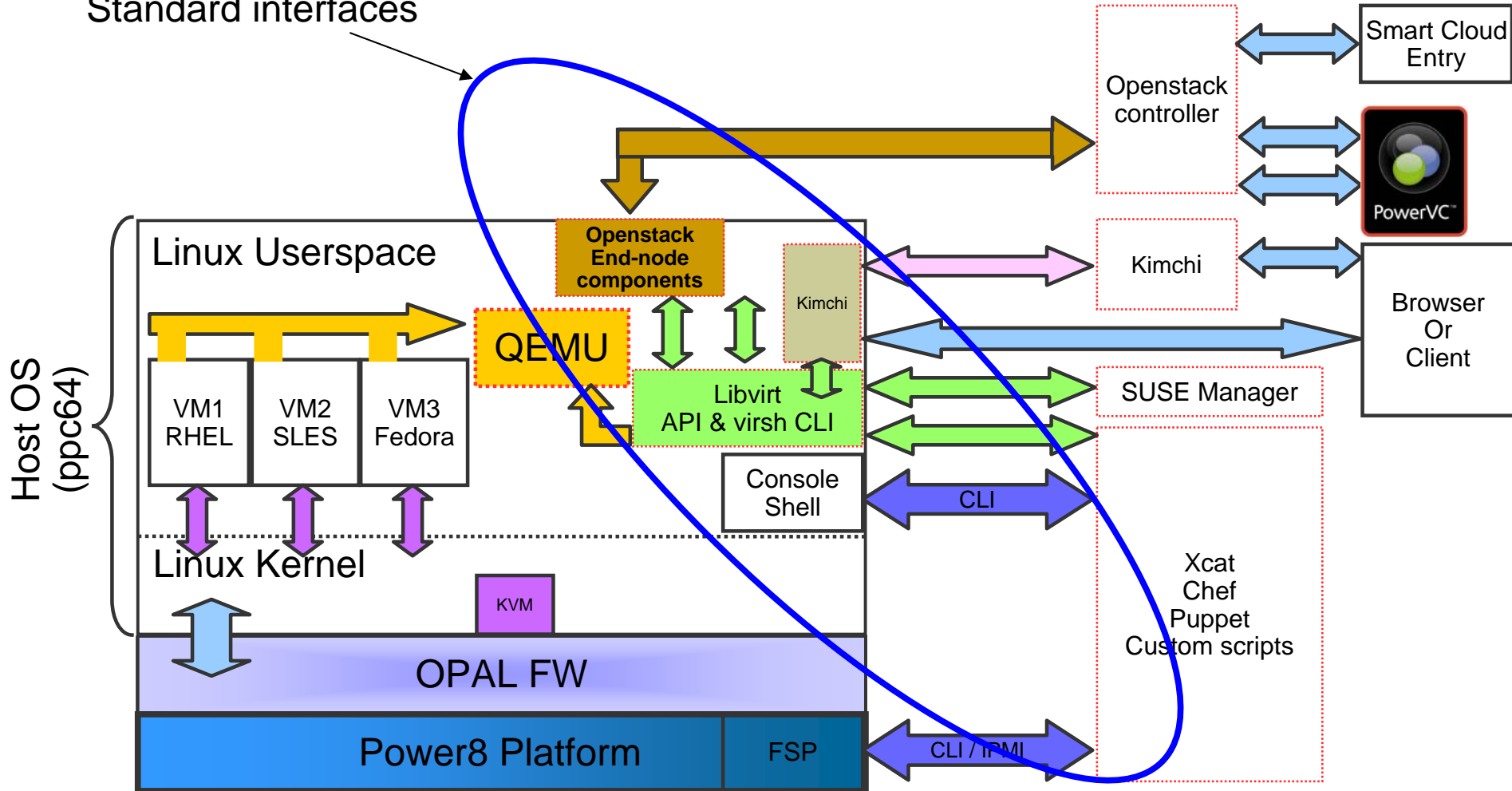**QEMU** – Quick Emulator virtualizes I/O

Guests run a normal Linux Process scheduled by the Linux Scheduler

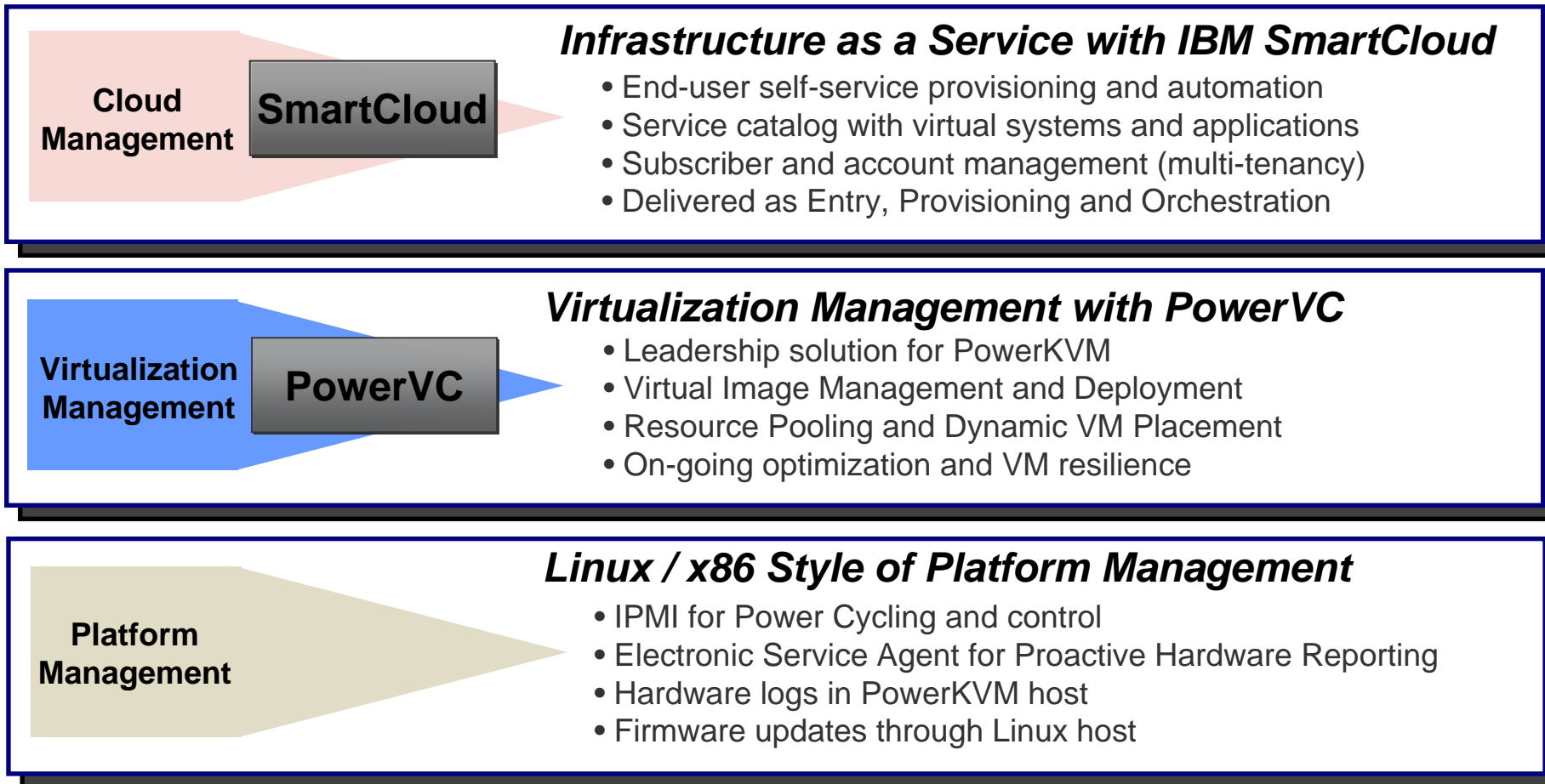Originally designed for x86 and uses hardware assists. Intel VT, AMD-V

KVM on Power – PowerKVM

Standard interfaces

- Smart Cloud Entry
- Openstack controller
- PowerVC™
- Openstack End-node components
- QEMU
- Kimchi
- Kimchi
- Libvirt API & virsh CLI
- SUSE Manager
- Browser Or Client
- Console Shell
- CLI
- Xcat Chef Puppet Custom scripts
- CLI / IPMI

Host OS (ppc64)

Linux Userspace

VM1 RHEL | VM2 SLES | VM3 Fedora

Linux Kernel

KVM

OPAL FW

Power8 Platform | FSP

IBM PowerKVM webinar, July 16, 2014

PowerKVM Management Strategy

Infrastructure as a Service with IBM SmartCloud

Cloud Management — SmartCloud

- End-user self-service provisioning and automation
- Service catalog with virtual systems and applications
- Subscriber and account management (multi-tenancy)
- Delivered as Entry, Provisioning and Orchestration

Virtualization Management with PowerVC

Virtualization Management — PowerVC

- Leadership solution for PowerKVM
- Virtual Image Management and Deployment
- Resource Pooling and Dynamic VM Placement
- On-going optimization and VM resilience

Linux / x86 Style of Platform Management

Platform Management

- IPMI for Power Cycling and control
- Electronic Service Agent for Proactive Hardware Reporting
- Hardware logs in PowerKVM host
- Firmware updates through Linux host

Just another KVM / Linux host. Normal open source tools & OpenStack can be used for management.

PowerVM vs. PowerKVM comparison

PowerKVM

PowerVM

	PowerKVM	PowerVM
Managers	PowerVC, OpenStack, libvirt, Open Source Tools	HMC, IVM, FSM, PowerVC, ISD VMControl
Guest VM Types	redhat Linux, SUSE, ubuntu	AIX, i for Business, redhat Linux, SUSE
Host Software	Linux KVM Hypervisor	VIO Server IO Virtualization
Firmware	OPAL Firmware Hardware Abstraction Boot services Standalone Diagnostics	Phyp Firmware - Hypervisor
Hardware	Power8 Linux only Hardware	P6, P7, P8 Hardware

PowerVM vs. PowerKVM comparison

	PowerVM	PowerKVM
GA Availability	Now since 2004	Q2 2014
Supported Hardware	All P6, P7, P7+, P8 Systems	S812L, S822L
Supported Guest OS	AIX, IBM i & Red Hat, SUSE Linux	Red Hat, SUSE & Ubuntu Linux
Workload Mobility	Supports AIX, IBM i & Linux	Linux
Basic Virtualization Management	IVM/HMC/FSM	virt-manager/libvirt/Kimchi
Advanced Virtualization Management	PowerVC/VMControl	PowerVC, Vanilla OpenStack
Admin Type	Power Centric	Linux/x86 Centric
Established Security Track Record on Power	Yes	No
Open Source Hypervisor	No	Yes
Complete Hardware Awareness & Exploitation	Yes	Partial

PowerVM vs. PowerKVM feature comparison

Feature	IBM PowerVM	IBM PowerKVM
Micro partitioning	Yes	Yes
Dynamic Logical Partition	Yes	Partial
SR-IOV support	Yes	No
Shared storage pools	Yes	Yes
Live partition mobility	Yes	Yes
Memory compression	Yes. (Active Memory™ Exploitation)	No. (zswap could be installed manually)
Memory page sharing	Yes (Active Memory Deduplication)	Yes (Kernel Same Page (KSM))
NPIV	Yes	No
License	Proprietary	No
PCI passthrough	Yes	Yes
Supported Machines	All IBM Power Systems	IBM Scale-out systems only
Supported Operating Systems in the guest	AIX, IBM i, Linux	Linux
Different editions	Yes (Standard and Enterprise)	No
Sparse disk storage	Yes (Thin provisioning)	Yes (qcow2 image)
Adding devices to the guest	DLPAR	Hot plug

Terminology comparing KVM and PowerVM

IBM PowerKVM	KVM on x86	IBM PowerKVM
Guest, Virtual Machine	Guest, Virtual Machine	LPAR
Hypervisor, Host	Hypervisor, Host	Hypervisor
Flexible Service Processor (FSP)	Integrated Management Module (IMM)	Flexible Service Processor (FSP)
Open Power Abstraction Layer	UEFI (Unified Extensible Firmware Interface) and BIOS	PowerVM hypervisor driver (pHyp)
KVM Host Userspace (QEMU)	KVM Host Userspace (QEMU)	VIOS (Virtual I/O Server)
Kimchi and virsh	Kimchi and virsh	HMC and IVM (Integrated Virtualization Manager)
IPMI	IPMI	HMC
Kernel Same-page Merge (KSM)	Kernel Same-page Merge (KSM)	Active Memory Deduplication
Zswap	zswap	Active Memory Expansion (AME)
SLOF	SeaBIOS	Open Firmware, SMS
Qcow2, raw and other image formats	Qcow2, raw and other image formats	Proprietary
Preboot eXecution Environment (PXE)	Preboot eXecution Environment (PXE)	BOOTP and TFTP, NIM
Virtio drivers, ibmvscsi and ibmveth	Virtio drivers	Ibmvscsi, ibmveth
Hot plug	Hot plug	DLPAR
VNC	VNC	VNC

IBM PowerKVM webinar, July 16, 2014

PowerKVM exploits POWER8 Micro-Threading

Traditional PowerVM and PowerKVM Dispatches the complete core to the VM

CPU Core

VM1

SMT1-8
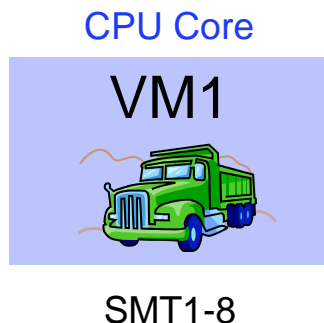
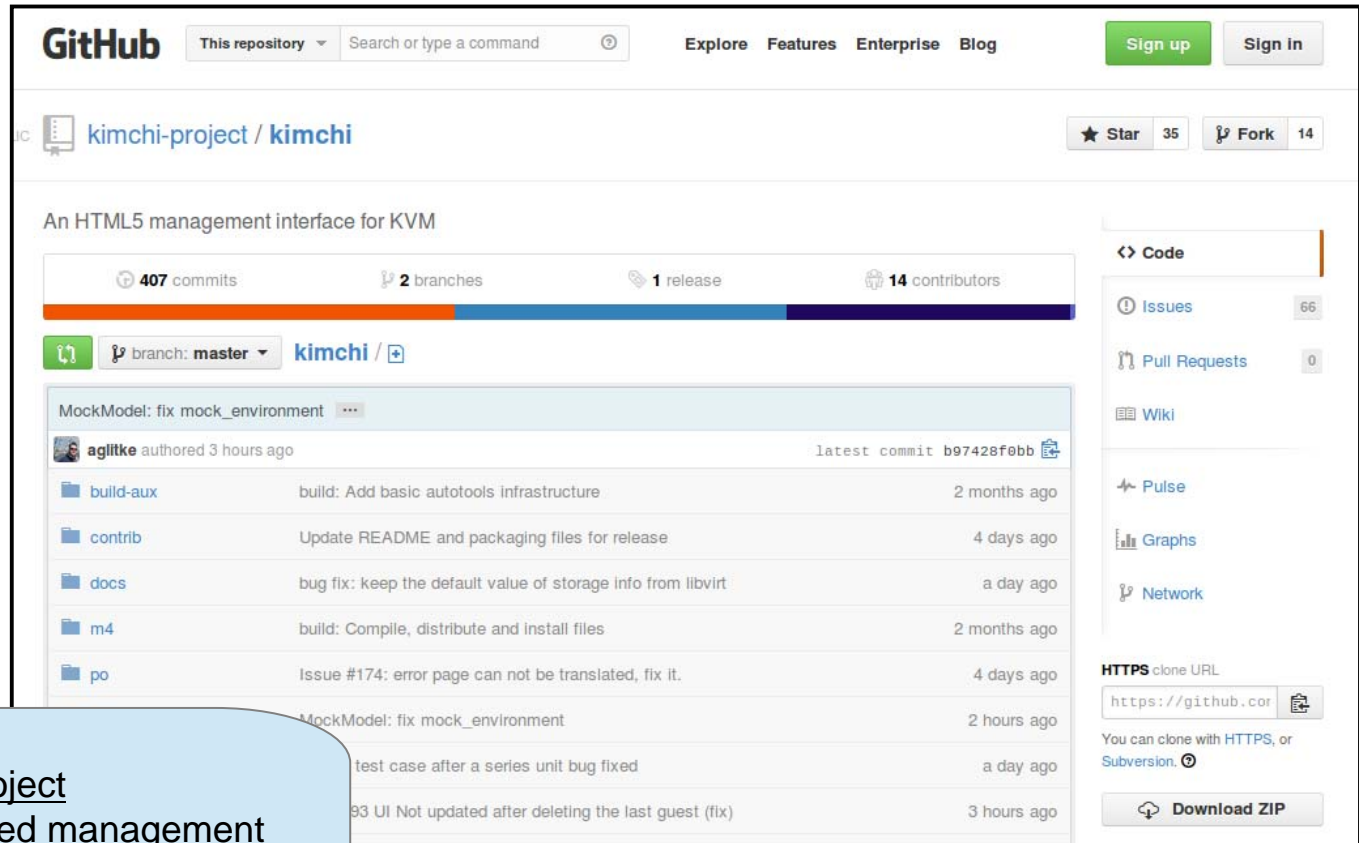**PowerKVM** with **Micro-Threading** Dispatches Multiple VMs on a single core at the **same time**.

CPU Core 4/1 Division

| VM1 | VM2 | VM3 | VM4 |

KVM

SMT1-2
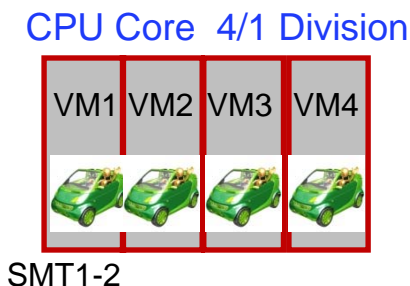
Good for many small VMs / Workloads. Enabled with the PowerKVM ppc64_cpu command. 4/1 Division is only option initially.

Project Kimchi – an emerging open source KVM management tool

Kimchi Project
Kimchi is an HTML5 based management tool for KVM. It is designed to make it as easy as possible to get started with KVM and create your first guest.

More information at https://github.com/kimchi-project/kimchi

IBM PowerKVM webinar, July 16, 2014

More Information on PowerKVM

- IBM **developerWorks** : [The Linux on Power Community – PowerKVM](#)

- IBM **Redbooks** : [IBM PowerKVM Configuration and Use](#)

- IBM Knowledge Center: [IBM PowerKVM](#)

- YouTube Videos:
  – [Demonstration of IBM PowerVC Standard V1.2.1 managing PowerKVM based POWER8 server](#)
  – [PowerKVM1 video - IBM Client Center Montpellier](#)
  – [KVM Forum 2013: Developments in KVM on Power by Paul Mackerras](#)
    • PDF: http://www.linux-kvm.org/wiki/images/7/70/Kvm-forum-2013-Mackerras.pdf

PowerKVM Demo

IBM PowerKVM webinar, July 16, 2014

PowerKVM Demo Setup

IBM "cloud"

9.X.X.X
network

IBM Munich

9.X.X.X network

10.0.X.X internal
network

- Connection to internal 10.0.X.X network in Munich with
  - **XRDP**
  - **SSH tunnel**
  - **VNC**

My KVM environment

Openfiler iSCSI storage

linuxppc NFS storage

IBM Munich 10.X.X.X network

DS4700 SAN storage

x3850-1
24 cores, 128 GB

x3850-2
24 cores, 128 GB

powerkvm
16 cores, 128 GB

x3650
2 cores, 30 GB
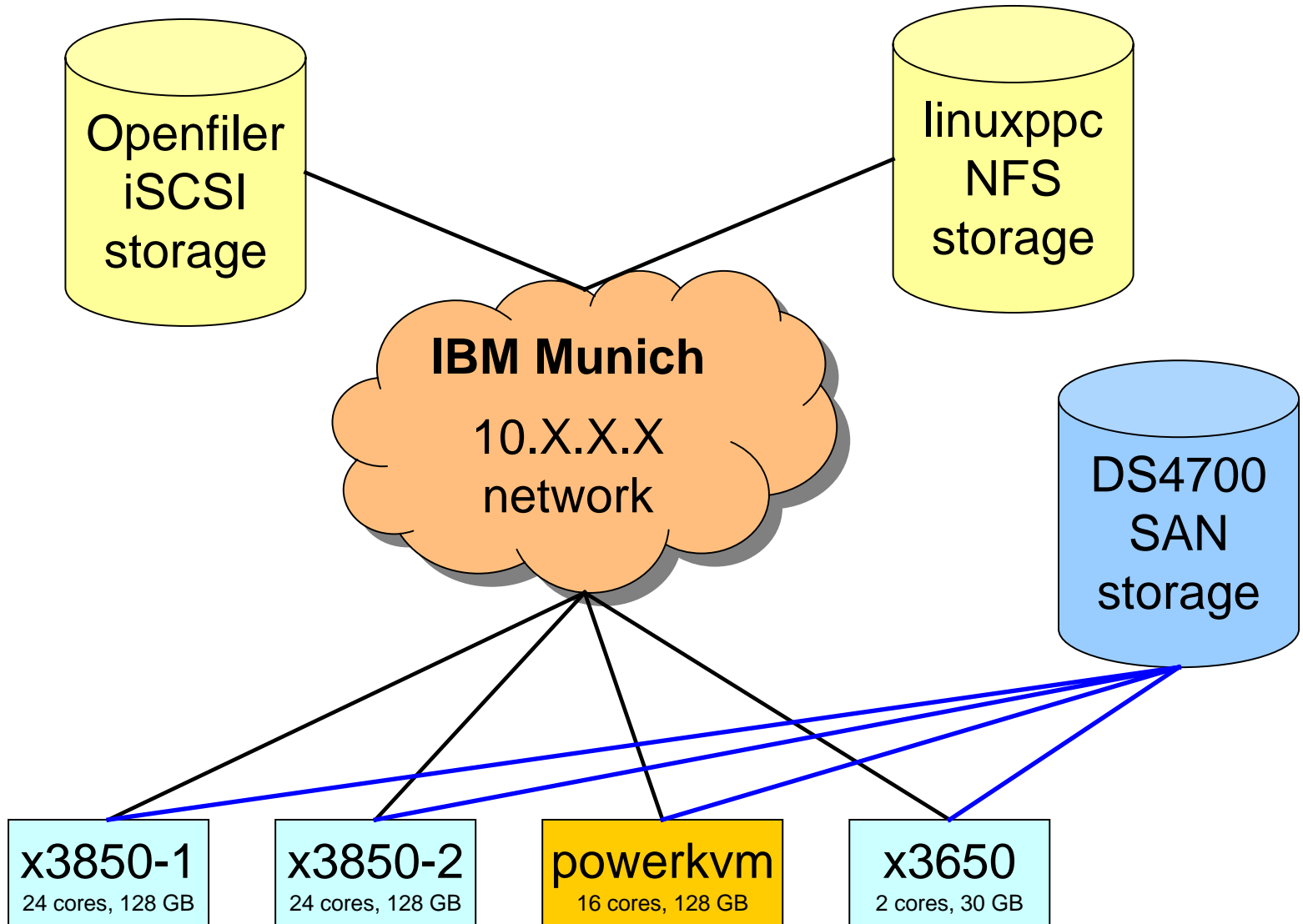
Special notices

This document was developed for IBM offerings in the United States as of the date of publication. IBM may not make these offerings available in other countries, and the information is subject to change without notice. Consult your local IBM business contact for information on the IBM offerings available in your area.

Information in this document concerning non-IBM products was obtained from the suppliers of these products or other public sources. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. Send license inquires, in writing, to IBM Director of Licensing, IBM Corporation, New Castle Drive, Armonk, NY 10504-1785 USA.

All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

The information contained in this document has not been submitted to any formal IBM test and is provided "AS IS" with no warranties or guarantees either expressed or implied.

All examples cited or described in this document are presented as illustrations of the manner in which some IBM products can be used and the results that may be achieved. Actual environmental costs and performance characteristics will vary depending on individual client configurations and conditions.

IBM Global Financing offerings are provided through IBM Credit Corporation in the United States and other IBM subsidiaries and divisions worldwide to qualified commercial and government clients. Rates are based on a client's credit rating, financing terms, offering type, equipment type and options, and may vary by country. Other restrictions may apply. Rates and offerings are subject to change, extension or withdrawal without notice.

IBM is not responsible for printing errors in this document that result in pricing or information inaccuracies.

All prices shown are IBM's United States suggested list prices and are subject to change without notice; reseller prices may vary.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

Any performance data contained in this document was determined in a controlled environment. Actual results may vary significantly and are dependent on many factors including system hardware configuration and software design and configuration. Some measurements quoted in this document may have been made on development-level systems. There is no guarantee these measurements will be the same on generally-available systems. Some measurements quoted in this document may have been estimated through extrapolation. Users of this document should verify the applicable data for their specific environment.

Revised October 2010

IBM PowerKVM webinar, July 16, 2014

Special notices (cont.)

IBM, the IBM logo, ibm.com AIX, AIX (logo), AIX 6 (logo), AS/400, Active Memory, BladeCenter, Blue Gene, CacheFlow, ClusterProven, DB2, ESCON, i5/OS, i5/OS (logo), IBM Business Partner (logo), IntelliStation, LoadLeveler, Lotus, Lotus Notes, Notes, Operating System/400, OS/400, PartnerLink, PartnerWorld, PowerPC, pSeries, Rational, RISC System/6000, RS/6000, THINK, Tivoli, Tivoli (logo), Tivoli Management Environment, WebSphere, xSeries, z/OS, zSeries, AIX 5L, Chiphopper, Chipkill, Cloudscape, DB2 Universal Database, DS4000, DS6000, DS8000, EnergyScale, Enterprise Workload Manager, General Purpose File System, , GPFS, HACMP, HACMP/6000, HASM, IBM Systems Director Active Energy Manager, iSeries, Micro-Partitioning, POWER, PowerExecutive, PowerVM, PowerVM (logo), PowerHA, Power Architecture, Power Everywhere, Power Family, POWER Hypervisor, Power Systems, Power Systems (logo), Power Systems Software, Power Systems Software (logo), POWER2, POWER3, POWER4, POWER4+, POWER5, POWER5+, POWER6, POWER7, pureScale, System i, System p, System p5, System Storage, System z, Tivoli Enterprise, TME 10, TurboCore, Workload Partitions Manager and X-Architecture are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml

The Power Architecture and Power.org wordmarks and the Power and Power.org logos and related marks are trademarks and service marks licensed by Power.org.
UNIX is a registered trademark of The Open Group in the United States, other countries or both.
Linux is a registered trademark of Linus Torvalds in the United States, other countries or both.
Microsoft, Windows and the Windows logo are registered trademarks of Microsoft Corporation in the United States, other countries or both.
Intel, Itanium, Pentium are registered trademarks and Xeon is a trademark of Intel Corporation or its subsidiaries in the United States, other countries or both.
AMD Opteron is a trademark of Advanced Micro Devices, Inc.
Java and all Java-based trademarks and logos are trademarks of Sun Microsystems, Inc. in the United States, other countries or both.
TPC-C and TPC-H are trademarks of the Transaction Performance Processing Council (TPPC).
SPECint, SPECfp, SPECjbb, SPECweb, SPECjAppServer, SPEC OMP, SPECviewperf, SPECapc, SPEChpc, SPECjvm, SPECmail, SPECimap and SPECsfs are trademarks of the Standard Performance Evaluation Corp (SPEC).
NetBench is a registered trademark of Ziff Davis Media in the United States, other countries or both.
AltiVec is a trademark of Freescale Semiconductor, Inc.
Cell Broadband Engine is a trademark of Sony Computer Entertainment Inc.
InfiniBand, InfiniBand Trade Association and the InfiniBand design marks are trademarks and/or service marks of the InfiniBand Trade Association.
Other company, product and service names may be trademarks or service marks of others.

Revised October 2010