

Hardware Deep Dive

April / May Announcements Scale out systems

Alberto Damigella

IBM Power systems presales tech support

Original slides:

Mark Olson

WW Offering manager, Power Systems



Scale-out Systems built with POWER8

- Leadership Data / Analytics performance in consumable Scale out form factor; Foundation for Big Data / Analytic solutions
- Portfolio is complemented by a full commitment to support an open stack of software
 - Ubuntu, SUSE, RedHat, PowerKVM, and Open Stack
- Offerings to compete head to head with x86 Linux



S812L	S822L	S822	S814	S824L	S824
<ul style="list-style-type: none"> • 1-socket, 2U • Linux only 	<ul style="list-style-type: none"> • 2-socket, 2U • Up to 24 cores • 1 TB memory • 9 PCI Gen3 slot • Linux only • NEBs option 	<ul style="list-style-type: none"> • 2-socket, 2U • Up to 20 cores • 1 TB memory • 9 PCIe Gen 3 • AIX & Linux • PowerVM only • NEBs option 	<ul style="list-style-type: none"> • 1-socket, 4U • Up to 8 cores • 512 GB memory • 7 PCIe Gen 3 • AIX, IBM i, Linux • PowerVM only • 4core/P05 (IBM i) 	<ul style="list-style-type: none"> • 2-socket, 4U • Up to 24 cores • Linux • NVidia GPU 	<ul style="list-style-type: none"> • 2-socket, 4U • Up to 24 cores • 2 TB memory • 11 PCIe Gen 3 • AIX, IBM i, Linux • PowerVM only

<p>New in 2Q15</p>	<ul style="list-style-type: none"> • 8-core, 4.15 GHz option • SOD: Water cooling 	<ul style="list-style-type: none"> • 8-core, 4.15 GHz option • SOD: Water cooling 	<ul style="list-style-type: none"> • 1 TB memory • 110 v power 	<ul style="list-style-type: none"> • Full offering (GPU not required) • 2TB of memory • 8/16-core 4.15GHz and 24-core 3.52GHz option • PowerKVM (2H15) • Nvidia K80 GPU 	
---------------------------	---	---	--	--	--

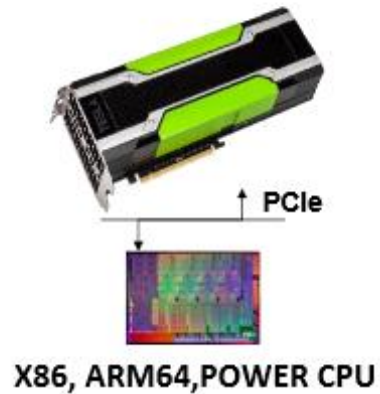
Announce / Availability

Scale-out announcements	Annc	General Available
110V power for rack mount S814	28 Apr 2015	now
4.1 GHz for S822 and S822L	28 Apr 2015	5 June 2015
Bigger memory for S814 and S824L	28 Apr 2015	5 June 2015
PCIe Gen3 I/O drawers	28 Apr 2015	5 June 2015
S814 Tower ← → Rack	28 Apr 2015	5 June 2015
S824L general non-GPU configs	28 Apr 2015	19 June 2015
K80 GPU on Sk24L	June	June

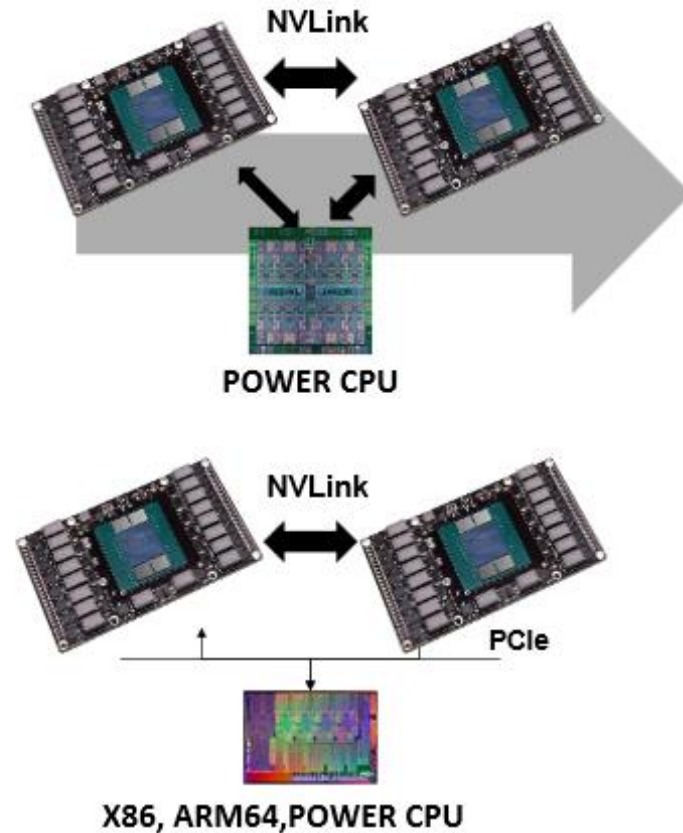
NVLink GPU Interconnect



KEPLER GPU



PASCAL GPU



Announce / Availability

Other announcements	Annc	General Available
4-port 10GbE adapter	28 Apr 2015	5 June 2015
2-port 10GbE NIC&RoCE adapter	28 Apr 2015	5 June 2015 Except some countries*
2-port 56Gb FDR IB adapter	28 Apr 2015	5 June 2015 Except some countries*
1.8TB 10k rpm HDD	28 Apr 2015	5 Jun 2015
4k block 6.TB & 1.2TB 10k rpm HDD	28 Apr 2015	5 June 2015

SAS SFF-3 Supported Options as of May 2015

SFF-3 HDD	Block size	Formatted with 512 or 4096 byte sectors	Formatted with 528 or 4224 byte sectors
10k	5xx	300 GB #ESDR	283 GB
10k	5xx	600 GB #ESD5	571 GB
10k	4k	600 GB #ESF5	571 GB
10k	5xx	1.2 TB #ESD9	1.1 TB
10k	4k	1.2 TB #ESF9	1.1 TB
10k	4k	1.8 TB #ESFV	1.7 TB
15k	5xx	146 GB #ESDT	139 GB
15k	5xx	300 GB #ESDB	283 GB
15k	4k	300 GB #ESFB	283 GB
15k	4k	600 GB #ESFF	571 GB
15k	5xx	600 GB #ESDF	571 GB

SFF-3 HDD shipped from IBM formatted in 528 or 4224 byte sectors. They can be reformatted to 512 or 4096 by the client if the extra protection is not desired.

4k drives (HDD or SSD) can NOT be reformatted to 5xx drives.

4k drives and 5xx drives can NOT be mixed on the same array. True for both HDD and SSD.

SFF-3 SSD	Block size	For AIX/Linux/ VIOS rules
eMLC3	528 byte	387 GB #ES0L
eMLC3	528 byte	775 GB #ES0N
eMLC3	4224 byte	387 GB #ES0U
eMLC3	4224 byte	775 GB #ES0W

The above features use SFF-3 or Gen3 carriers/trays and can be placed in the SFF SAS bays of the POWER8 servers

PCIe Gen3 I/O Expansion Drawer

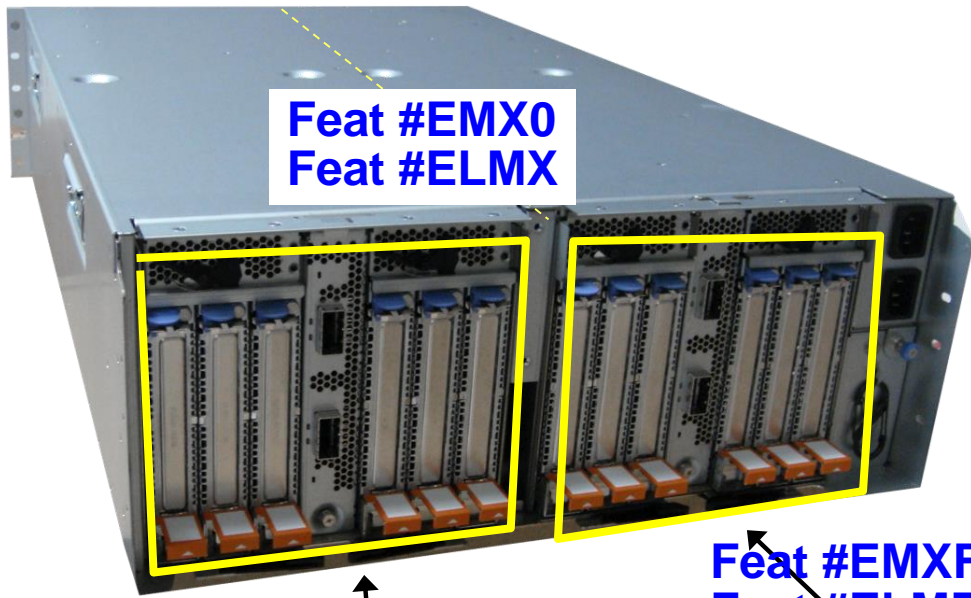


PCIe Gen3 I/O Expansion Drawer for Scale-out



Front view

- 6 or 12 PCIe Gen3 slots
- 4U drawer
- Full high PCIe slots
- Hot plug PCIe slots
- Modules not hot plug



Rear view

Feat #EMX0
Feat #ELMX

Feat #EMXF
Feat #ELMF

Feat #EMXF
Feat #ELMF

Fan-out Module
6 PCIe Gen3 Slots
4 x8 & 2 x16

Fan-out Module
6 PCIe Gen3 Slots
4 x8 & 2 x16

#EMX0/EMXF for S814, S824, S822
#ELMX/ELMF for S824L, S812L, S822L

Use same Blind Swap Cassette (BSC) as used in #5802/5877/5803/5873 I/O drawer

4U System to PCIe Gen3 I/O Drawer (S814, S824, S824L, E850)

PCIe3 Optical Cable Adapter (#EJ08) **single-wide card**

- One #EJ08 per fan-out module
- Can be in any of system unit's x16 PCIe slots

AOC cable pairs

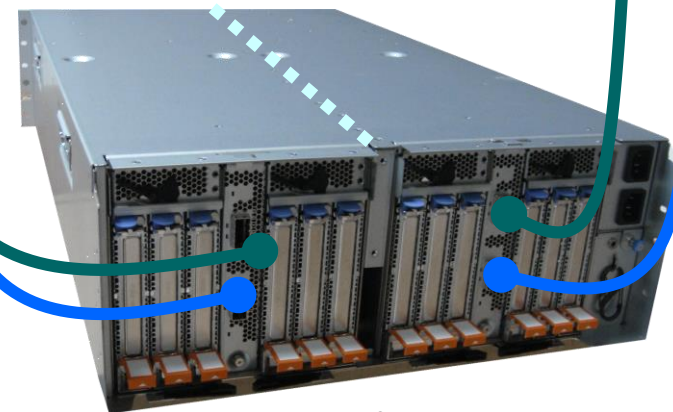
- #ECC7 3 meter length
- #ECC8 10 meter length
- One feat code ships two identical cables
- Connect top CXP port of #EJ08 to top CXP port of fan-out module. Likewise connect bottom port to bottom port. Do NOT reverse !!
- Do NOT mix lengths of AOC cables for the same fan-out module
- Do NOT cross cables connecting one fan-out module to two different #EJ08 adapters
- Each cable has two electrical-to-optical converters in the transceiver/CXP module at each cable end. 12-lane optical cables are used for bandwidth and for incremental redundancy.



Conceptual diagram. Does not depict actual x16 slot locations
Rear of 4U2S picture shown above

Active Optical
Cable Pair

Active Optical
Cable Pair



fan-out module

fan-out module

2U Scale-out System to PCIe Gen3 I/O Drawer

PCIe3 Optical Cable Adapter (#EJ05) **double-wide card**

- One #EJ05 per fan-out module
- Can be in any of node's x16 PCIe slots – will cover up another slot

AOC cable pairs

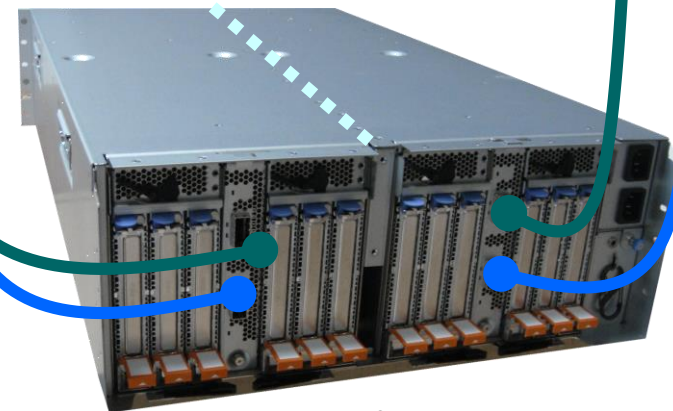
- #ECC7 3 meter length
- #ECC8 10 meter length
- One feat code ships two identical cables
- Connect top CXP port of #EJ05 to top CXP port of fan-out module. Likewise connect bottom port to bottom port. Do NOT reverse !!
- Do NOT mix lengths of AOC cables for the same fan-out module
- Do NOT cross cables connecting one fan-out module to two different #EJ05 adapters
- Each cable has two electrical-to-optical converters in the transceiver/CXP module at each cable end. 12-lane optical cables are used for bandwidth and for incremental redundancy.



Conceptual diagram of two socket. Does not depict actual x16 slot locations
With one processor: use C6 (Plus overlap)
With two processors: Use C3 ,C6 (plus overlap)

Active Optical
Cable Pair

Active Optical
Cable Pair



fan-out module

fan-out module

PCIe Gen3 I/O expansion drawer PCIe Slot Math

	2U 1-socket	2U 2-socket One filled	2U 2-socket Two filled	4U 1-socket ***	4U 2-socket One filled	4U 2-socket Two filled
	S812L	S822	S822, S822L	S814	S824	S824, S824L*
PCIe slots in system unit	6	6	9	7	7	11
x16 slots in system unit	2	2	4	2	2	4
Max PCIe Gen3 I/O Drawer	½	½	1	1	1	2
PCIe Gen3 Drawer slots	6	6	12	12	12	24
PCIe slots used for Optical Cable Adapter	2**	2**	4**	2	2	4
Total Max PCIe slots	6-2+6 = 10	6-2+6 = 10	9-4+12 = 17	7-2+12 = 17	7-2+12 = 17	11-4+24 = 31

Nearly 2x more slots

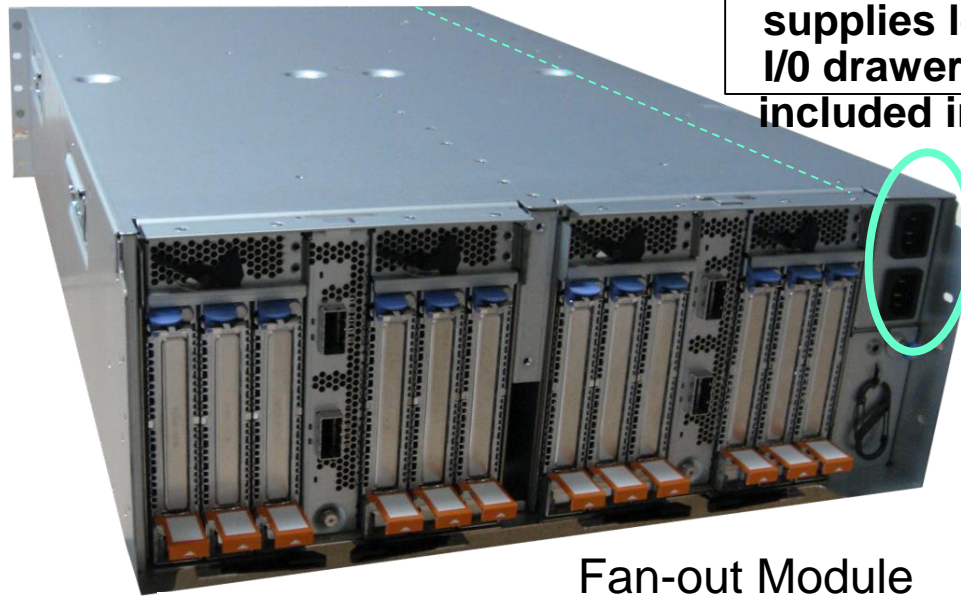
Over 2x more

nearly 3x

* If S824L using a GPU, can not have I/O drawer
 ** 2U uses double-wide Optical Cable Adapter which uses two PCIe slots per fan-out module. Because of x16 slot location, can only use half of the x16 slots for attaching an I/O drawer.
 *** requires a 6-core or 8-core server. 4-core server doesn't support I/O drawer

PCIe Gen3 I/O Drawer Electrical Power

Redundant, hot plug power supplies located in front of PCIe3 I/O drawer (no separate feat code, included in PCIe Gen3 I/O drawer)



Feat #EMXA
AC Power Supply Conduit

- Provides connector for two power cords (separate feature codes)
- Carries power from back of drawer to front where power supplies are located

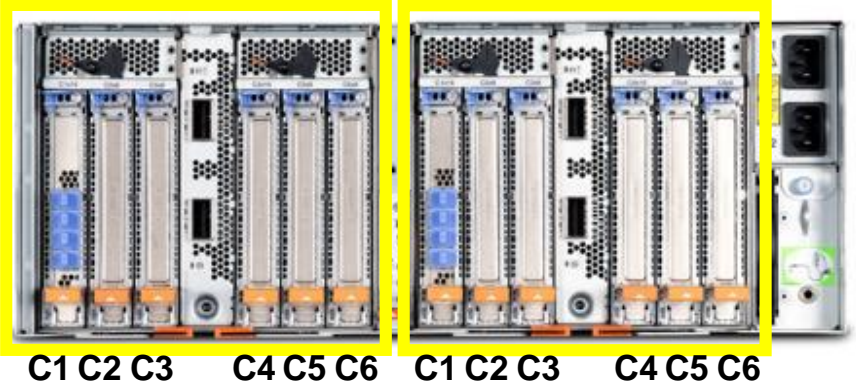
Rear view

Fan-out Module

Fan-out Module

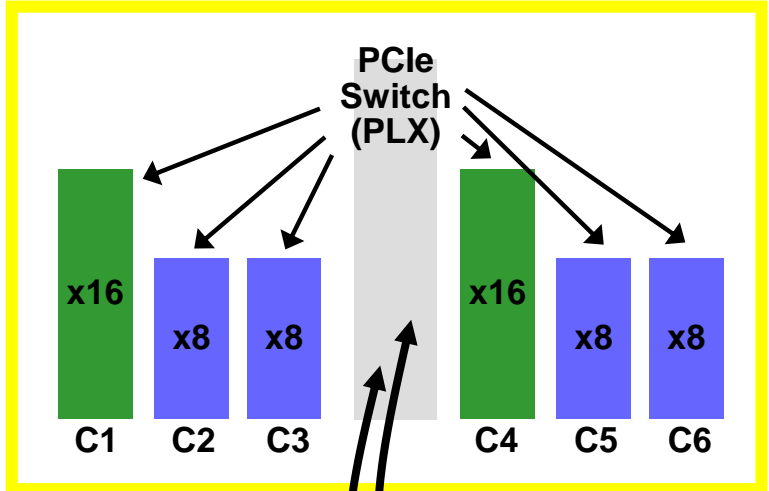
PCIe3 6-slot Fanout Modules

Rear view



Fan-out Module
6 PCIe Gen3 Slots
4 x8 & 2 x16

Fan-out Module
6 PCIe Gen3 Slots
4 x8 & 2 x16



Block diagram of one fanout module

AOC cables to x16 slot in system

For optimal thermal plugging (and ignoring any specific adapter slot requirements / limitations), first fill slots C2 and C5. Then fill C3 and C6. Finally fill C1 and C4

Power S814, S824, S822 PCIe Adapters Supported
as of June 2015 (page 3) in PCIe Gen3 I/O Drawer

Infiniband (IB)	2-port QDR IB SR optical	#5285
Graphics	2D graphics for general use	#5748
Graphics	3D graphics for RHEL7	#EC42
Encryption	Crypto Coprocessor 4765-001	#EJ28
USB	4-port USB-3	#EC46

New PCIe Adapters



Three New PCIe3 Adapters for POWER8 servers

4-port 10Gb Ethernet Adapter

- Save PCIe slots. Use instead of 2-port 10Gb adapters
- Supported on POWER8 servers except 2U system units
- SR-IOV
- AIX, IBM i, Linux, VIOS note -IBM i native (VIOS optional)

2-port 10Gb Ethernet NIC+RoCE Adapter

- Refresh of existing 2-port 10Gb NIC+RoCE adapter
- AIX, Linux, VIOS (IBM i via VIOS)
- Supports AIX NIM and Linux Network Install

2-port 56Gb FDR InfiniBand Adapter x16

- Bigger bandwidth, great server interconnectivity
- Supported in system units of POWER8 Scale-out servers in x16 slot
- Linux only

* preliminary pricing

New 4-Port 10Gb Ethernet Adapter

1st 4-port ALL 10Gb PCIe adapter

- Ethernet NIC
- Save PCIe slots vs using two 2-port 10GbE cards
- Supports SR-IOV for additional virtualization
- IBM i native support (VIOS optional)
- Supports AIX NIM and Linux Network Install

Four feature codes:

	LP for 2U system units	LP+* for E870/E880 system node	Full high tail stock for 4U servers and for PCIe Gen3 I/O Drawer
SR Optical Fiber	n/a	#EN16	#EN15
Copper twinax	n/a	#EN18	#EN17

*physically too large for 2U server's low profile slot

2-port 56Gb FDR InfiniBand Adapter x16

InfiniBand FDR

- Bigger bandwidth, great server interconnectivity compared to predecessor 40Gb EDR IB adapter
- Two QSFP+ ports, each up to 56Gb
- Order transceivers separately
- Cabling: Use either FDR IB or 40GbE cabling
- Supported in system units of POWER8 Scale-out servers in x16 slot
- Linux only

Four feature codes:	Low profile tail stock	Full high tail stock
Linux-only servers	#EL3D	#EL50
Multi-OS servers (running Linux)	#EC32	#EC33

THANKS



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.

Special notices (cont.)

IBM, the IBM logo, ibm.com AIX, AIX (logo), AIX 6 (logo), AS/400, BladeCenter, Blue Gene, 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, POWER6+, System i, System p, System p5, System Storage, System z, Tivoli Enterprise, TME 10, 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, SPECcapc, SPECchpc, 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.