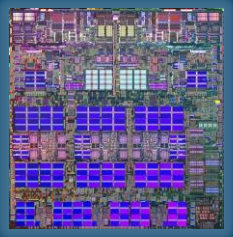
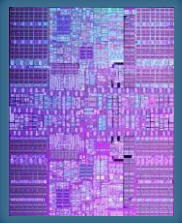
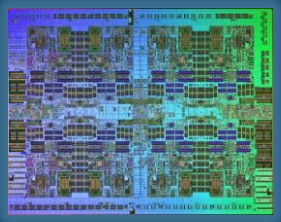
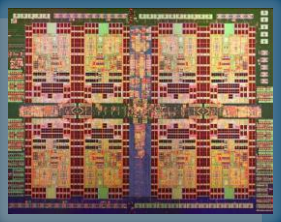
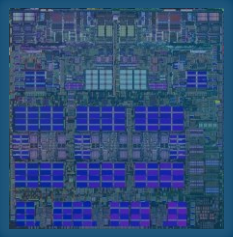
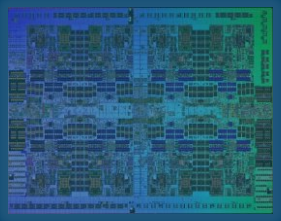
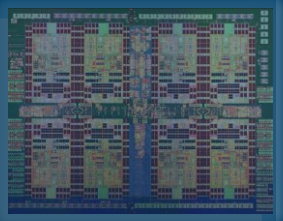
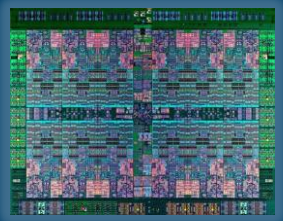
POWER8
Open Innovation for Big Data & Cloud

Jeff Stuecheli, PhD
IBM Power Systems
IBM Systems & Technology Group Development

History...

	POWER5 *2004*	POWER6 *2007*	POWER7 *2010*	POWER7+ *2012*
				
Technology	130nm SOI	65nm SOI	45nm SOI eDRAM	32nm SOI eDRAM
Compute				
Cores	2	2	8	8
Threads	SMT2	SMT2	SMT4	SMT4
Caching				
On-chip	1.9MB	8MB	2 + 32MB	2 + 80MB
Off-chip	36MB	32MB	None	None
Die Bandwidth				
Sust. Mem.	15GB/s	30GB/s	100GB/s	100GB/s
Peak I/O	6GB/s	20GB/s	40GB/s	40GB/s

History...

	POWER5 *2004*	POWER6 *2007*	POWER7 *2010*	POWER7+ *2012*	POWER8
					
Technology	130nm SOI	65nm SOI	45nm SOI eDRAM	32nm SOI eDRAM	
Compute					
Cores	2	2	8	8	***Today's Topic***
Threads	SMT2	SMT2	SMT4	SMT4	
Caching					
On-chip	1.9MB	8MB	2 + 32MB	2 + 80MB	
Off-chip	36MB	32MB	None	None	
Die Bandwidth					
Sust. Mem.	15GB/s	30GB/s	100GB/s	100GB/s	
Peak I/O	6GB/s	20GB/s	40GB/s	40GB/s	

POWER8 Vision

Leadership Performance

- Increase core throughput at single thread, SMT2, SMT4, and SMT8 level
- Large step in per socket performance
- Enable more robust multi-socket scaling

System Innovation

- Higher capacity cache hierarchy and highly threaded processor
- Enhanced memory bandwidth, capacity, and expansion
- Dynamic code optimization
- Hardware-accelerated virtual memory management

Open System Innovation

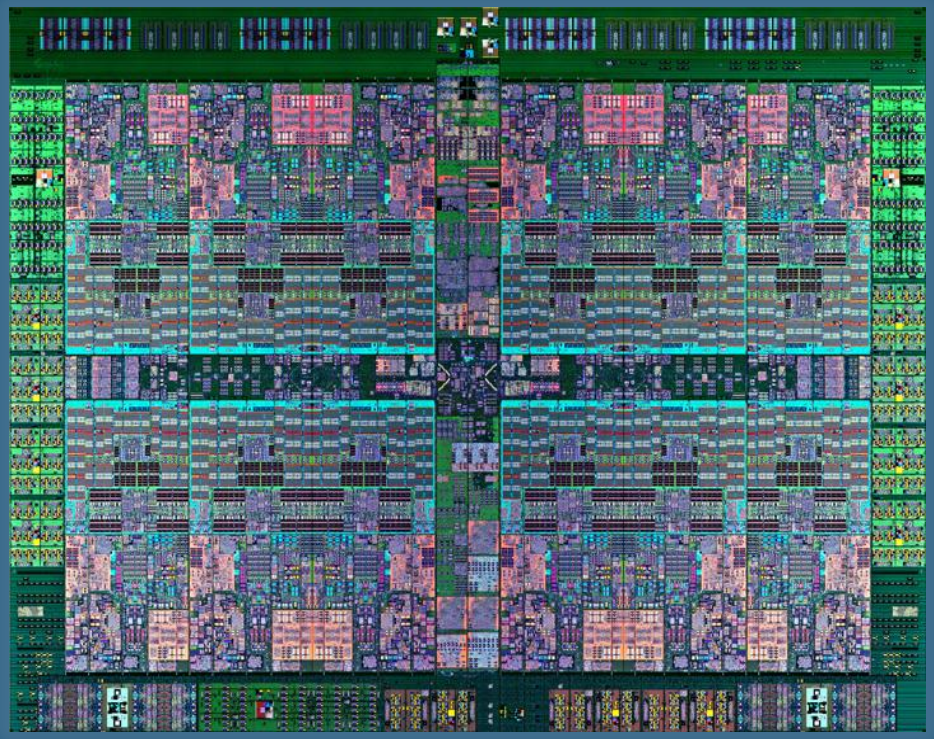
- Coherent Accelerator Processor Interface (CAPI)
- Agnostic Memory interface
- Open system software

Optimize Analytics & Big Data

Enhance Cloud Efficiency

Enable Open Innovation on POWER

POWER8 Processor



POWER8 Processor

Technology

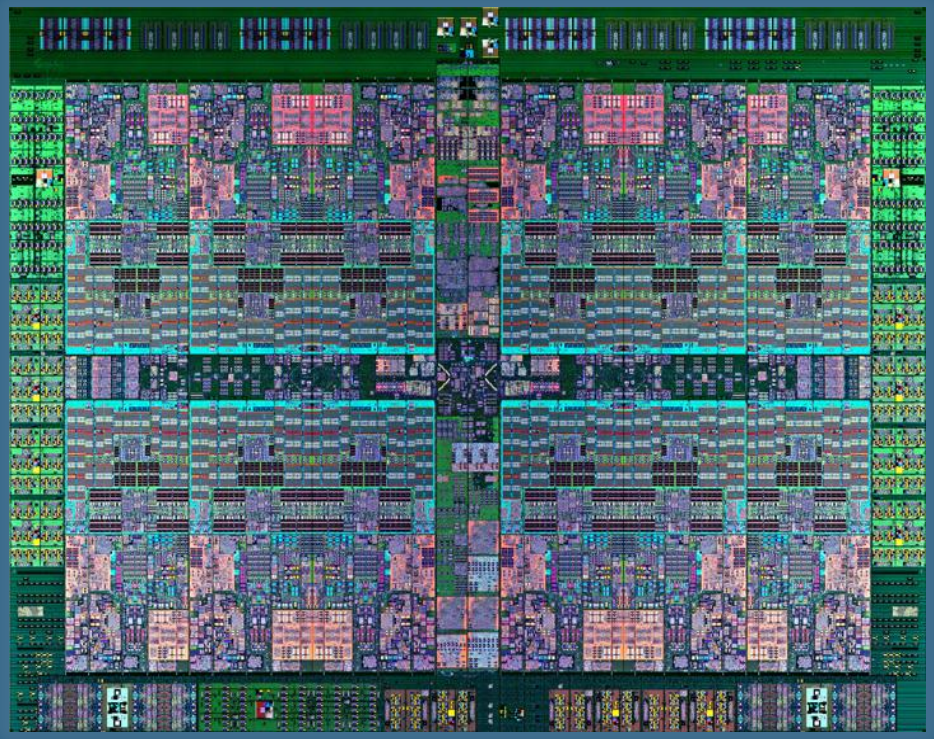
- 22nm SOI, eDRAM, 15 ML 650mm2

Cores

- 12 cores (SMT8)
- 8 dispatch, 10 issue, 16 exec pipe
- 2X internal data flows/queues
- Enhanced prefetching
- 64K data cache, 32K instruction cache

Accelerators

- Crypto & memory expansion
- Transactional Memory
- VMM assist
- Data Move / VM Mobility



Caches

- 512 KB SRAM L2 / core
- 96 MB eDRAM shared L3
- Up to 128 MB eDRAM L4 (off-chip)

Memory per Die

- Up to 230 GB/s sustained bandwidth

Die Bus Interfaces

- Durable open memory attach interface
- Integrated PCIe Gen3
- SMP Interconnect
- CAPI (Coherent Accelerator Processor Interface)

Energy Management

- On-chip Power Management Micro-controller
- Integrated Per-core VRM
- Critical Path Monitors

POWER8 Processor

Technology

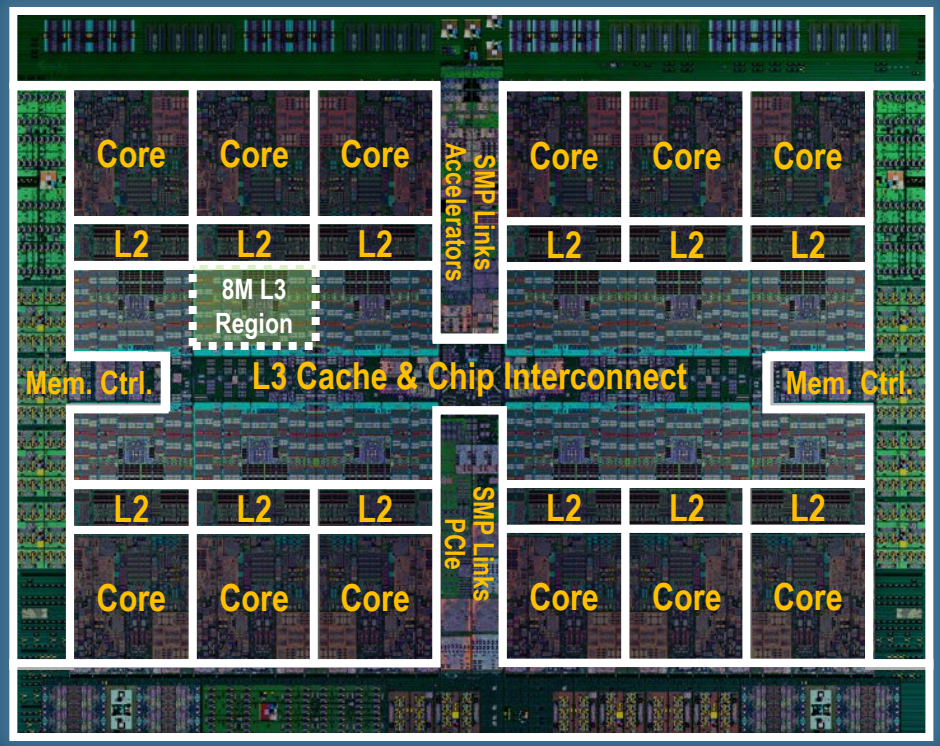
- 22nm SOI, eDRAM, 15 ML 650mm2

Cores

- 12 cores (SMT8)
- 8 dispatch, 10 issue, 16 exec pipe
- 2X internal data flows/queues
- Enhanced prefetching
- 64K data cache, 32K instruction cache

Accelerators

- Crypto & memory expansion
- Transactional Memory
- VMM assist
- Data Move / VM Mobility



Energy Management

- On-chip Power Management Micro-controller
- Integrated Per-core VRM
- Critical Path Monitors

Caches

- 512 KB SRAM L2 / core
- 96 MB eDRAM shared L3
- Up to 128 MB eDRAM L4 (off-chip)

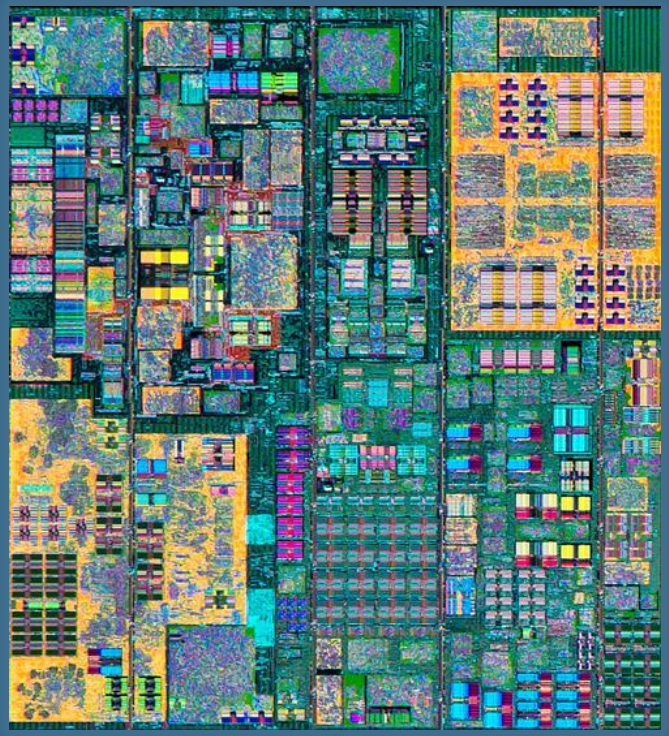
Memory per Die

- Up to 230 GB/s sustained bandwidth

Die Bus Interfaces

- Durable open memory attach interface
- Integrated PCIe Gen3
- SMP Interconnect
- CAPI (Coherent Accelerator Processor Interface)

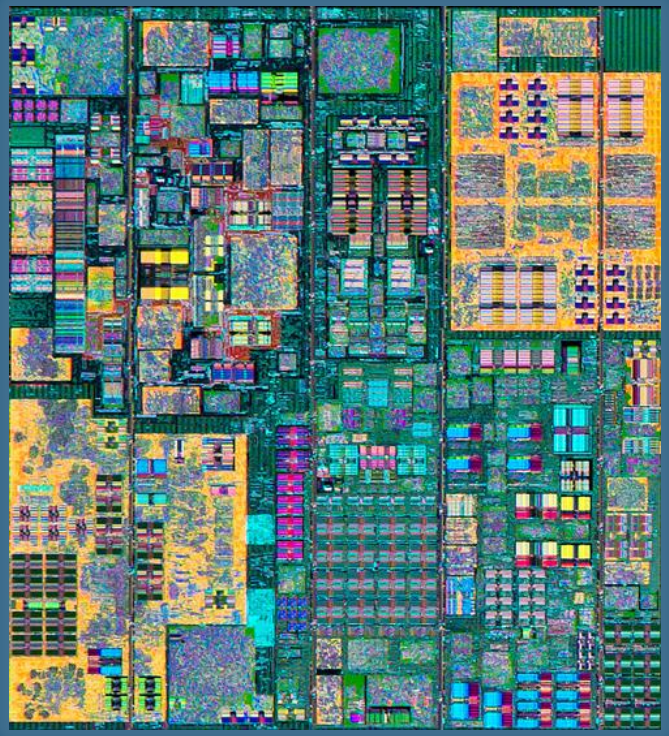
POWER8 Core



POWER8 Core

Execution Improvement vs. POWER7

- SMT4 → SMT8
- 8 dispatch
- 10 issue
- 16 execution pipes:
  - 2 FXU, 2 LSU, 2 LU, 4 FPU, 2 VMX, 1 Crypto, 1 DFU, 1 CR, 1 BR
- Larger Issue queues (4 x 16-entry)
- Larger global completion, Load/Store reorder
- Improved branch prediction
- Improved unaligned storage access



Larger Caching Structures vs. POWER7

- 2x L1 data cache (64 KB)
- 2x outstanding data cache misses
- 4x translation Cache

Wider Load/Store

- 32B → 64B L2 to L1 data bus
- 2x data cache to execution dataflow

Enhanced Prefetch

- Instruction speculation awareness
- Data prefetch depth awareness
- Adaptive bandwidth awareness
- Topology awareness

Core Performance vs. POWER7

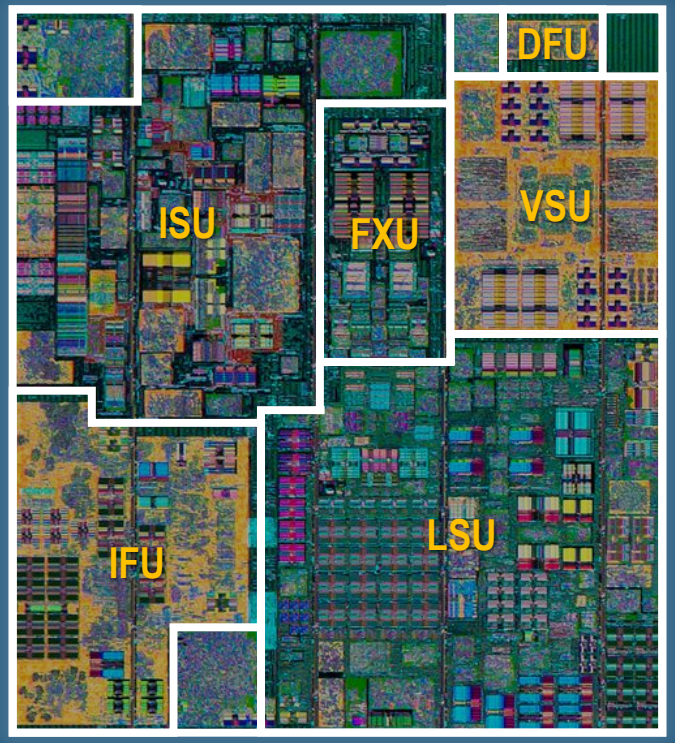
~1.6x Single Thread

~2x Max SMT

POWER8 Core

Execution Improvement vs. POWER7

- SMT4 → SMT8
- 8 dispatch
- 10 issue
- 16 execution pipes:
  - 2 FXU, 2 LSU, 2 LU, 4 FPU, 2 VMX, 1 Crypto, 1 DFU, 1 CR, 1 BR
- Larger Issue queues (4 x 16-entry)
- Larger global completion, Load/Store reorder
- Improved branch prediction
- Improved unaligned storage access



Larger Caching Structures vs. POWER7

- 2x L1 data cache (64 KB)
- 2x outstanding data cache misses
- 4x translation Cache

Wider Load/Store

- 32B → 64B L2 to L1 data bus
- 2x data cache to execution dataflow

Enhanced Prefetch

- Instruction speculation awareness
- Data prefetch depth awareness
- Adaptive bandwidth awareness
- Topology awareness

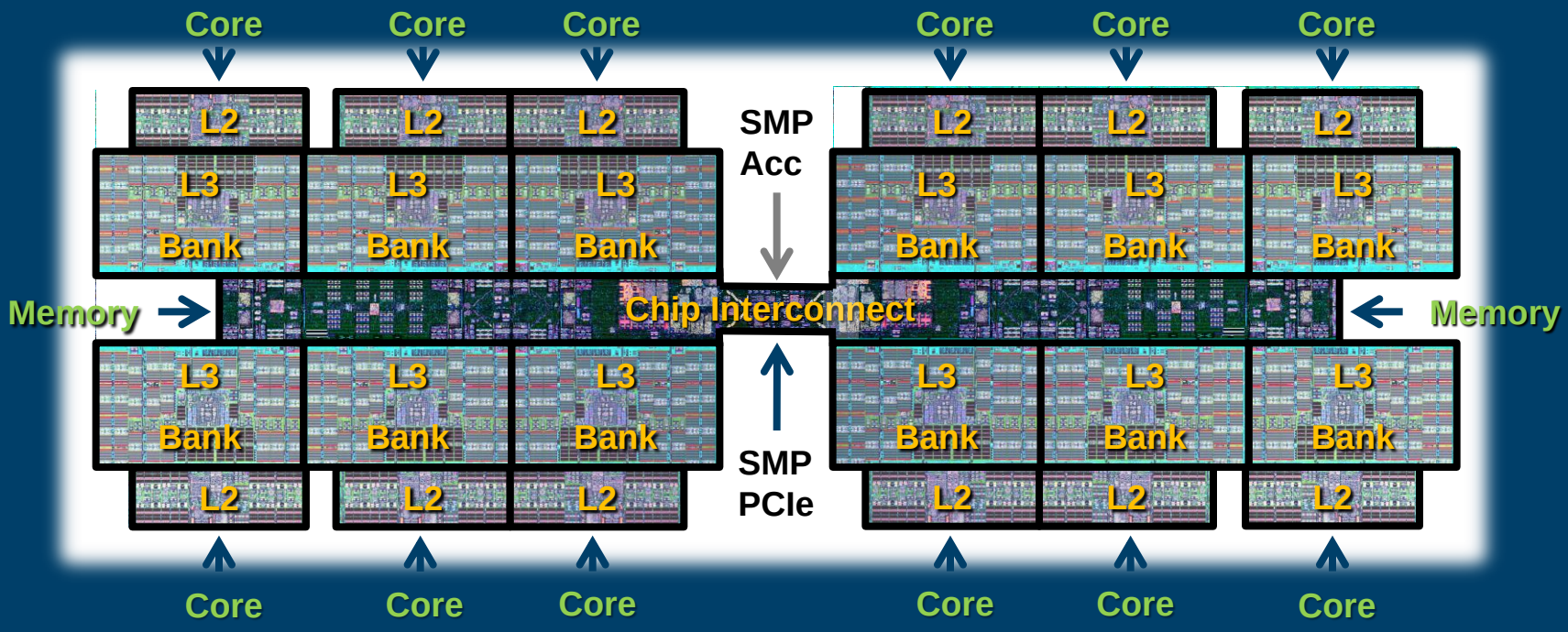
Core Performance vs. POWER7

~1.6x Single Thread

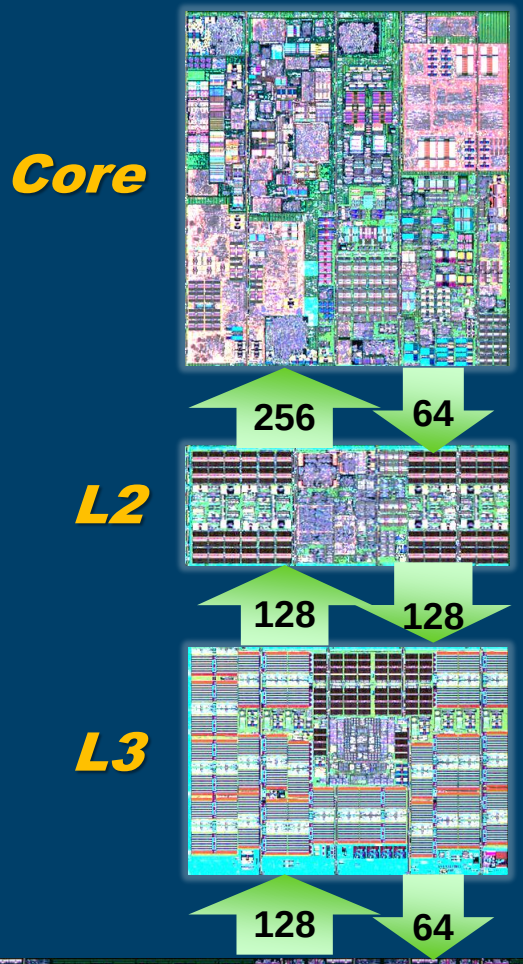
~2x Max SMT

POWER8 On Chip Caches

- **L2: 512 KB 8 way per core**
- **L3: 96 MB (12 x 8 MB 8 way Bank)**
- **"NUCA" Cache policy (**Non-Uniform Cache Architecture)
  - Scalable bandwidth and latency
  - Migrate "hot" lines to local L2, then local L3 (replicate L2 contained footprint)
- **Chip Interconnect: 150 GB/sec x 12 segments per direction = 3.6 TB/sec**

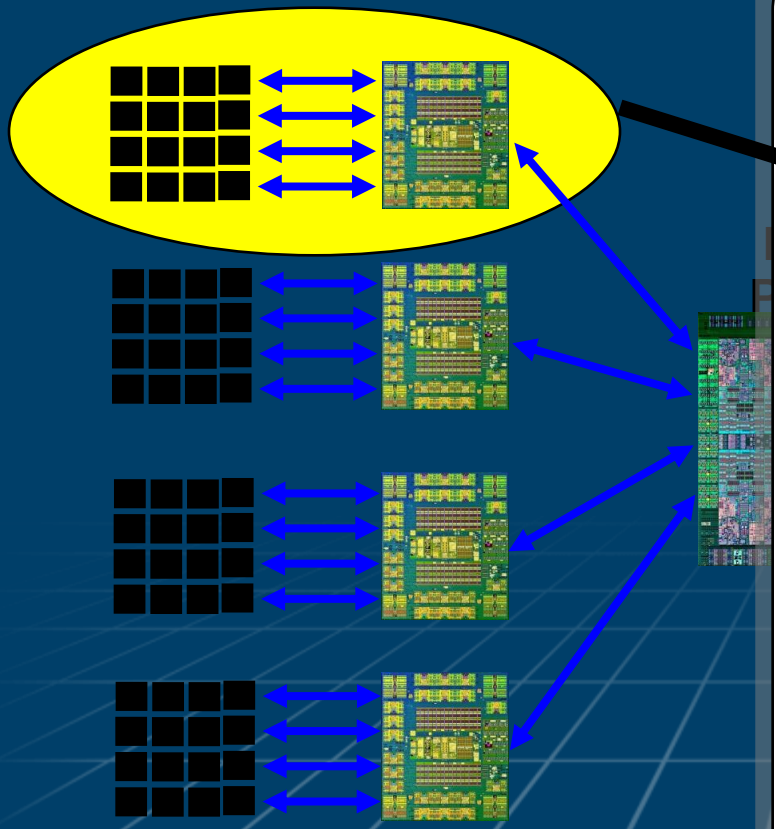


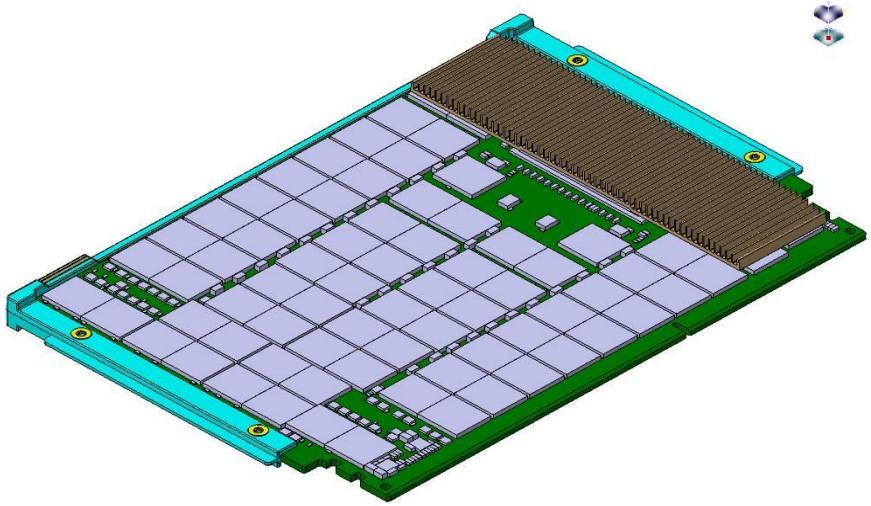
Cache Bandwidths



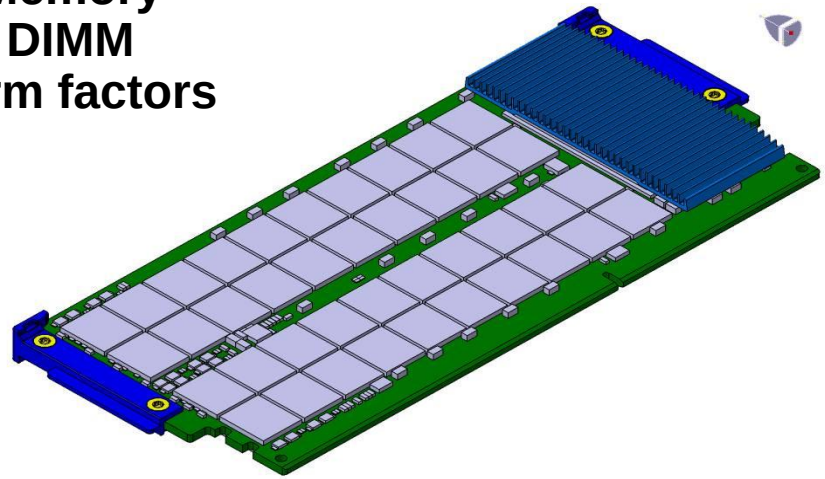
- **GB/sec shown assuming 4 GHz**
  - Product frequency will vary based on model type
- **Across 12 core chip**
  - 4 TB/sec L2 BW
  - 3 TB/sec L3 BW

POWER8 Memory Organization



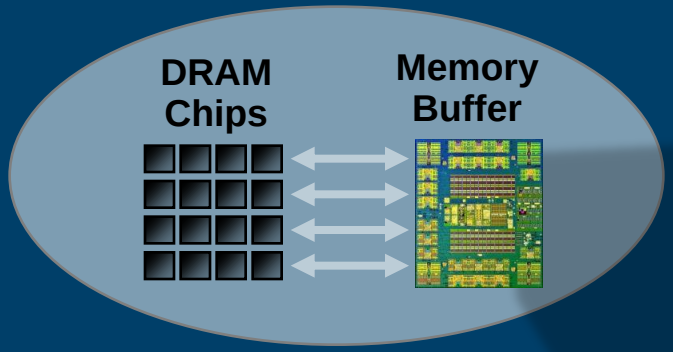


Memory DIMM Form factors



- *Up to 4 high speed channels, each running up to 9.6 Gb/s*
- *Up to 16 total DDR ports*

POWER8 Memory Buffer Chip *…with 16MB of Cache…*



Intelligence Moved into Memory

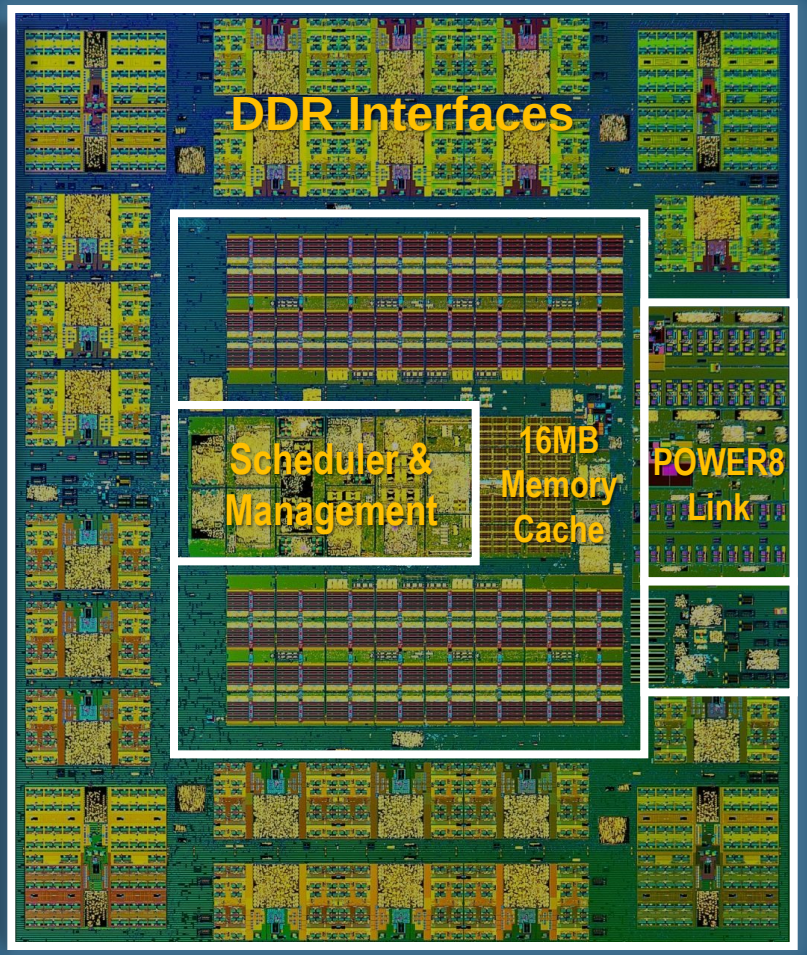
- Scheduling logic, caching structures
- Energy Mgmt, RAS decision point
  - Formerly on Processor
  - Moved to Memory Buffer

Processor Interface

- 9.6 GB/s high speed interface
- More robust RAS
- “On-the-fly” lane isolation/repair
- Extensible for innovation build-out

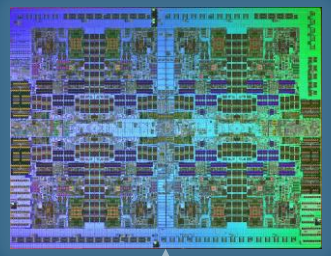
Performance Value

- End-to-end fastpath and data retry (latency)
- Cache → latency/bandwidth, partial updates
- Cache → write scheduling, prefetch, energy
- 22nm SOI for optimal performance / energy
- 15 metal levels (latency, bandwidth)



POWER8 Integrated PCIe Gen 3

POWER7



GX Bus

I/O Bridge

PCIe G2

PCI Device

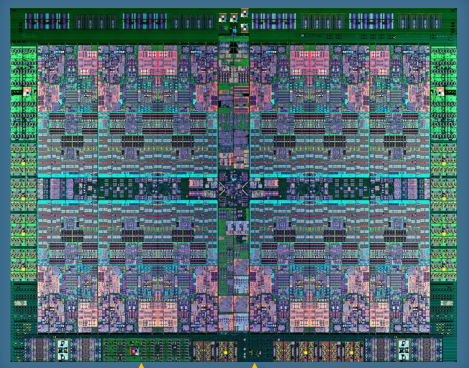
Native PCIe Gen 3 Support

- Direct processor integration
- Replaces proprietary GX/Bridge
- Low latency
- Gen3 x16 bandwidth (16 Gb/s)

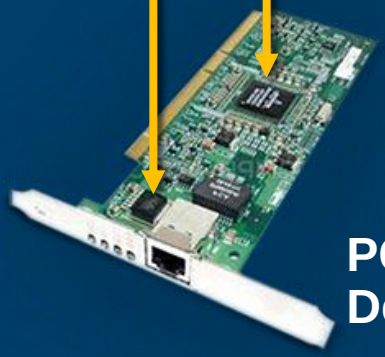
Transport Layer for CAPI Protocol

- Coherently Attach Devices connect to processor via PCIe
- Protocol encapsulated in PCIe

POWER8



PCIe G3

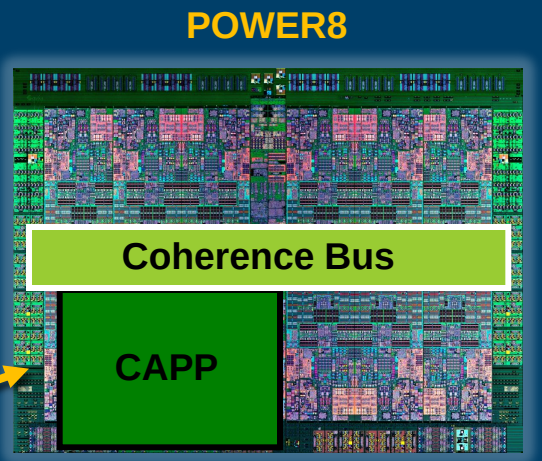


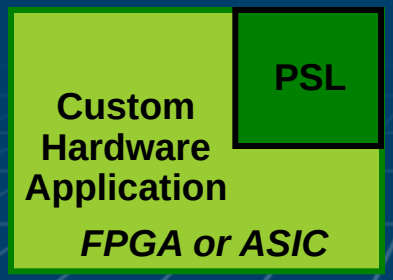
PCI Device

POWER8 CAPI Coherent Accelerator Processor Interface

Virtual Addressing
- Accelerator can work with same memory addresses that the processors use
- Pointers de-referenced same as the host application
- Removes OS & device driver overhead

Hardware Managed Cache Coherence
- Enables the accelerator to participate in “Locks” as a normal thread Lowers Latency over IO communication model



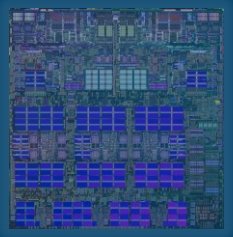
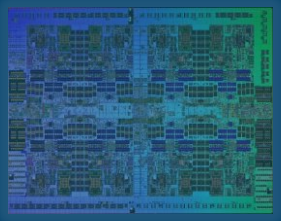
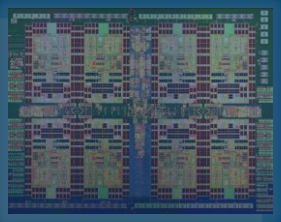
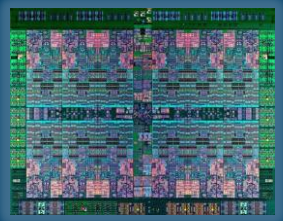


PCIe Gen 3
Transport for encapsulated messages

Customizable Hardware Application Accelerator
- Specific system SW, middleware, or user application
- Written to durable interface provided by PSL

Processor Service Layer (PSL)
- Present robust, durable interfaces to applications
- Offload complexity / content from CAPP

POWER8 Innovation

	POWER5 *2004*	POWER6 *2007*	POWER7 *2010*	POWER7+ *2012*	POWER8
					
Technology	130nm SOI	65nm SOI	45nm SOI eDRAM	32nm SOI eDRAM	22nm SOI eDRAM
Compute					
Cores	2	2	8	8	12
Threads	SMT2	SMT2	SMT4	SMT4	SMT8
Caching					
On-chip	1.9MB	8MB	2 + 32MB	2 + 80MB	6 + 96MB
Off-chip	36MB	32MB	None	None	128MB
Die Bandwidth					
Sust. Mem.	15GB/s	30GB/s	100GB/s	100GB/s	Up to 230GB/s
Peak I/O	6GB/s	20GB/s	40GB/s	40GB/s	Up to 96GB/s

Power Systems *…a New Conversation…*

Solutions

Industry Solutions

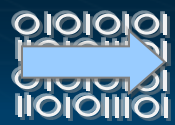
Business & Predictive Analytics

Cognitive Computing

Stream Computing ***Real-time Analytics*** ***Natural Language Learning***

Systems Design

Continuous data load	Massive IO bandwidth	Parallel processing	1,000+ Concurrent Queries	Large-scale memory processing



Open & flexible infrastructure
Available on premise or through the Cloud

POWER8 Enabling: *…Big Data, Analytics, Cognitive Computing…*

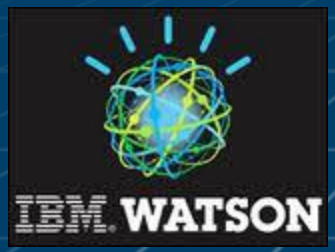
POWER8 Differentiation for Analytics

- Massive capacity and bandwidth to memory and IO
- Large caches with massive bandwidth
- Strong Single thread
- SMT8, Many threads to hide memory latency
  - Graph traversals
  - Transactional memory enables efficient thread scaling

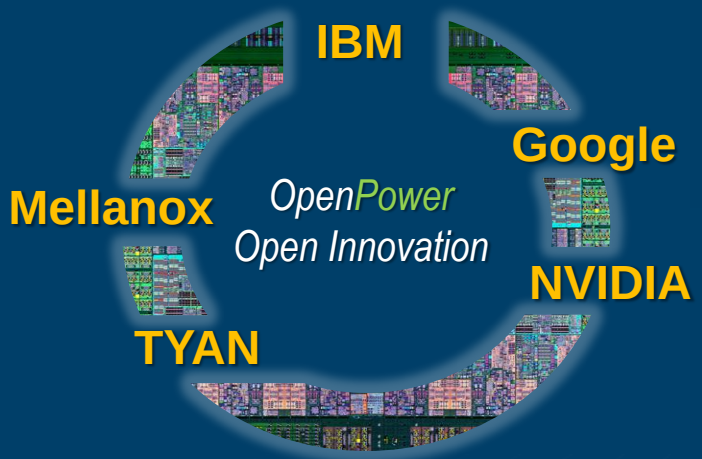
CAPI Accelerators

- Enables heterogeneous compute (GPU, FPGA, etc.)

Synergy with IBM Software, Driving Optimization Across the Stack



OpenPOWER *…giving ecosystem partners a license to innovate…*

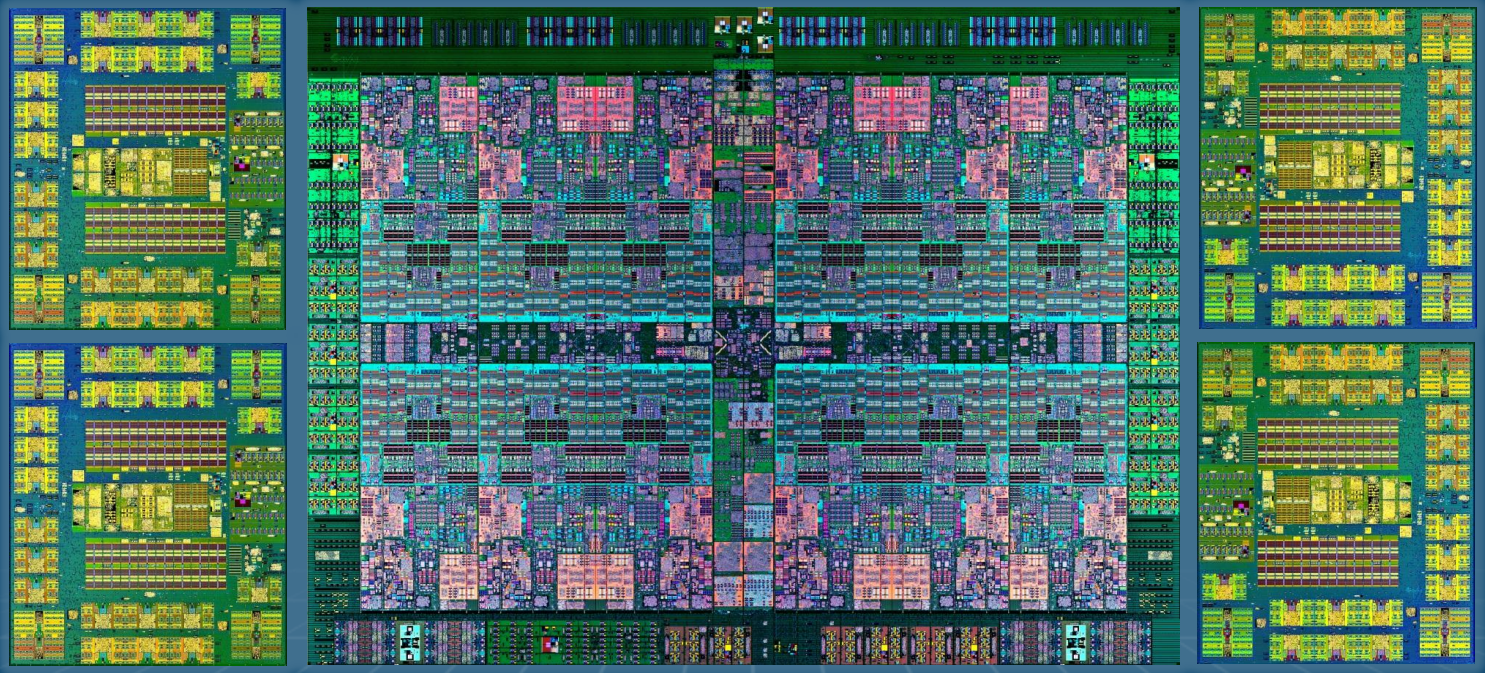


OpenPOWER will enable hyper-scale cloud data centers to rethink their approach to technology.

Member companies will use **POWER** for custom open servers and components for Linux based cloud data centers.

For the first time, **OpenPOWER** ecosystem partners can optimize the interactions of server building blocks – microprocessors, networking, I/O & other components – to tune performance.

POWER8



- *Significant Performance at Thread, Core, and System*
- *Optimization for VM Density & Efficiency*
- *Strong Enablement of Autonomic System Optimization*
- *Excellent Big Data Analytics Capability*

Thank You!

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.

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 5L, 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, Active Memory, Balanced Warehouse, CacheFlow, Cool Blue, IBM Watson, IBM Systems Director VMControl, pureScale, TurboCore, Chiphopper, Cloudscape, DB2 Universal Database, DS4000, DS6000, DS8000, EnergyScale, Enterprise Workload Manager, General Parallel File System, , GPFS, HACMP, HACMP/6000, HASM, IBM Systems Director Active Energy Manager, iSeries, Micro-Partitioning, POWER, PowerLinux, 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+, POWER7, POWER7+, Systems, System i, System p, System p5, System Storage, System z, 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 full list of U.S. trademarks owned by IBM may be found at: http://www.**ibm.com**/legal/copytrade.shtml.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

AltiVec is a trademark of Freescale Semiconductor, Inc.

AMD Opteron is a trademark of Advanced Micro Devices, Inc.

InfiniBand, InfiniBand Trade Association and the InfiniBand design marks are trademarks and/or service marks of the InfiniBand Trade Association.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency which is now part of the Office of Government Commerce.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries or both.

PowerLinux™ uses the registered trademark Linux® pursuant to a sublicense from LMI, the exclusive licensee of Linus Torvalds, owner of the Linux® mark on a world-wide basis.

Microsoft, Windows and the Windows logo are registered trademarks of Microsoft Corporation in the United States, other countries or both.

NetBench is a registered trademark of Ziff Davis Media in the United States, other countries or both.

SPECint, SPECfp, SPECjbb, SPECweb, SPECjAppServer, SPEC OMP, SPECviewperf, SPECapc, SPEChpc, SPECjvm, SPECmail, SPECimap and SPECsfs are trademarks of the Standard Performance Evaluation Corp (SPEC).

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.

TPC-C and TPC-H are trademarks of the Transaction Performance Processing Council (TPPC).

UNIX is a registered trademark of The Open Group in the United States, other countries or both.

Other company, product and service names may be trademarks or service marks of others.