Future Commodity Chip Called CELL for HPC

Piotr Pietrzak
Systems Architect
IBM Systems & Technology Group

Short agenda

- **An overview of curent HPC situation**
- **PowerPC architecture for HPC**
- **An overview of**
  - Cell history
  - Cell microprocessor highlights
  - Hardware components
  - Software environment
  - Software development considerations
  - Cell performance characteristics
  - Cell blade server
  - Cell application affinity and target opportunities
- **Curiosity**

IBM POWER Architecture™ leads in TOP500 supercomputers

Semiannual independent ranking of top 500 supercomputers in the world*

Technology	# of Processors	Share
POWER3™	16,768	2.3%
POWER4™	8,608	1.2%
POWER4+™	21,866	3.0%
POWER5™	22,208	3.0%
PowerPC® (970)	14,460	2.0%
Power PC 440	243,712	33.3%
IBM POWER™ Total	**327,622**	**44.8%**
Pentium® 4 Xeon™	184,908	25.3%
Xeon EM64T	58,204	8.0%
Itanium® 2	45,064	6.2%
Intel Total	**288,176**	**39.4%**
HP (PA-RISC)	14,784	2.0%
Opteron	68,789	9.4%
Cray X1	3,034	0.4%
NEC	6,072	0.8%
SPARC	6,112	0.8%
Alpha	15,160	2.1%
Hitachi SR8000	1,320	0.2%
Other Total	**115,271**	**15.8%**

Source: http://www.top500.org

p5-575 Is already top UNIX® system in TOP500!

TOP500[1]	Installation	Processor	Rmax TF/s
1	DOE BlueGene®/L LLNL	700 MHz PPC 440	280.6
2	BlueGene at Watson	700 MHz PPC 440	91.2
3	**ASC Purple LLNL**	**1.9 GHz p5-575**	**63.4**

BASED ON CURRENT TOP500 LIST, p5-575 IS THE WORLD'S MOST POWERFUL UNIX SYSTEM!

16-core 1.9 GHz p5-575 delivers nearly 2X the LINPACK performance of the 8-core 1.9 GHz p5-575[2]

*Sources:
1. http://www.top500.org
2. http://www.netlib.org/benchmark/performance.pdf

Source: www.top500.org

p5-575 and Blue Gene/L differentiation

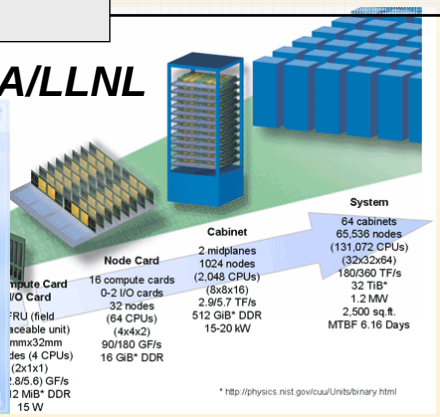
p5-575:	Blue Gene/L:
64-bit AIX 5L/Linux cluster nodes suitable for applications requiring high memory bandwidth and large memory (32GB) per 64-bit processor.	32-bit Linux and custom kernel clusters suitable for highly parallel applications with limited memory requirements (256MB per 32-bit processor) and limited or highly parallelized I/O.
Scalable systems: 16 to 2,048 1.9 GHz POWER5+ CPUs	Very large systems: up to 100,000+ 667 MHz PPC440 CPUs
“Off-the-shelf” and custom configurations	Custom configurations
Standard IBM service and support	Custom service and support
1,000s of applications supported	Highly effective in target applications

CURRENT DEPLOYMENT EXAMPLES

ASC PURPLE – NNSA/LLNL

Source: http://www.netlib.org/benchmark/performance.pdf as of 02/14/2006

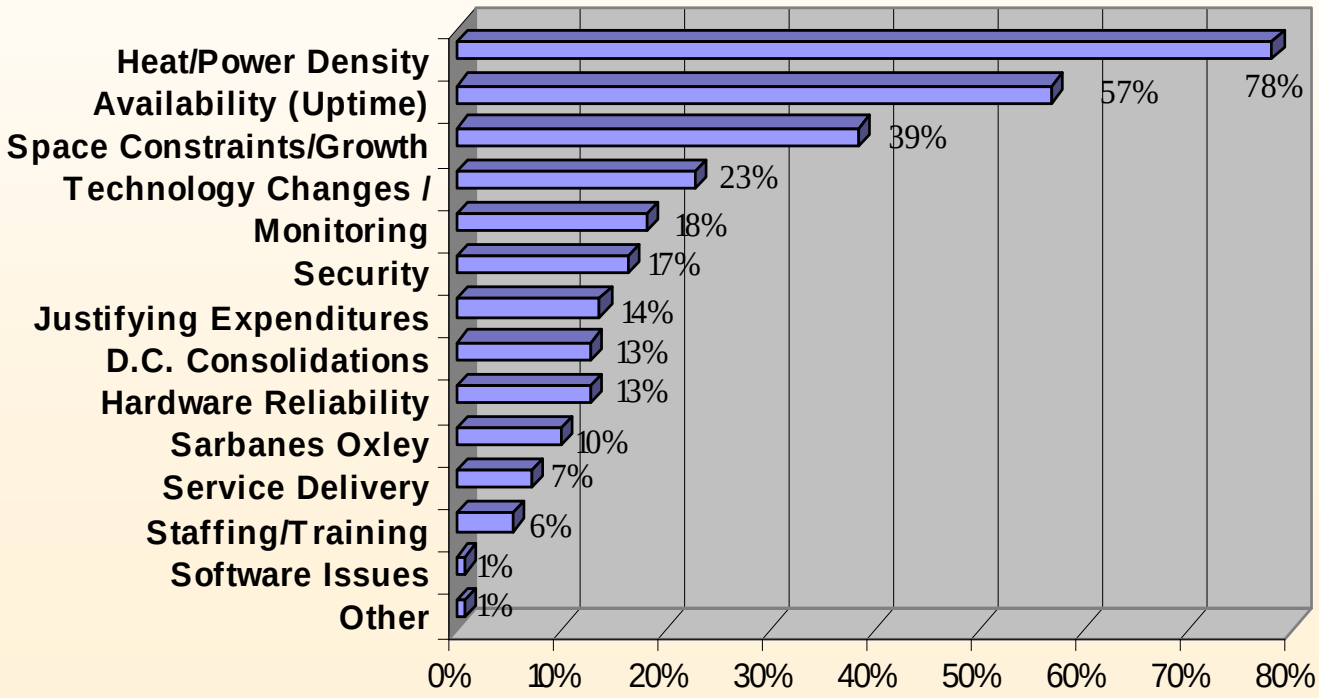
Blue Gene/L – NNSA/LLNL



Source: http://www.top500.org/lists/plists.php?Y=2005&M=06 as of 02/14/2006

Technology: Heat/power density tops list of IT management concerns

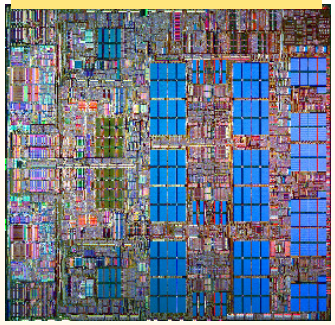
Spring 2005 Data Center Users' Group conference
- The Adaptive Data Center: Managing Dynamic Technologies.



Source: http://www.liebert.ws/liebertmadara/liebertmadara_files/Default.htm#nopreload=1

Enhanced POWER processor capabilities

POWER5+



POWER5+ Enhancements

- Higher frequencies
- 37% reduction in size vs. POWER5
- Reduction in power consumption
- Large page size support
- Memory controller improvements
- Quad-Core Module support
- Better performance

IBM PowerPC 970MP

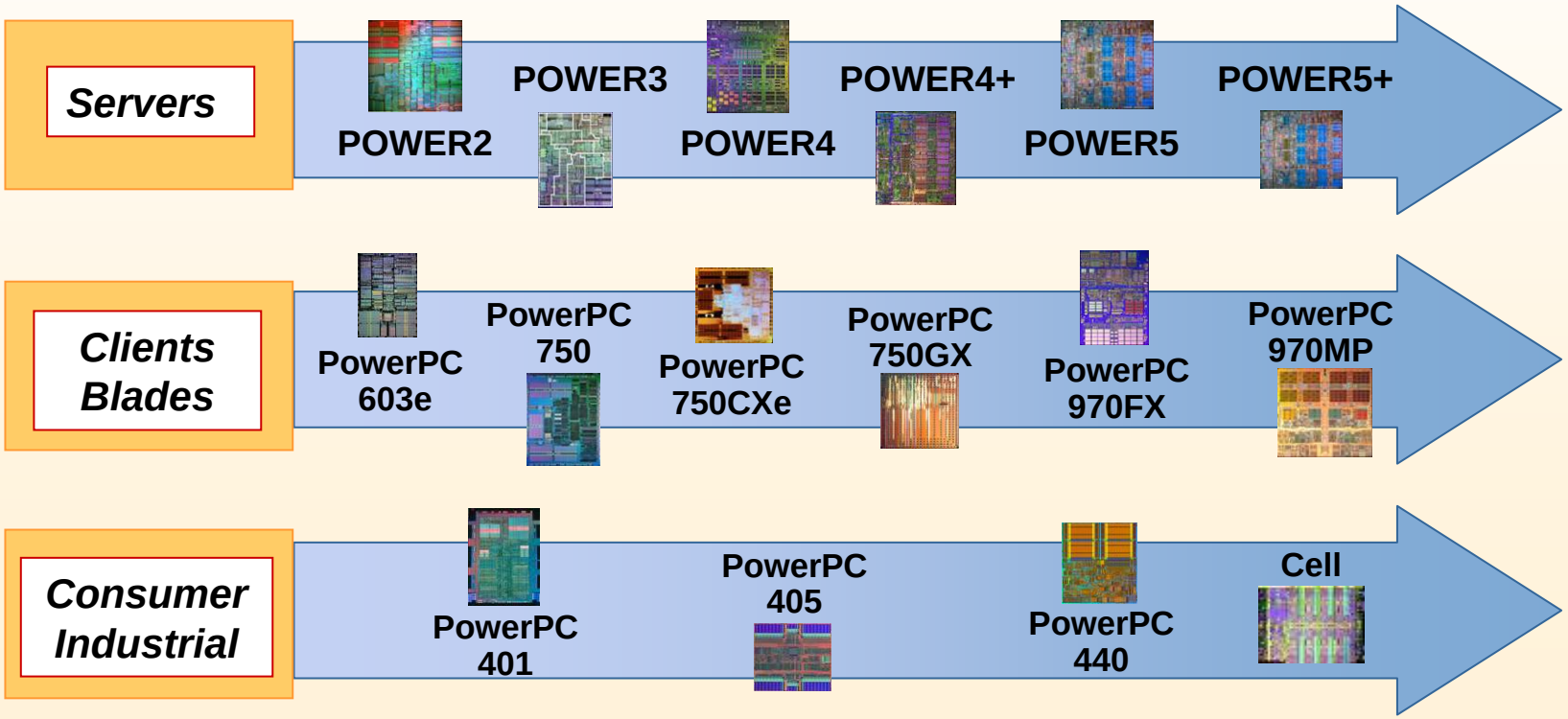
IBM PowerPC 970MP

- High performance and advaced function - over 50 million transistors
- Single instruction multiple data (SIMD) units accelerate HPC workloads
- Low power consumption
- Combines silicon-on-insulator, strained silicon, and copper wiring technologies

IBM POWER Architecture

From consumer electronics to supercomputers

- A common architecture . . . the most scalable technology



Source: http://www.ibm.com/chips/power/aboutpower/

IBM System p5 575: *THE CRUNCHER!*

New 8-core
DATA CRUNCHER

p5-575 [8-core] 2.2 GHz
66.4 GFLOPS/node
LINPACK HPC
100.5 GB/sec
STREAM Traid (Tuned)

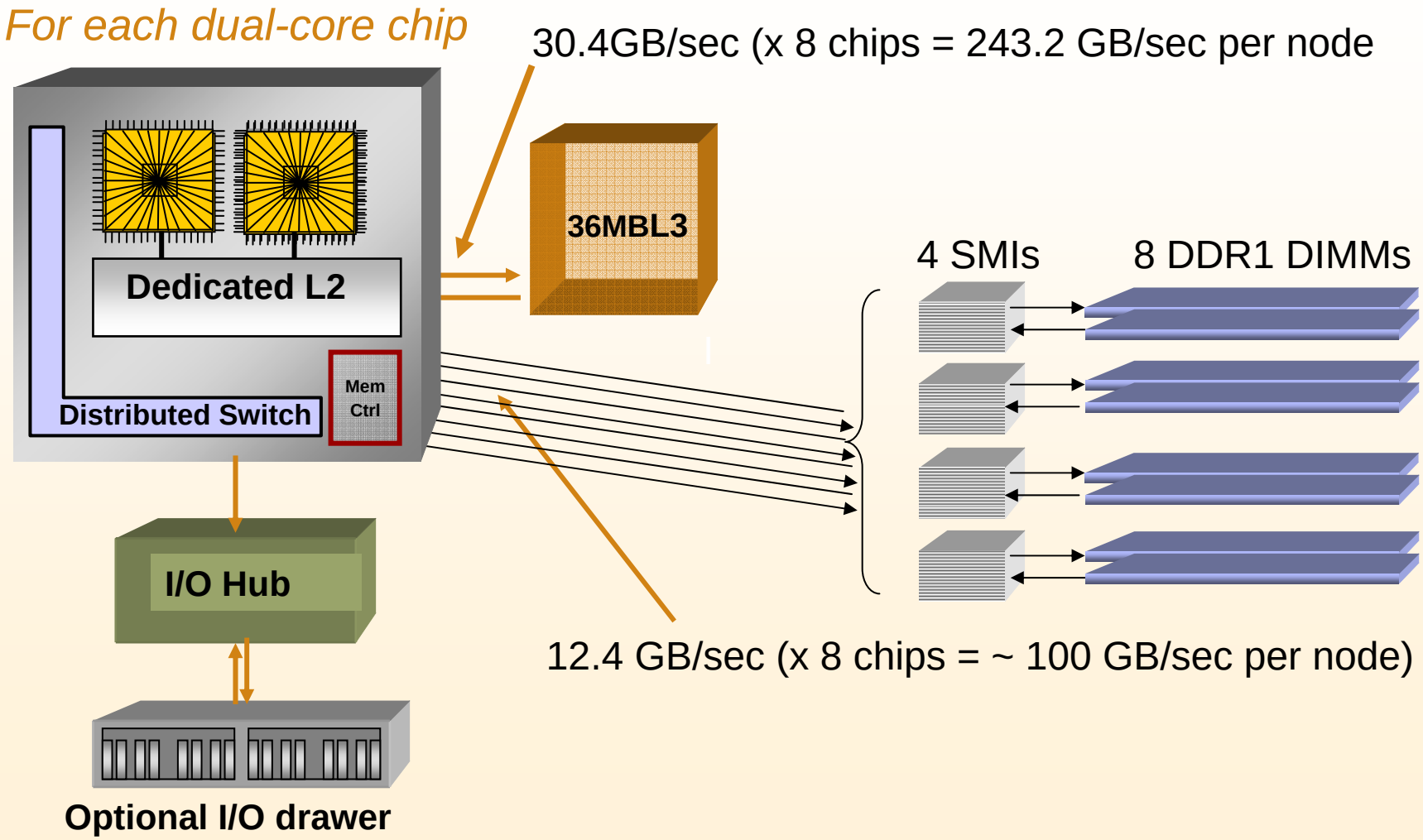
New 16-core
NUMBER CRUNCHER

p5-575 [16-core] 1.9 GHz
111.4 GFLOPS/node
LINPACK HPC
86.3 GB/sec
STREAM Triad (Tuned)

Sources: http://www.cs.virginia.edu/stream/ submitted February 14, 2006
http://www.netlib.org/benchmark/performance.pdf submitted February 14,2006

1: Node hardware only, 1GB memory 2: Node hardware only, 16GB memory

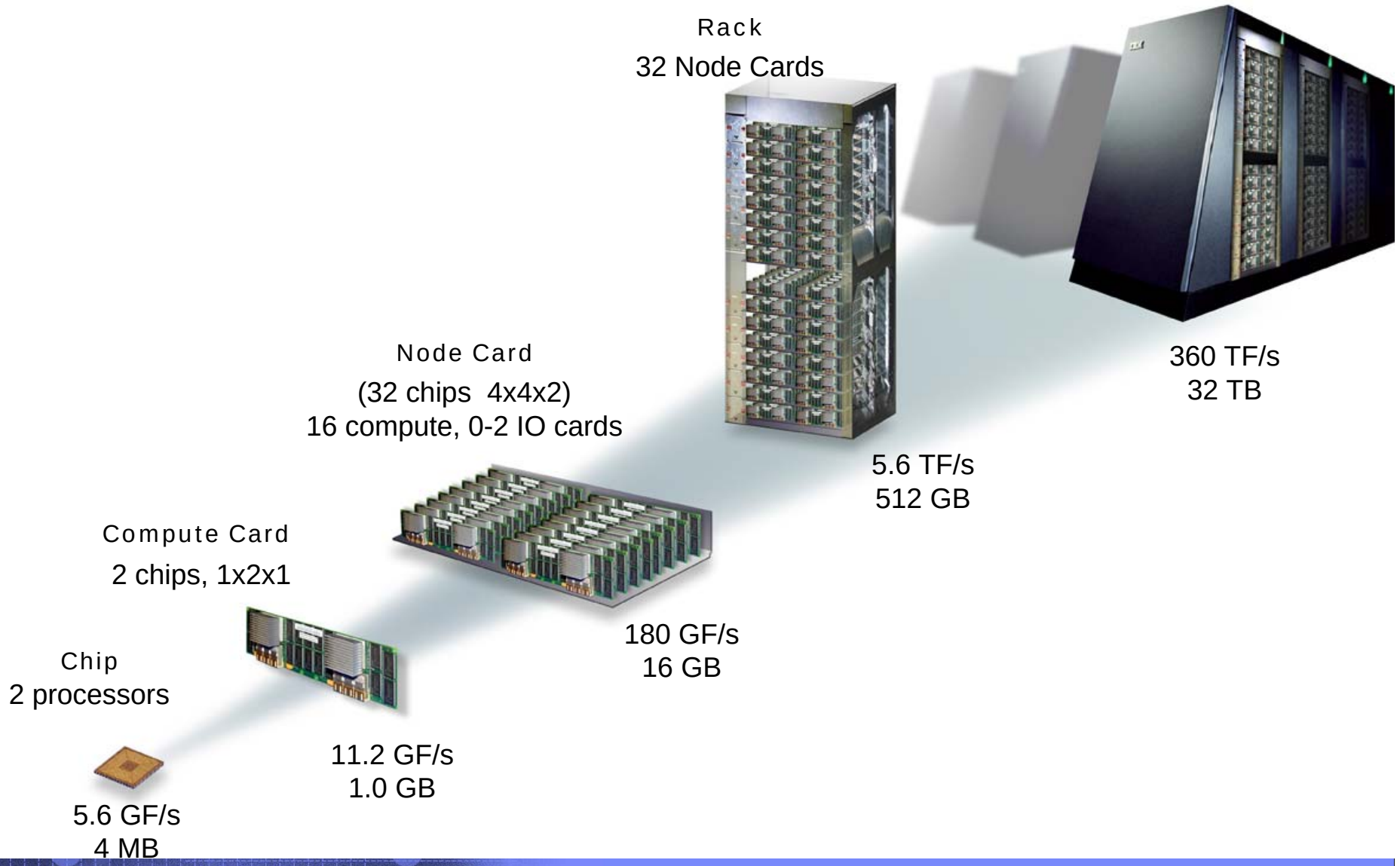
p5-575 Peak Bandwidths per 1.5 GHz 2-way Chip

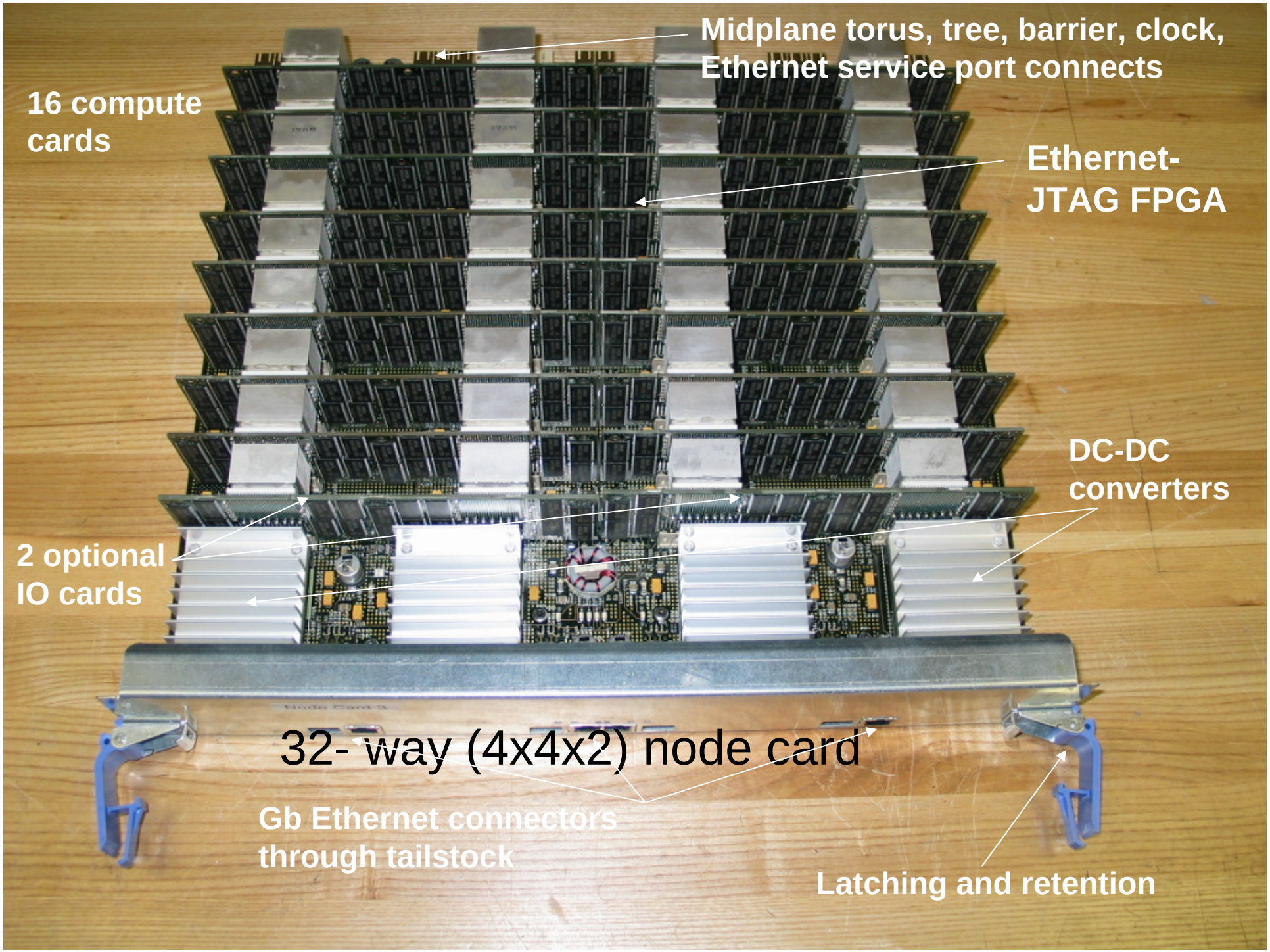


Facts About the Blue Gene® Program

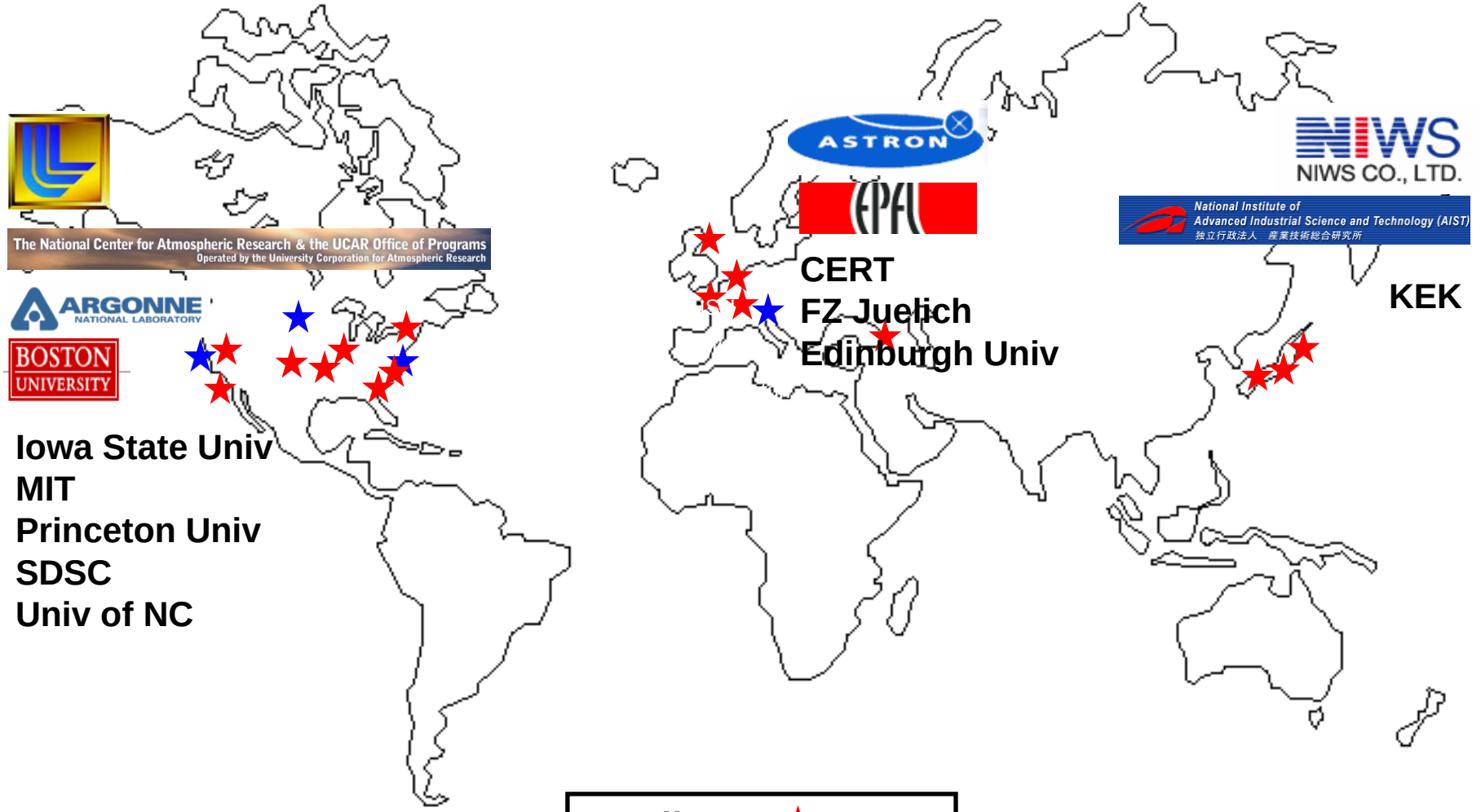
1) ***Delivered on the promise of leadership HPC performance***
   - On target to reach 360 peak Teraflops
   - Other application and industry performance studies as available
2) ***Broad applicability of Blue Gene to important HPC workloads helps advance science, engineering and business***
   - Topics at the forefront of international concern:
     - Defense and Homeland Security
     - Healthcare and Life Sciences
     - Environmental and Climate Modeling; Weather Forecasting
     - Energy Production and Resource Management
     - Product Safety and Efficiency
   - Cultivating ISV support in key application areas
3) ***Enhanced accessibility to world-class computing***
   - Inclusion of Blue Gene in the Deep Computing Capacity On Demand center
   - Introduction of reduced-sized Blue Gene systems (limited scope)
   - IBM Global Financing - Leasing and Financing Programs
4) ***Commercialized and aligned with the IBM server portfolio***
   - Leveraging and advancing the strengths of Power technology

BlueGene® /L System Buildup




Midplane torus, tree, barrier, clock, Ethernet service port connects
16 compute cards
Ethernet-JTAG FPGA
DC-DC converters
2 optional IO cards
32- way (4x4x2) node card
Gb Ethernet connectors through tailstock
Latching and retention

Where in the world is Blue Gene®?



Blue Gene® Performance and Density Guidance

Performance

Performance Metric	Single Rack Blue Gene Rating
Peak Teraflops (Virtual Node mode)	5.73
Peak Teraflops (Coprocessor mode)	2.86
Linpack Teraflops	4.53
MTOPS (Exceeding 190,000 requires Gov't Export License into Tier 3 countries)	>1.5M

Density

Metric	ASCI White	ASCI Q	Earth Simulator	BG/L
Memory/Space (GB/sq.m)	8.6	17	3.1	***140***
Speed/Space (GFlops*/sq.m)	13	16	13	***1600***
Speed/Power (GFlops*/kW)	12	7.9	4	***300***

*** Peak**

As of Nov 2005

Blue Gene® Award-Winning Performance

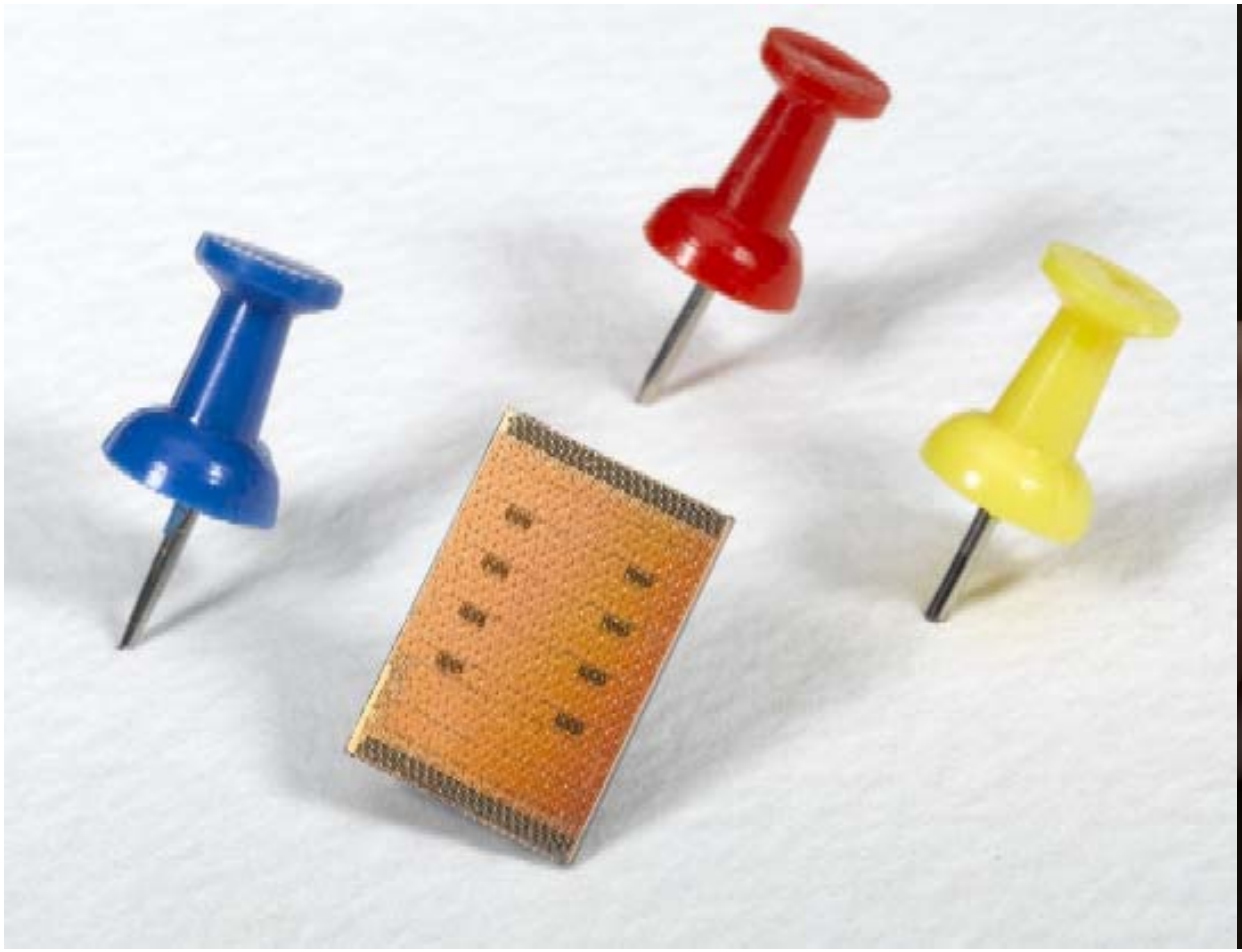
- 3 in Top10 (#1 and #2)
- 7 in Top50 (11-50)
- 9 in Top100 (51-100)
- 19 in Top500

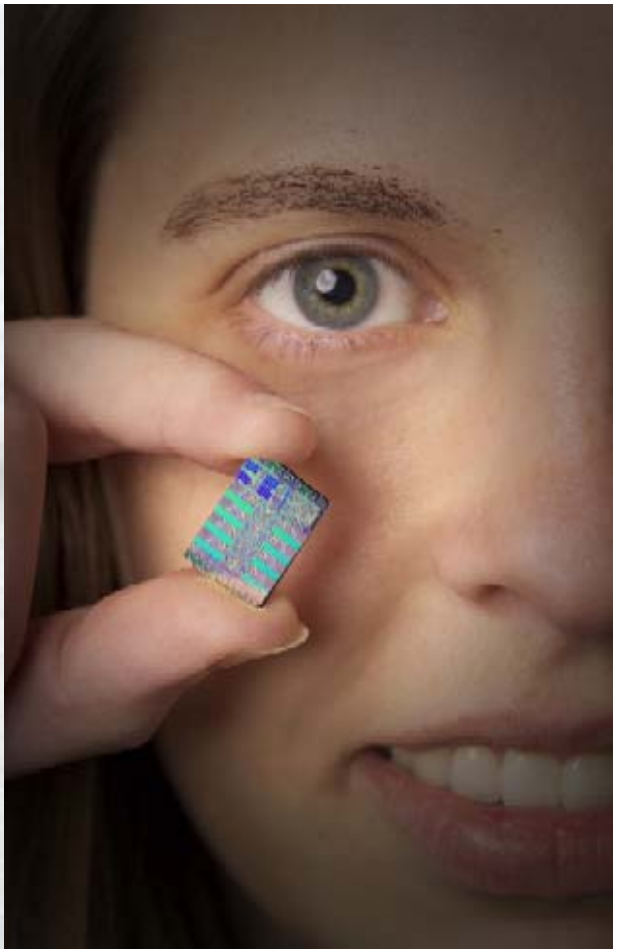
www.top500.org

HPC Challenge Benchmarks	IBM Blue Gene 65,536 nodes 64 racks
HPL (TFLOP/s) – Linpack TPP benchmark which measures floating point rate of execution for solving linear system of equations	259.21
RANDOMACCESS (GUP/s) – measures rate of integer random updates of memory	35.46
FETE (GFLOP/s) – measures floating point rate of execution of double precision complex one-dimensional Discrete Fourier Transform (DFT)	2311.09
STREAM (GB/s) – simple synthetic that measures sustainable memory bandwidth and corresponding computation rate for simple vector kernel	160,064

As of Nov 2006

Cell





Cell History

- **IBM, SCEI/Sony, Toshiba Alliance formed in 2000**
- **Design Center opened in March 2001**
  - Based in Austin, Texas
- **Single CellBE operational Spring 2004**
- **2-way SMP operational Summer 2004**
- **February 7, 2005: First technical disclosures**
- **October 6, 2005: Mercury Announces Cell Blade**
- **November 9, 2005: Open Source SDK & Simulator Published**
- **November 14, 2005: Mercury Announces Turismo Cell Offering**
- **February 8, 2006 IBM Announced Cell Blade**

Cell Basic Design Concept

Cell Basic Concept

- Compatibility with 64b Power Architecture™
  - Builds on and leverages IBM investment and community
- Increased efficiency and performance
  - Attacks on the "Power Wall"
    - Non Homogenous Coherent Multiprocessor
    - High design frequency @ a low operating voltage with advanced power management
  - Attacks on the "Memory Wall"
    - Streaming DMA architecture
    - 3-level Memory Model: Main Storage, Local Storage, Register Files
  - Attacks on the "Frequency Wall"
    - Highly optimized implementation
    - Large shared register files and software controlled branching to allow deeper pipelines
- Interface between user and networked world
  - Image rich information, virtual reality
  - Flexibility and security
- Multi-OS support, including RTOS / non-RTOS
  - Combine real-time and non-real time worlds

Cell Design Goals

- **Cell is an accelerator extension to Power**
  - Built on a Power ecosystem
  - Used best know system practices for processor design
- **Sets a new performance standard**
  - Exploits parallelism while achieving high frequency
  - Supercomputer attributes with extreme floating point capabilities
  - Sustains high memory bandwidth with smart DMA controllers
- **Designed for natural human interaction**
  - Photo-realistic effects
  - Predictable real-time response
  - Virtualized resources for concurrent activities
- **Designed for flexibility**
  - Wide variety of application domains
  - Highly abstracted to highly exploitable programming models
  - Reconfigurable I/O interfaces
  - Virtual trusted computing environment for security

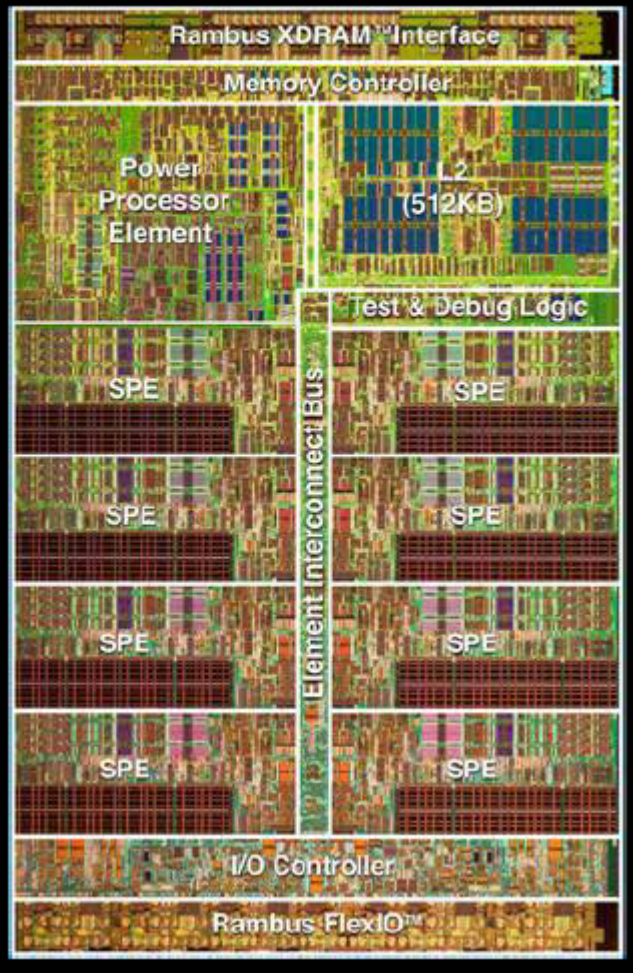
Cell Synergy

- **Cell is not a collection of different processors, but a synergistic whole**
  - Operation paradigms, data formats and semantics consistent
  - Share address translation and memory protection model
- **PPE for operating systems and program control**
- **SPE optimized for efficient data processing**
  - SPEs share Cell system functions provided by Power Architecture
  - MFC implements interface to memory
    - Copy in/copy out to local storage
- **PowerPC provides system functions**
  - Virtualization
  - Address translation and protection
  - External exception handling
- **EIB integrates system as data transport hub**

Cell Hardware Components

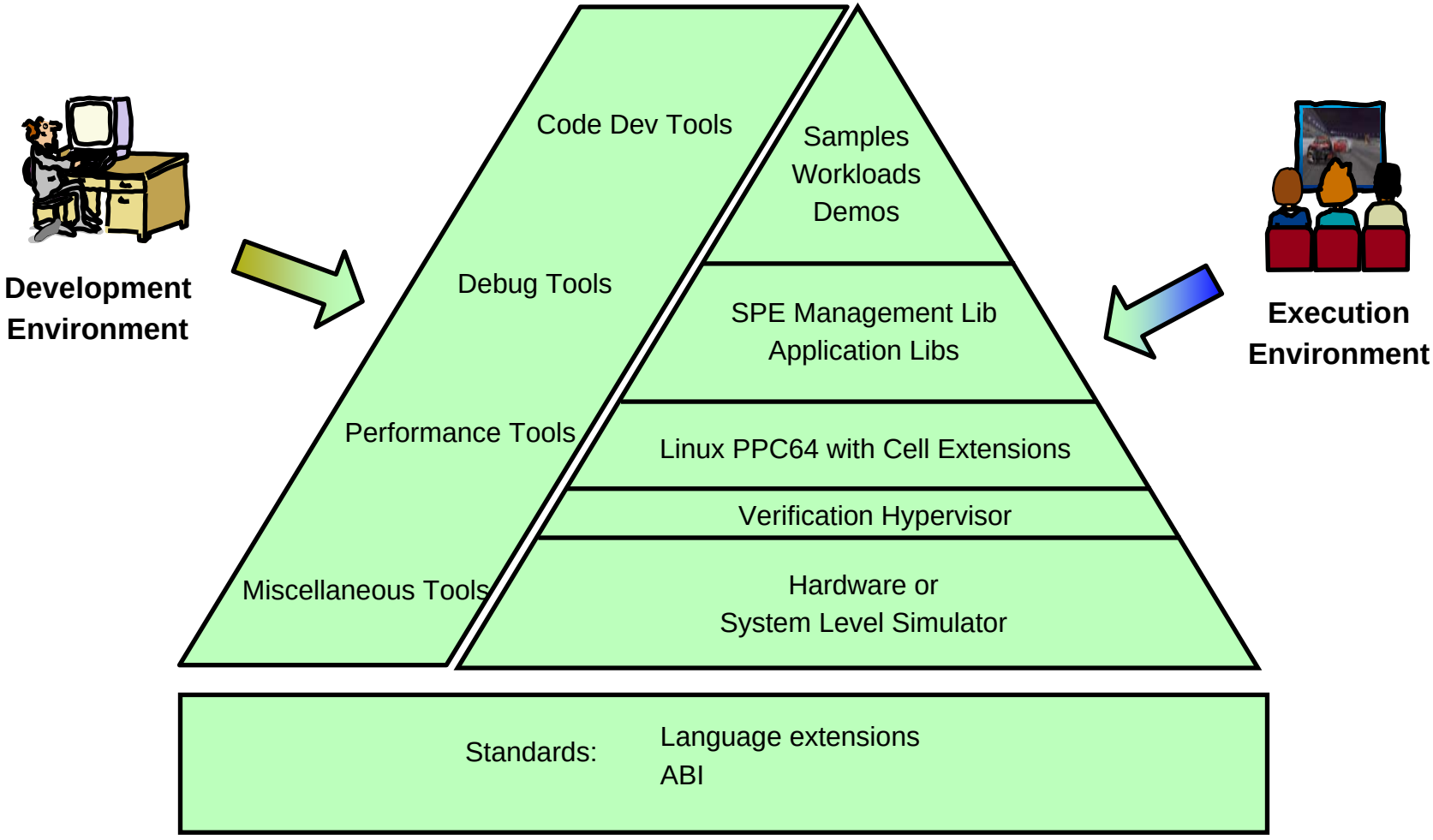
Highlights (3.2 GHz)

- 241M transistors
- 235mm2
- 9 cores, 10 threads
- >200 GFlops (SP)
- >20 GFlops (DP)
- Up to 25 GB/s memory B/W
- Up to 75 GB/s I/O B/W
- >300 GB/s EIB
- Top frequency >4GHz
  (observed in lab)

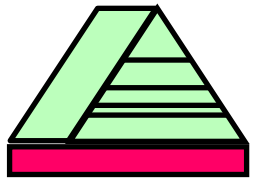


Cell Software Environment

Cell Software Environment

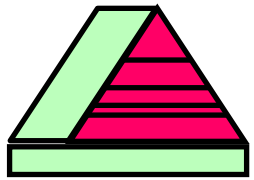


CBE Standards



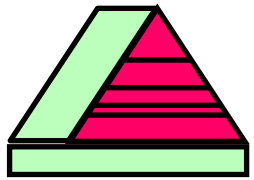
- **Application Binary Interface Specifications**
  - Defines such things as data types, register usage, calling conventions, and object formats to ensure compatibility of code generators and portability of code.
    - SPE ABI
    - Linux for CBE Reference Implementation ABI
- **SPE C/C++ Language Extensions**
  - Defines standardized data types, compiler directives, and language intrinsics used to exploit SIMD capabilities in the core.
  - Data types and Intrinsics styled to be similar to Altivec/VMX.
- **SPE Assembly Language Specification**

System Level Simulator



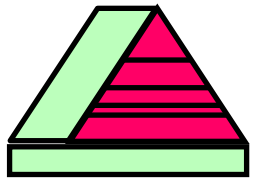
- **Cell BE – full system simulator**
  - Uni-Cell and multi-Cell simulation
  - User Interfaces – TCL and GUI
  - Cycle accurate SPU simulation (pipeline mode)
  - Emitter facility for tracing and viewing simulation events

SPE Management Library



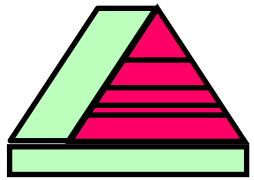
- **SPEs are exposed as threads**
  - SPE thread model interface is similar to POSIX threads.
  - SPE thread consists of the local store, register file, program counter, and MFC-DMA queue.
  - Associated with a single Linux task.
  - Features include:
    - **Threads** - create, groups, wait, kill, set affinity, set context
    - **Thread Queries** - get local store pointer, get problem state area pointer, get affinity, get context
    - **Groups** - create, set group defaults, destroy, memory map/unmap, madvise.
    - **Group Queries** - get priority, get policy, get threads, get max threads per group, get events.
    - **SPE image files** - opening and closing
- **SPE Executable**
  - Standalone SPE program managed by a PPE executive.
  - Executive responsible for loading and executing SPE program. It also services assisted requests for I/O (eg, fopen, fwrite, fprintf) and memory requests (eg, mmap, shmat, …).

Optimized SPE and Multimedia Extension Libraries



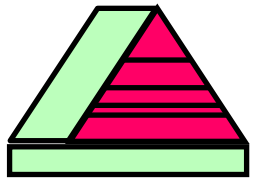
- **Standard SPE C library subset**
  - optimized SPE C99 functions including stdlib c lib, math and etc.
  - subset of POSIX.1 Functions – PPE assisted
- **Audio resample - resampling audio signals**
- **FFT - 1D and 2D fft functions**
- **gmath - mathematic functions optimized for gaming environment**
- **image - convolution functions**
- **intrinsics - generic intrinsic conversion functions**
- **large-matrix - functions performing large matrix operations**
- **matrix - basic matrix operations**
- **mpm - multi-precision math functions**
- **noise - noise generation functions**
- **oscillator - basic sound generation functions**
- **sim – simulator only function including print, profile checkpoint, socket I/O, etc …**
- **surface - a set of bezier curve and surface functions**
- **sync - synchronization library**
- **vector - vector operation functions**

Sample Source



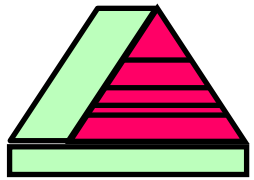
- cesof - the samples for the CBE embedded SPU object format usage
- spu_clean - cleans the SPU register and local store
- spu_entry - sample SPU entry function (crt0)
- spu_interrupt - SPU first level interrupt handler sample
- spulet - direct invocation of a spu program from Linux shell
- sync
- simpleDMA / DMA
- tutorial - example source code from the tutorial
- SDK test suite

Workloads



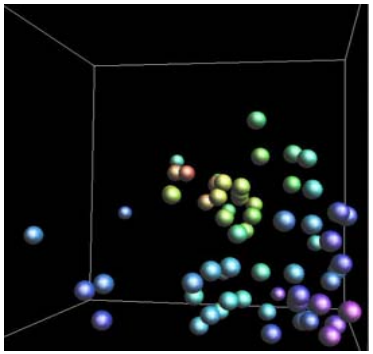
- FFT16M – optimized 16 M point complex FFT
- Oscillator - audio signal generator
- Matrix Multiply – matrix multiplication workload
- VSE_subdiv - variable sharpness subdivision algorithm

Bringup Workloads / Demos

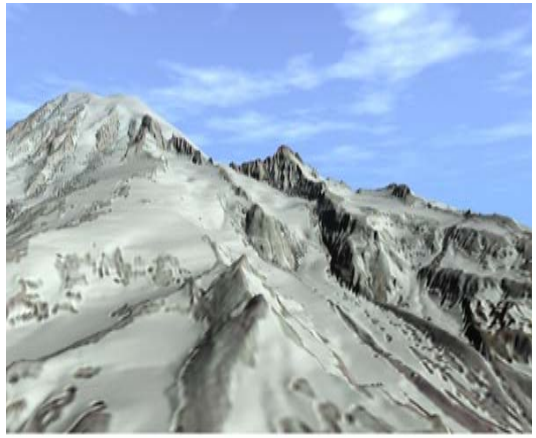


- **Numerous code samples provided to demonstrate system design constructs**
- **Complex workloads and demos used to evaluate and demonstrate system performance**

Geometry Engine

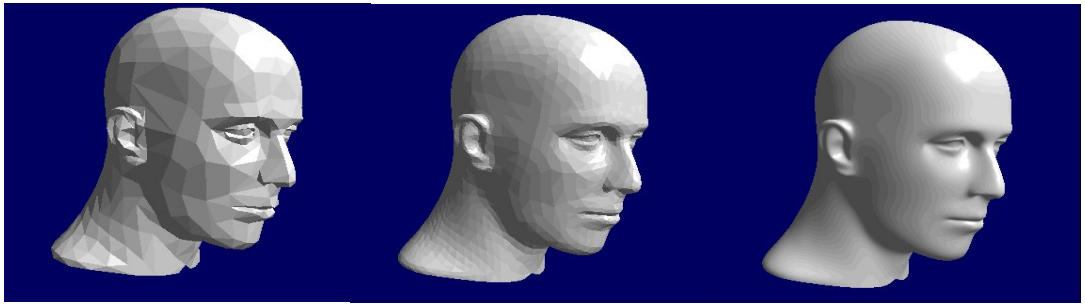


Physics Simulation

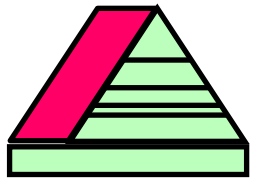


Terrain Rendering Engine

Subdivision Surfaces

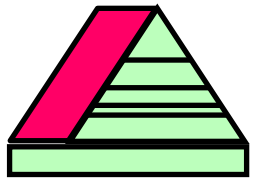


Code Development Tools



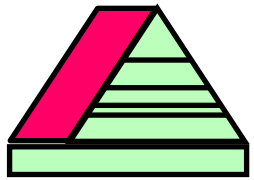
- **GNU based binutils**
  - From Sony Computer Entertainment
  - gas SPE assembler
  - gld SPE ELF object linker
    - ppu-embedspu script for embedding SPE object modules in PPE executables
  - misc bin utils (ar, nm, ...) targeting SPE modules
- **GNU based C/C++ compiler targeting SPE**
  - From Sony Computer Entertainment
  - retargeted compiler to SPE
  - Supports common SPE Language Extensions and ABI (ELF/Dwarf2)
- **Cell Broadband Engine Optimizing Compiler (executable)**
  - IBM XLC C/C++ for PowerPC (Tobey)
  - IBM XLC C retargeted to SPE assembler (including vector intrinsics) - highly optimizing
  - Prototype CBE Programmer Productivity Aids
    - Auto-Vectorization (auto-SIMD) for SPE and PPE Multimedia Extension code
  - spu_timing Timing Analysis Tool

Bringup Debug Tools



- **GNU gdb**
  - Multi-core Application source level debugger supporting PPE multithreading, SPE multithreading, interacting PPE and SPE threads
  - Three modes of debugging SPU threads
    - Standalone SPE debugging
    - Attach to SPE thread
      - Thread ID output when SPU_DEBUG_START=1

SPE Performance Tools (executables)



- **Static analysis (spu_timing)**
  - Annotates assembly source with instruction pipeline state
- **Dynamic analysis (CBE System Simulator)**
  - Generates statistical data on SPE execution
    - Cycles, instructions, and CPI
    - Single/Dual issue rates
    - Stall statistics
    - Register usage
    - Instruction histogram

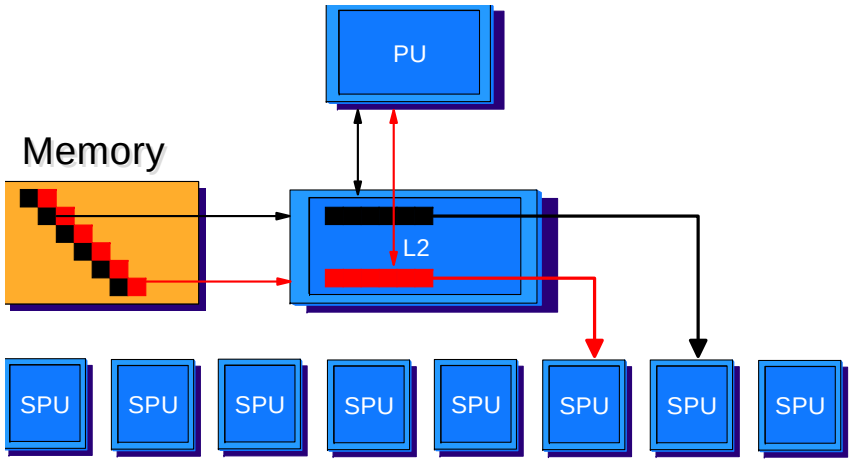
Cell Performance Characteristics

Why Cell processor is so fast?

Key Architectural Reasons

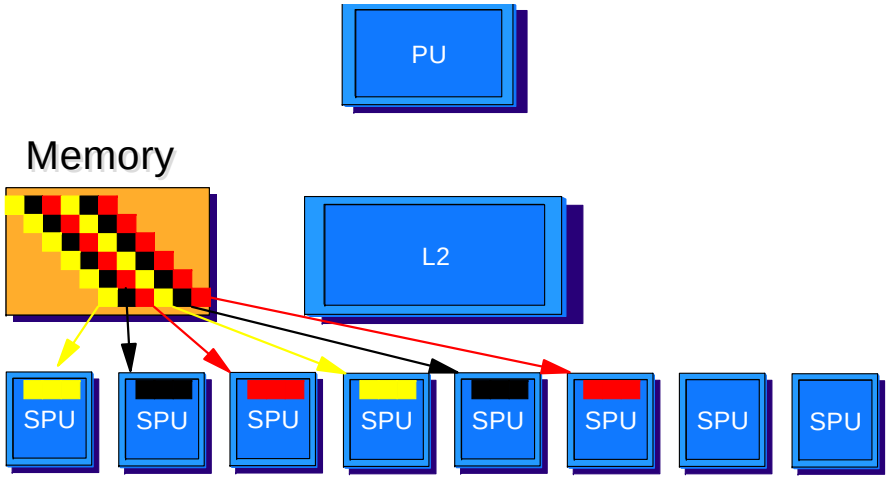
- Parallel processing inside chip
- Fully parallelized and concurrent operations
- Functional offloading
- High frequency design
- High bandwidth for memory and IO accesses
- Fine tuning for data transfer

Staging

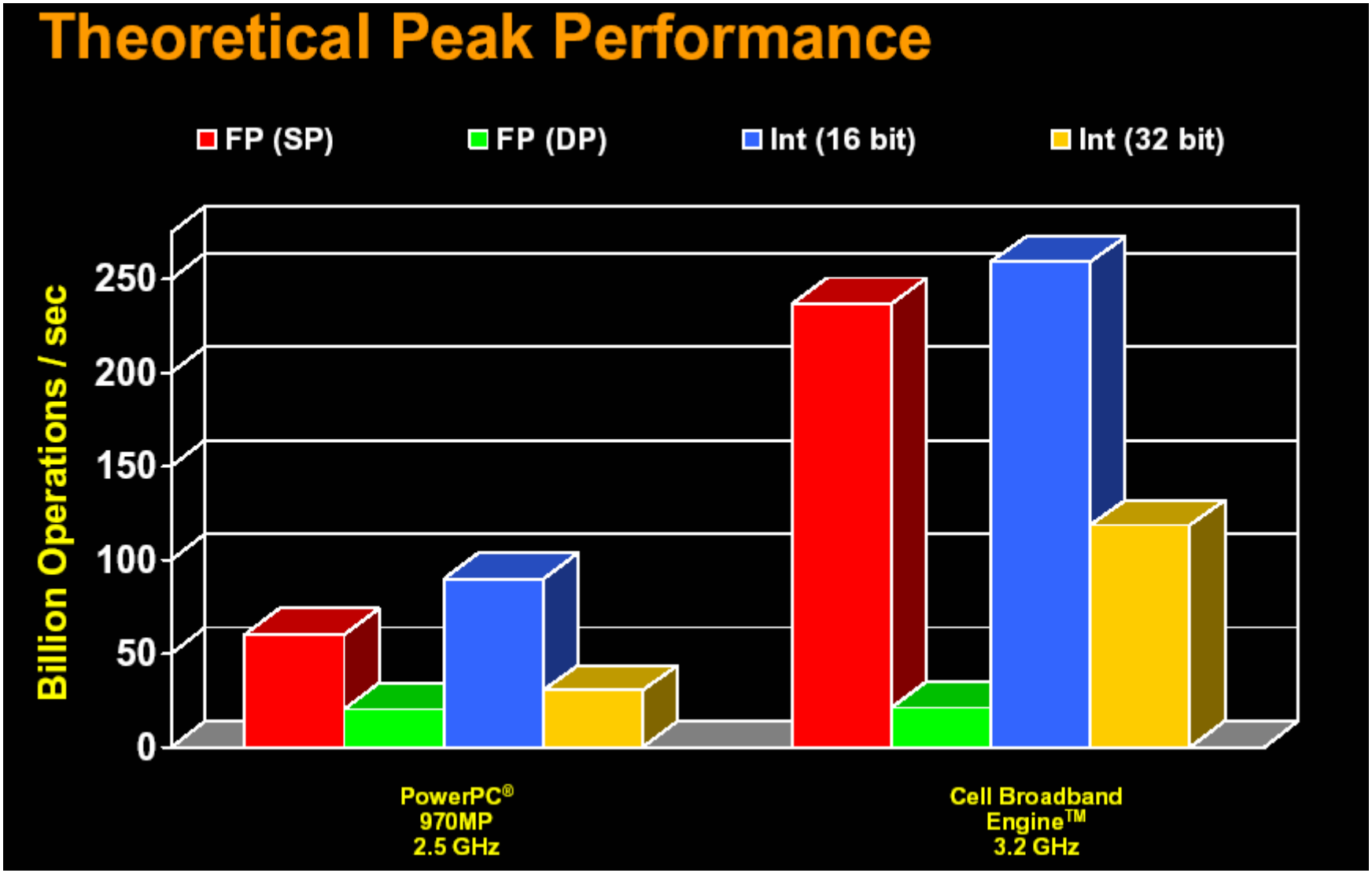


L2 - 4 outstanding loads + 2 prefetch

Data



SPU - 16 outstanding loads per SPU


Theoretical Peak Performance
FP (SP)
FP (DP)
Int (16 bit)
Int (32 bit)
Billion Operations / sec
250
200
150
100
50
0
PowerPC®
970MP
2.5 GHz
Cell Broadband
Engine™
3.2 GHz

Cell BE Performance Summary

Type	Algorithm	3.2 GHz GPP	3.2 GHz Cell	Cell Perf Advantage
HPC	Matrix Multiplication (S.P.)	24 Gflops (w/SIMD)	200 GFlops* (8SPEs)	8x
	Linpack (S.P.)	16 GFlops (w/SIMD)	156 GFlops* (8SPEs)	9x
	Linpack (D.P.): 1kx1k matrix	7.2 GFlops (IA32/SSE3)	9.67 GFLops* (8SPEs)	1.3x
graphics	Transform-light	170 MVPS (G5/VMX)	256 MVPS** (per SPE)	12x
	TRE	1 fps (G5/VMX)	30 fps* (Cell)	30x
security	AES encryp. 128-bit key	1.03 Gbps	2.06Gbps** (per SPE)	16x
	AES decryp. 128-bit key	1.04 Gbps	1.5Gbps** (per SPE)	11x
	TDES	0.12 Gbps	0.16 Gbps** (per SPE)	10x
	DES	0.43 Gbps	0.49 Gbps** (per SPE)	9x
	SHA-1	0.85 Gbps	1.98 Gbps** (per SPE)	18x
video processing	mpeg2 decoder (CIF)	----	1267 fps* (per SPE)	--
	mpeg2 decoder (SDTV)	354 fps (IA32)	365 fps** (per SPE)	8x
	mpeg2 decoder (HDTV)	----	73 fps* (per SPE)	--

Notes: * Hardware measurement ** Simulation results

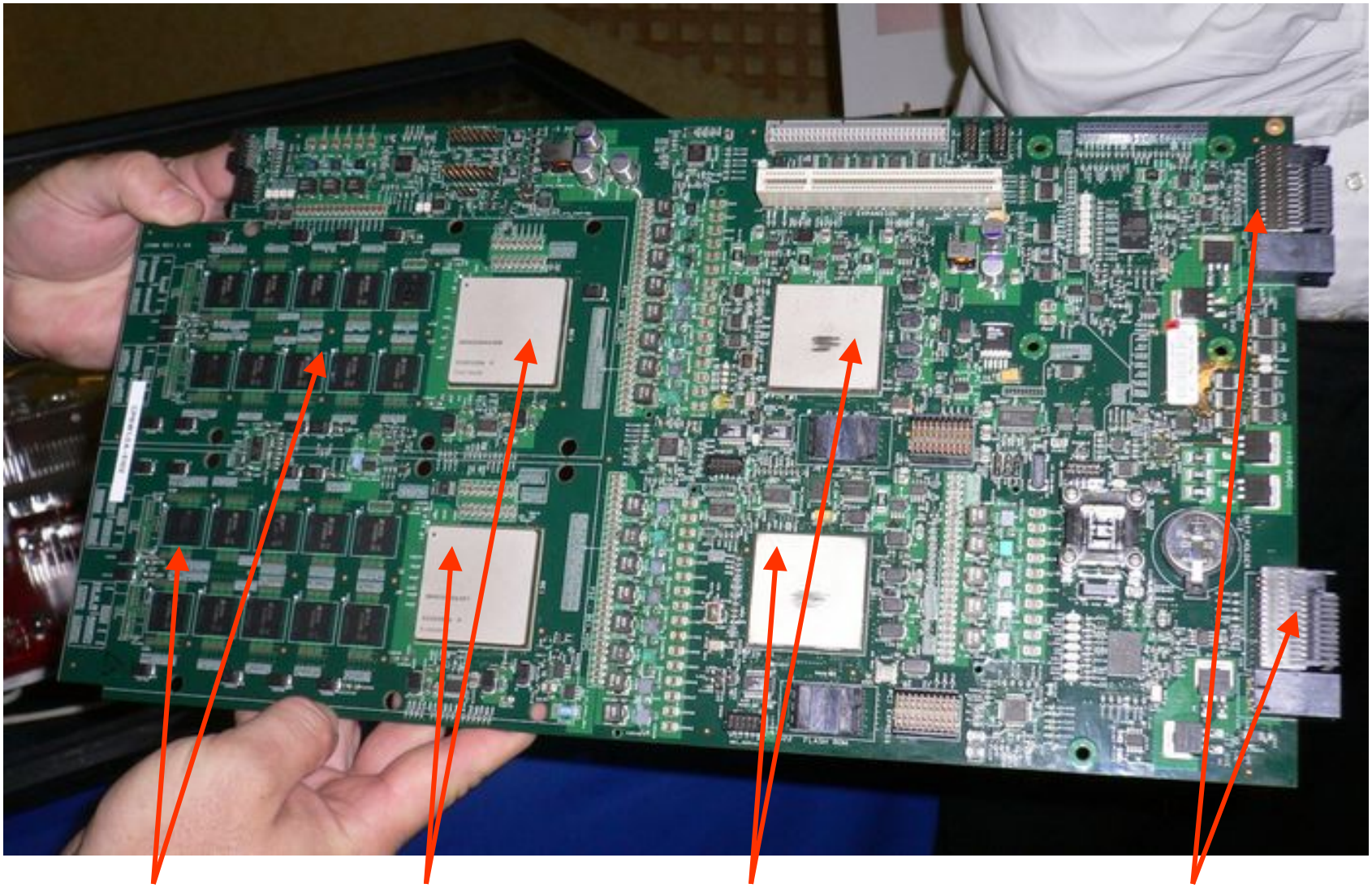
Source: Cell Broadband Engine Architecture and its first implementation – A performance view, http://www-128.ibm.com/developerworks/library/pa-cellperf/

Key Performance Characteristics

- Cell's performance is about an order of magnitude better than GPP for media and other applications that can take advantage of its SIMD capability
  - Performance of its simple PPE is comparable to a traditional GPP performance
  - its each SPE is able to perform mostly the same as, or better than, a GPP with SIMD running at the same frequency
  - key performance advantage comes from its 8 de-coupled SPE SIMD engines with dedicated resources including large register files and DMA channels
- Cell can cover a wide range of application space with its capabilities in
  - floating point operations
  - integer operations
  - data streaming / throughput support
  - real-time support
- Cell microarchitecture features are exposed to not only its compilers but also its applications
  - performance gains from tuning compilers and applications can be significant
  - tools/simulators are provided to assist in performance optimization efforts

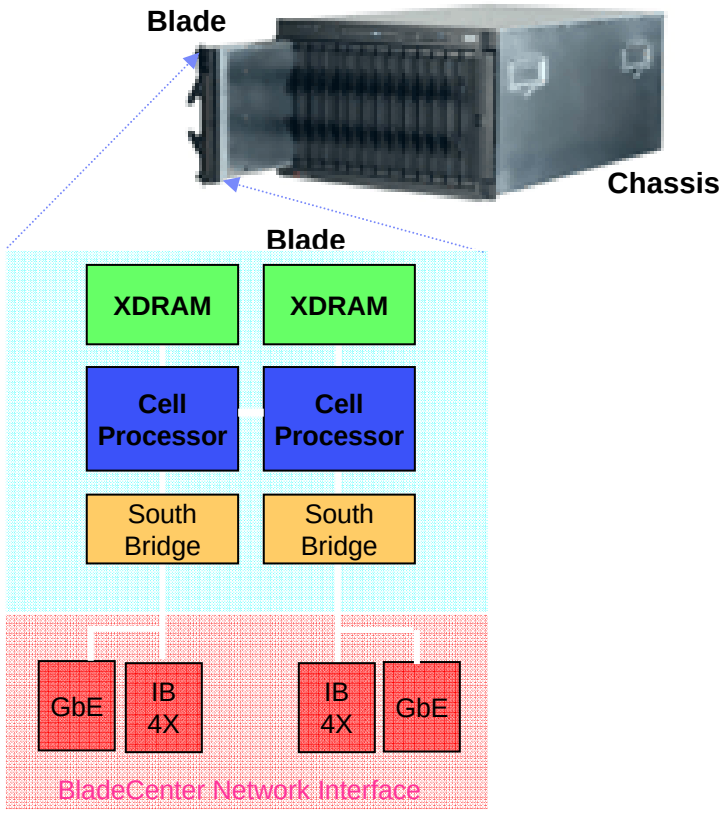
Cell Blade

The First Generation Cell Blade



Cell Blade Overview

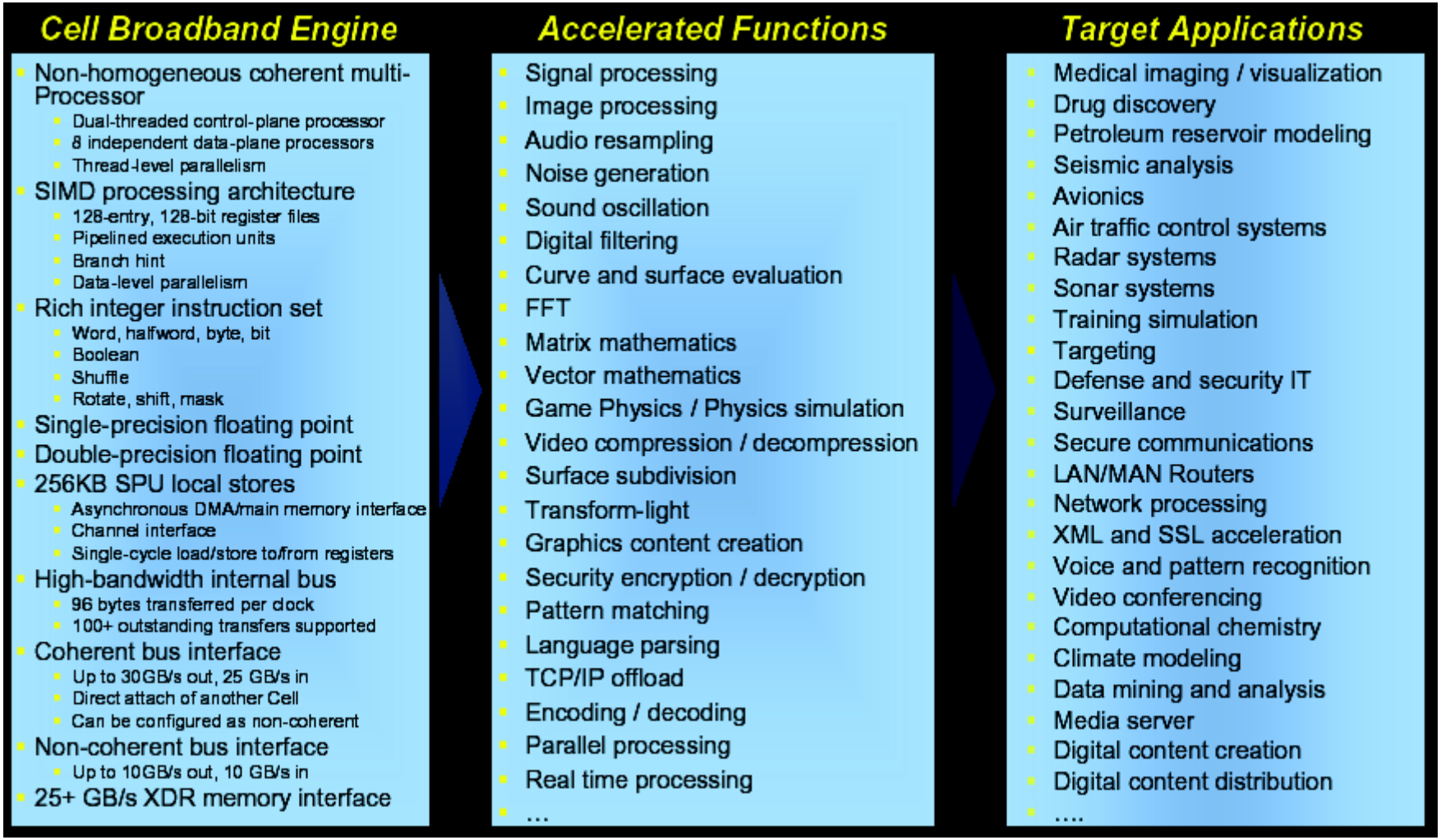
- **Blade**
  - Two Cell BE Processors
  - 1GB XDRAM
  - BladeCenter Interface (Based on IBM JS20)
- **Chassis**
  - Standard IBM BladeCenter form factor with:
    - 7 Blades (for 2 slots each) with full performance
    - 2 switches (1Gb Ethernet) with 4 external ports each
  - Updated Management Module Firmware.
  - External Infiniband Switches with optional FC ports.



- **Typical Configuration (available today from E&TS)**
  - eServer 25U Rack
  - 7U Chassis with Cell BE Blades, OpenPower 710
  - Nortel GbE switch
  - GCC C/C++ (Barcelona) or XLC Compiler for Cell (alphaworks)
  - SDK Kit on http://www-128.ibm.com/developerworks/power/cell/

Cell Application Affinity

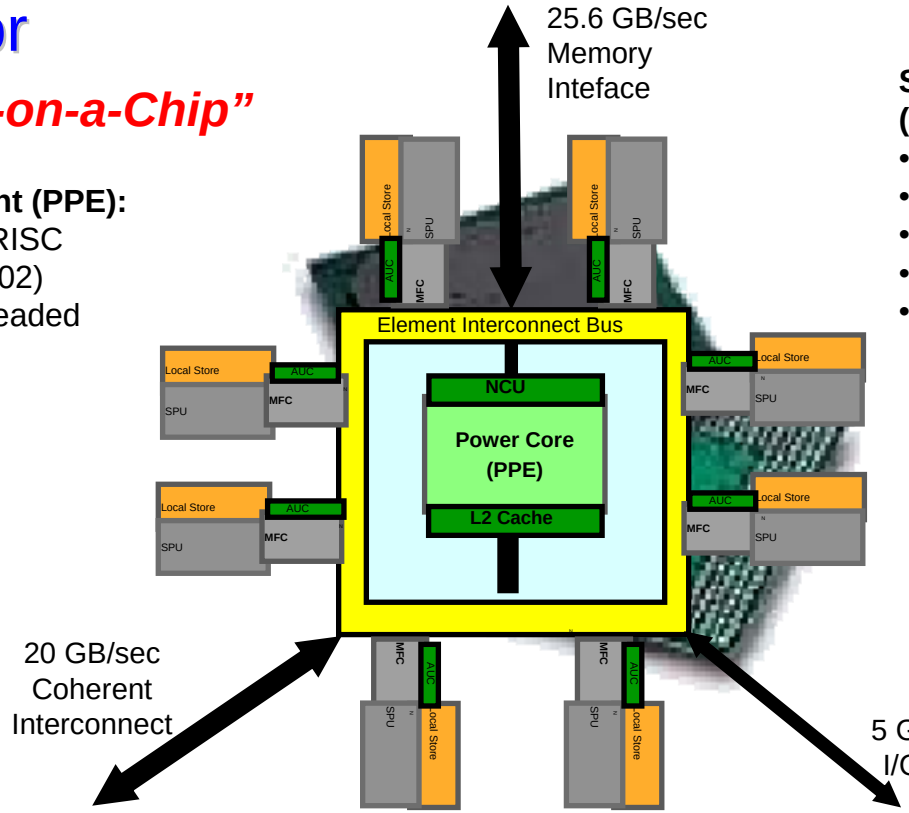
Cell Application Affinity – Target Applications



Cell Processor

"Supercomputer-on-a-Chip"

Power Processor Element (PPE):
- General Purpose, 64-bit RISC Processor (PowerPC 2.02)
- 2-Way Hardware Multithreaded
- L1 : 32KB I ; 32KB D
- L2 : 512KB
- Coherent load/store
- VMX
- 3.2 GHz



Synergistic Processor Elements (SPE):
- 8 per chip
- 128-bit wide SIMD Units
- Integer *and* Floating Point capable
- 256KB Local Store
- Up to 25.6 GF/s per SPE --- 200GF/s total *

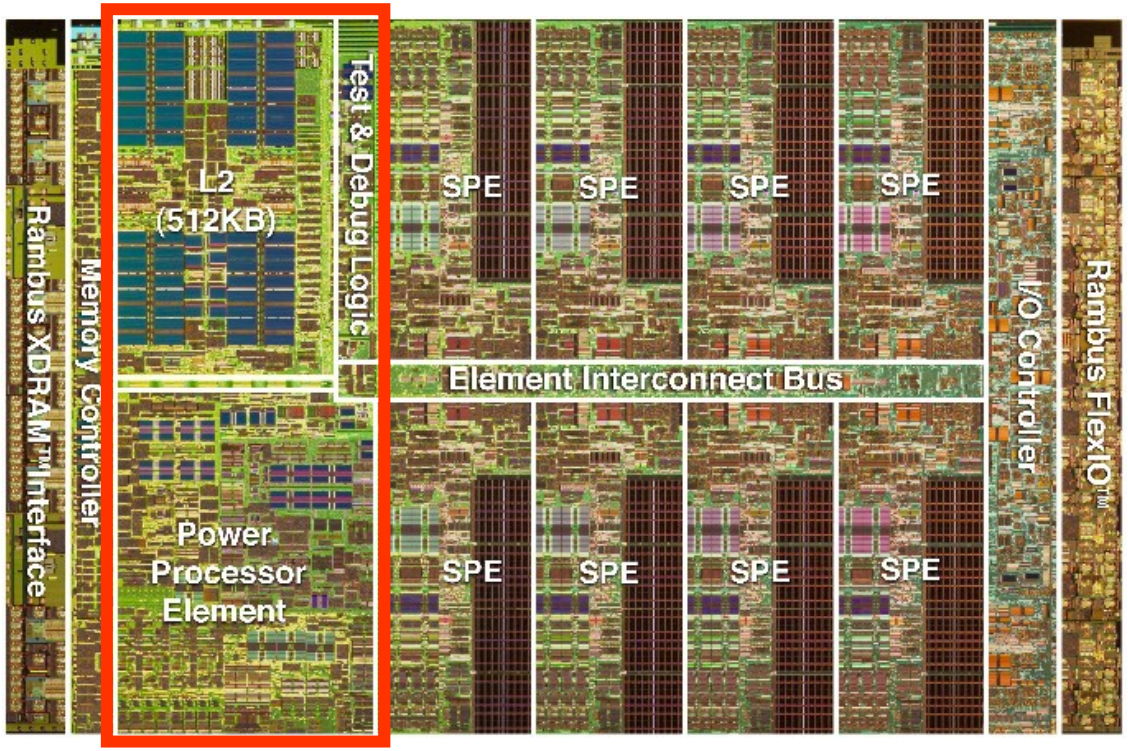
Internal Interconnect:
- Coherent ring structure
- 300+ GB/s total internal interconnect bandwidth
- DMA control to/from SPEs supports >100 outstanding memory requests

External Interconnects:
- 25.6 GB/sec BW memory interface
- 2 Configurable I/O Interfaces
  - Coherent interface (SMP)
  - Normal I/O interface (I/O & Graphics)
  - Total BW configurable between interfaces
  - Up to 35 GB/s out
  - Up to 25 GB/s in

Memory Management & Mapping
- SPE Local Store aliased into PPE system memory
- MFC/MMU controls SPE DMA accesses
  - Compatible with PowerPC Virtual Memory architecture
  - S/W controllable from PPE MMIO
- Hardware or Software TLB management
- SPE DMA access protected by MFC/MMU

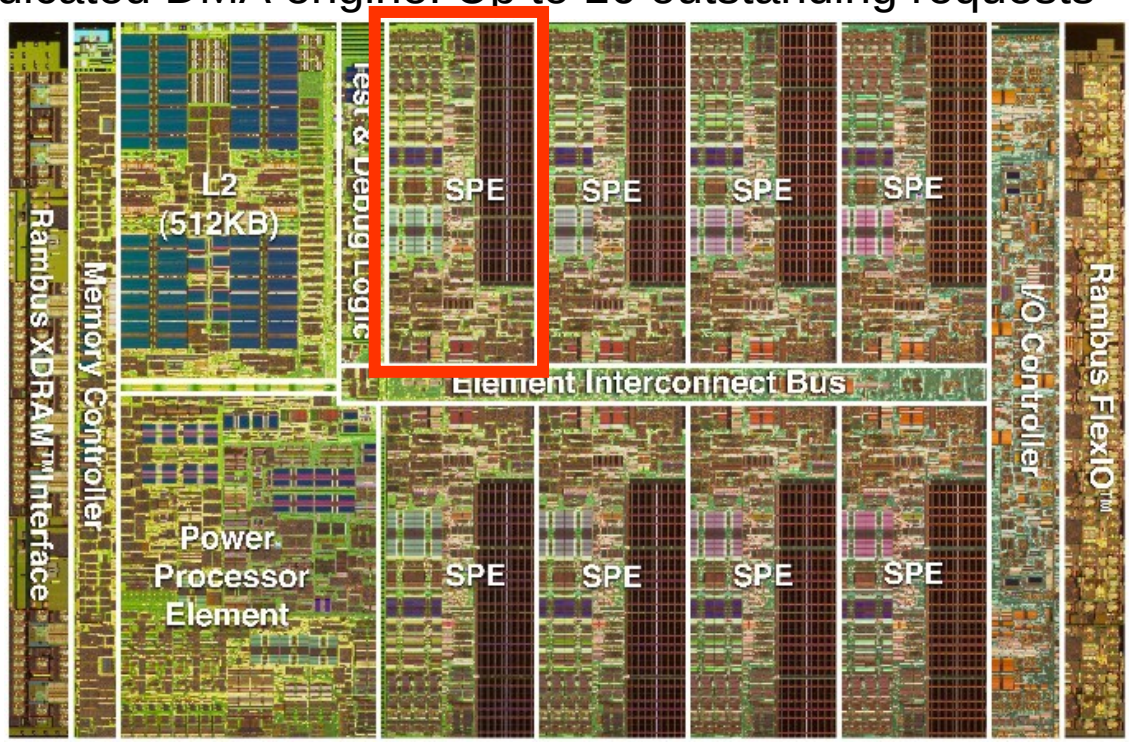
Power Processor Element

- **PPE handles operating system and control tasks**
  - 64-bit Power Architecture™ with VMX
  - In-order, 2-way hardware simultaneous multi-threading (SMT)
  - Coherent Load/Store with 32KB I & D L1 and 512KB L2



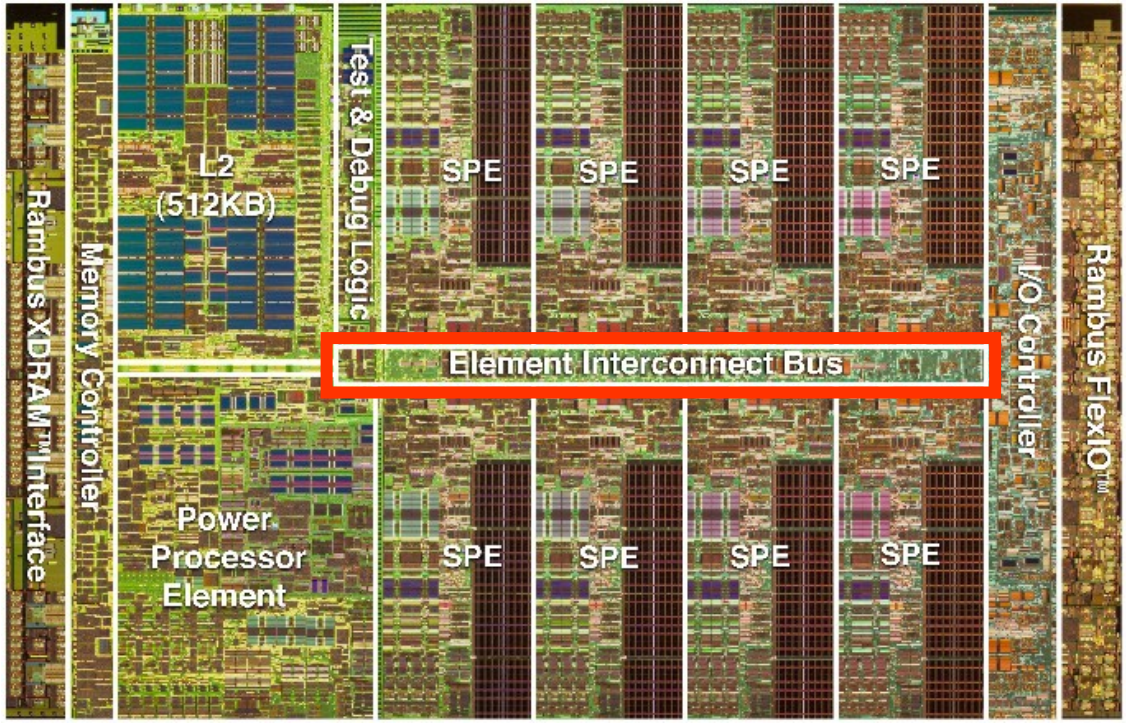
Synergistic Processor Element

- **SPE provides computational performance**
  - Dual issue, up to 16-way 128-bit SIMD
  - Dedicated resources: 128 128-bit RF, 256KB Local Store
  - Each can be dynamically configured to protect resources
  - Dedicated DMA engine: Up to 16 outstanding requests



Element Interconnect Bus

- **EIB data ring for internal communication**
  - Four 16 byte data rings, supporting multiple transfers
  - 96B/cycle peak bandwidth
  - Over 100 outstanding requests



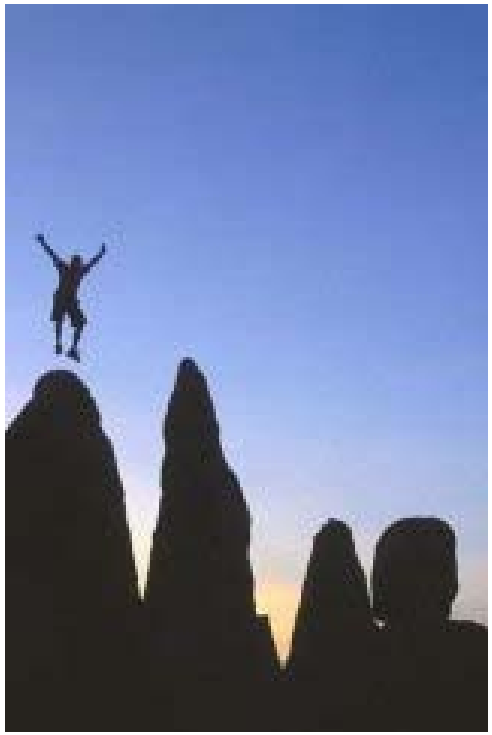
Internal Bandwidth Capability

- **Each EIB Bus data port supports 25.6GBytes/sec* *in each direction***
- **The EIB Command Bus streams commands fast enough to support 102.4 GB/sec for coherent commands, and 204.8 GB/sec for non-coherent commands.**
- **The EIB data rings can sustain 204.8GB/sec for certain workloads, with transient rates as high as 307.2GB/sec between bus units**

Despite all that available bandwidth…

* The above numbers assume a 3.2GHz core frequency – internal bandwidth scales with core frequency

Why IBM for HPC



- Advanced processor technologies, innovative system designs, and elegant packaging
- Leadership performance and top positioning in worldwide supercomputing*
- A continuum of granular solution platforms that scale from the desktop to the world's highest performance HPC systems*

Select the system sized to your needs and designed with features for:

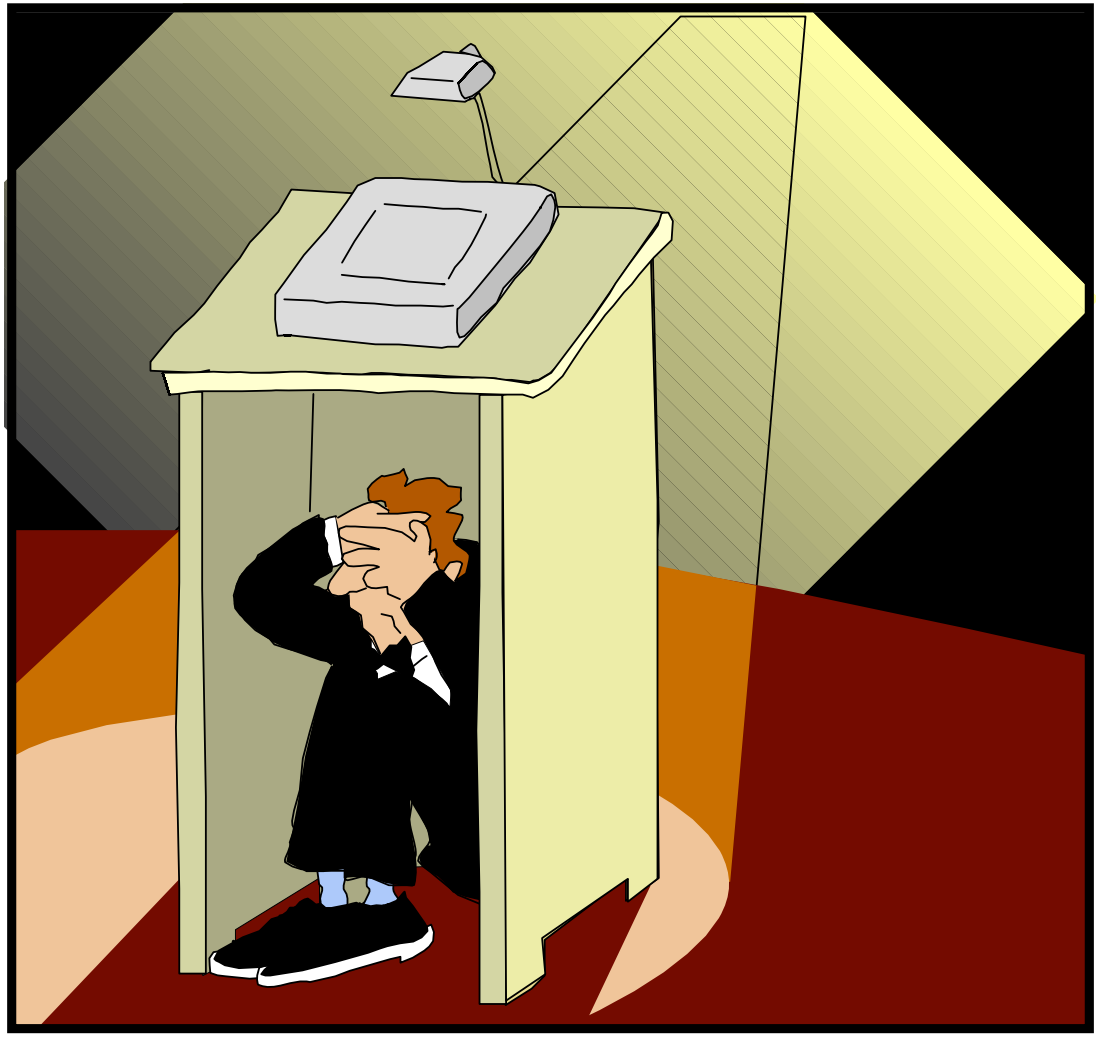
- Meeting the challenges
- High performance and versatility

of your most demanding applications

*http://www.top500.org

Questions... ?

piotr.pietrzak@pl.ibm.com



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.

Many of the features described in this document are operating system dependent and may not be available on Linux. For more information, please check: http://www.ibm.com/systems/p/software/whitepapers/linux_overview.html

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 January 19, 2006

Special notices (cont.)

The following terms are registered trademarks of International Business Machines Corporation in the United States and/or other countries: AIX, AIX/L, AIX/L(logo), alphaWorks, AS/400, BladeCenter, Blue Gene, Blue Lightning, C Set++, CICS, CICS/6000, ClusterProven, CT/2, DataHub, DataJoiner, DB2, DEEP BLUE, developerWorks, DirectTalk, Domino, DYNIX, DYNIX/ptx, e business(logo), e(logo)business, e(logo)server, Enterprise Storage Server, ESCON, FlashCopy, GDDM, i5/OS, IBM, IBM(logo), ibm.com, IBM Business Partner (logo), Informix, IntelliStation, IQ-Link, LANStreamer, LoadLeveler, Lotus, Lotus Notes, Lotusphere, Magstar, MediaStreamer, Micro Channel, MQSeries, Net.Data, Netfinity, NetView, Network Station, Notes, NUMA-Q, Operating System/2, Operating System/400, OS/2, OS/390, OS/400, Parallel Sysplex, PartnerLink, PartnerWorld, Passport Advantage, POWERparallel, Power PC 603, Power PC 604, PowerPC, PowerPC(logo), PowerPC 601, Predictive Failure Analysis, pSeries, PTX, ptx/ADMIN, RETAIN, RISC System/6000, RS/6000, RT Personal Computer, S/390, Scalable POWERparallel Systems, SecureWay, Sequent, ServerProven, SpaceBall, System/390, The Engines of e-business, THINK, Tivoli, Tivoli(logo), Tivoli Management Environment, Tivoli Ready(logo), TME, TotalStorage, TURBOWAYS, VisualAge, WebSphere, xSeries, z/OS, zSeries.

The following terms are trademarks of International Business Machines Corporation in the United States and/or other countries: Advanced Micro-Partitioning, AIX 5L, AIX PVMe, AS/400e, Chipkill, Chiphopper, Cloudscape, DB2 OLAP Server, DB2 Universal Database, DFDSM, DFSORT, e-business(logo), e-business on demand, eServer, Express Middleware, Express Portfolio, Express Servers, Express Servers and Storage, GigaProcessor, HACMP, HACMP/6000, I5/OS (logo), IBMLink, IBM TotalStorage Proven, IMS, Intelligent Miner, iSeries, Micro-Partitioning, NUMACenter, ON DEMAND BUSINESS logo, OpenPower, POWER, Power Architecture, Power Everywhere, Power Family, Power PC, PowerPC Architecture, PowerPC 603, PowerPC 603e, PowerPC 604, PowerPC 750, POWER2, POWER2 Architecture, POWER3, POWER4, POWER4+, POWER5, POWER5+, POWER6, POWER6+, Redbooks, Sequent (logo), SequentLINK, Server Advantage, ServeRAID, Service Director, SmoothStart, SP, System i, System i5, System p, System p5, System Storage, System z, System z9, S/390 Parallel Enterprise Server, Tivoli Enterprise, TME 10, TotalStorage Proven, Ultramedia, VideoCharger, Virtualization Engine, Visualization Data Explorer, X-Architecture, z/Architecture, z/9.

A full list of U.S. trademarks owned by IBM may be found at: http://www.ibm.com/legal/copytrade.shtml.
UNIX is a registered trademark in the United States, other countries or both.
Linux is a trademark of Linus Torvalds in the United States, other countries or both.
Microsoft, Windows, Windows NT and the Windows logo are registered trademarks of Microsoft Corporation in the United States and/or other countries.
Intel, Intel Xeon, Itanium and Pentium are registered trademarks or trademarks of Intel Corporation in the United States and/or other countries.
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 and/or other countries.
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.
Other company, product and service names may be trademarks or service marks of others.

Revised January 19, 2006

Notes on benchmarks and values

The IBM benchmarks results shown herein were derived using particular, well configured, development-level and generally-available computer systems. Buyers should consult other sources of information to evaluate the performance of systems they are considering buying and should consider conducting application oriented testing. For additional information about the benchmarks, values and systems tested, contact your local IBM office or IBM authorized reseller or access the Web site of the benchmark consortium or benchmark vendor.

IBM benchmark results can be found in the IBM System p5, ~ p5, pSeries, OpenPower and IBM RS/6000 Performance Report at http://www.ibm.com/servers/systems/p/hardware/system_perf.html.

All performance measurements were made with AIX or AIX 5L operating systems unless otherwise indicated to have used Linux. For new and upgraded systems, AIX Version 4.3 or AIX 5L were used. All other systems used previous versions of AIX. The SPEC CPU2000, LINPACK, and Technical Computing benchmarks were compiled using IBM's high performance C, C++, and FORTRAN compilers for AIX 5L and Linux. For new and upgraded systems, the latest versions of these compilers were used: XL C Enterprise Edition V7.0 for AIX, XL C/C++ Enterprise Edition V7.0 for AIX, XL FORTRAN Enterprise Edition V9.1 for AIX, XL C/C++ Advanced Edition V7.0 for Linux, and XL FORTRAN Advanced Edition V9.1 for Linux. The SPEC CPU95 (retired in 2000) tests used preprocessors, KAP 3.2 for FORTRAN and KAP/C 1.4.2 from Kuck & Associates and VAST-2 v4.01X8 from Pacific-Sierra Research. The preprocessors were purchased separately from these vendors. Other software packages like IBM ESSL for AIX, MASS for AIX and Kazushige Goto's BLAS Library for Linux were also used in some benchmarks.

For a definition/explanation of each benchmark and the full list of detailed results, visit the Web site of the benchmark consortium or benchmark vendor.

TPC	http://www.tpc.org
SPEC	http://www.spec.org
LINPACK	http://www.netlib.org/benchmark/performance.pdf
Pro/E	http://www.proe.com
GPC	http://www.spec.org/gpc
NotesBench	http://www.notesbench.org
VolanoMark	http://www.volano.com
STREAM	http://www.cs.virginia.edu/stream/
SAP	http://www.sap.com/benchmark/
Oracle Applications	http://www.oracle.com/apps_benchmark/
PeopleSoft - To get information on PeopleSoft benchmarks, contact PeopleSoft directly	
Siebel	http://www.siebel.com/crm/performance_benchmark/index.shtm
Baan	http://www.ssaglobal.com
Microsoft Exchange	http://www.microsoft.com/exchange/evaluation/performance/default.asp
Veritest	http://www.veritest.com/clients/reports
Fluent	http://www.fluent.com/software/fluent/index.htm
TOP500 Supercomputers	http://www.top500.org/
Ideas International	http://www.ideasinternational.com/benchmark/bench.html
Storage Performance Council	http://www.storageperformance.org/results

Revised January 19, 2006