IBM Systems and Technology Group

And we also have hardware . . .

17th Machine Evaluation Workshop
Daresbury 6 Dec 2006

Crispin Keable
crispin.keable@uk.ibm.com

Agenda

- IBM Deep Computing systems
- Research projects
  - Blue Brain
  - Heart Modelling
  - Road Runner
  - Collaborations
- Future directions
  - Systems
  - HPCS

What is Deep Computing?

- As deeper understanding of physics and biology lifts the human spirit,
- As better physical and biological models are devised,
- As volumes of experimental data are collected,
- As the Internet grows to encompass more people and institutions,
- As pervasive devices connect to the network,
- As more business is done online,
  - *A wealth of data is becoming available in digital form.*

- Finding the value buried in that data will be an increasingly powerful tool
  - For business and for society
  - Extending the scale of simulations to more accurately model phenomenon

- **Deep computing** combines several techniques to solve extremely complex problems in this sea of digital data:
  - Advanced mathematics
  - Domain-specific knowledge
  - Specialized software
  - Powerful hardware

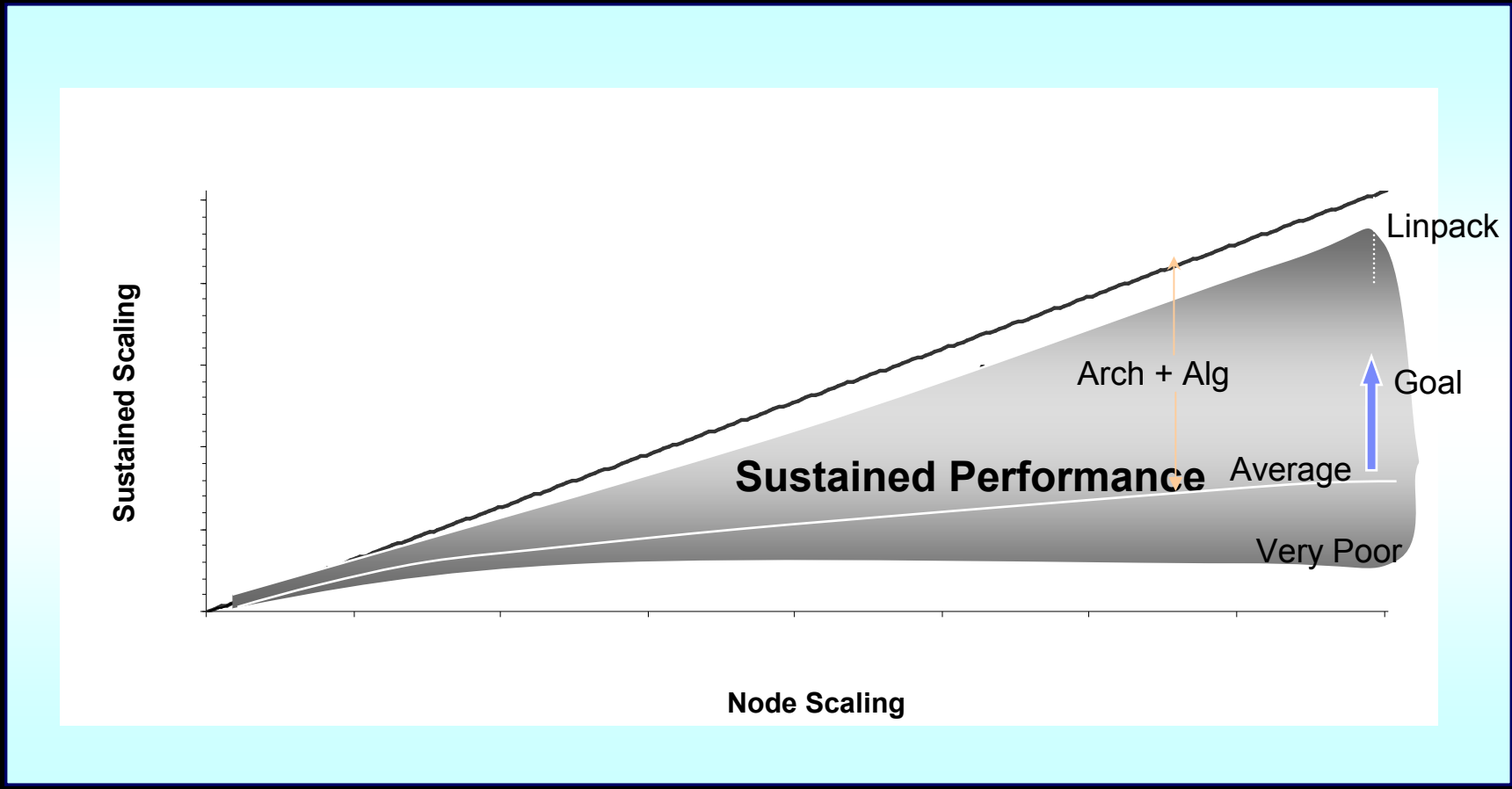
Why Does IBM Do Deep Computing?

- **Good for our customers:**
  - Biological and Life Sciences
  - Pharmaceutical Discovery
  - Materials and Manufacturing
  - Automotive and aerospace design
  - Environment and Energy
  - Finance
    - Inventory optimization
    - Value at risk
    - Portfolio and trading optimization
  - Security
    - Homeland security
    - Military simulations

- **Good for the world:**
  - Stop epidemics
  - Invent new drugs
  - Understand nature

- **Good for IBM:**
  - Good business
  - Drives innovation
  - Better design of our chips, servers, software

Sustained Performance vs. Peak



Key Problem: Gap between peak and sustained continues to widen

Goal: Understand reasons for the gap and attack the problems

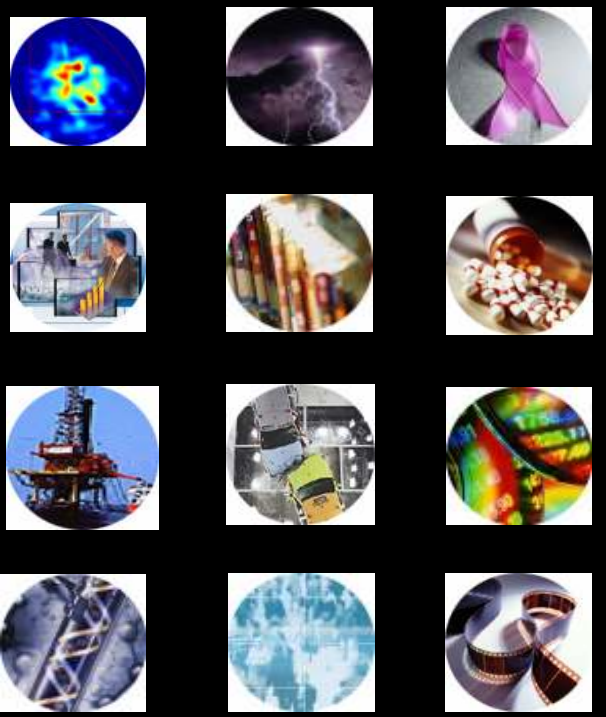
Deep Computing Innovation
Addressing Challenges Beyond Computation

- **System Design**
  - Scalability
  - Packaging & density
  - Network infrastructure
  - Power consumption & cooling
- **Software**
  - System management
  - Security
  - Software integration
  - Programming models & productivity
- **Data**
  - Data management
  - Archival & compliance
  - Performance & reliability
  - Simulation & modeling
  - Data warehousing & mining
  - Capacity management & virtualization
- **Economics**
  - Hybrid financial & delivery models
  - Software licensing

IBM's Deep Computing Strategy

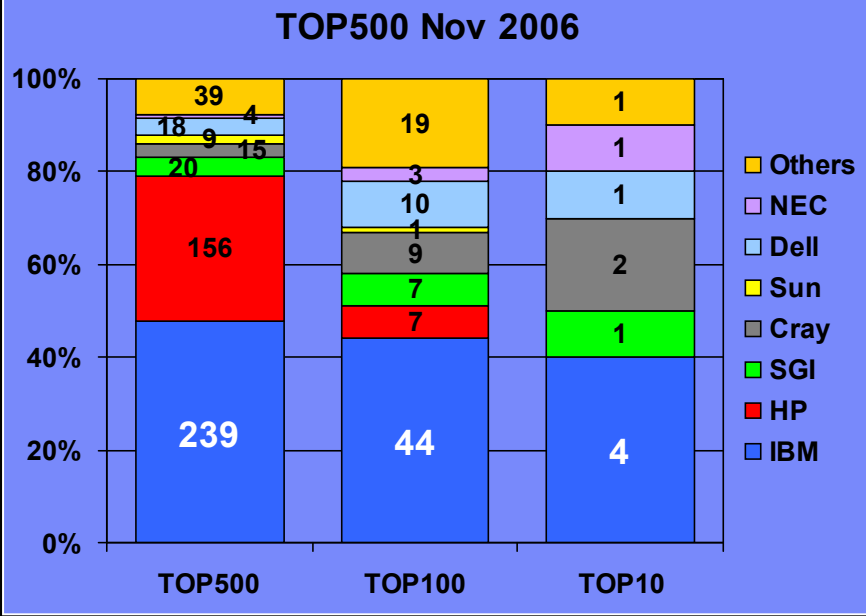
Solving Mission-Critical, Compute and Data Intensive Problems More Quickly at Lower Cost, for Strategic Value

- Aggressively evolve POWER-based Deep Computing product line
- Develop advanced systems based on loosely coupled clusters
- Research and overcome obstacles to parallelism and other revolutionary approaches to supercomputing
- Increase means of accessing supercomputing with on demand capabilities



IBM Supercomputing Leadership

Semiannual independent ranking of top 500 supercomputers in the world



IBM is clear leader ...

- ✓ #1 System – DOE/LLNL - BlueGene/L (280.6 TF)
- ✓ Most entries on TOP500 list with 239 (47.8%)
- ✓ Most installed aggregate throughput with over 1,763 Teraflops (49.8%)
- ✓ Most in TOP10 with 4 systems (40%)
- ✓ Most in TOP20 with 6 systems (30%)
- ✓ Most in TOP100 systems with 44 (44%)
- ✓ Most Commodity Clusters with 171 of 359 (47.6%)
- ✓ Fastest machines in USA (BG/L) and Europe (MareNostrum)

"IBM remains the dominant vendor of supercomputers with almost half of the list carrying its label. "

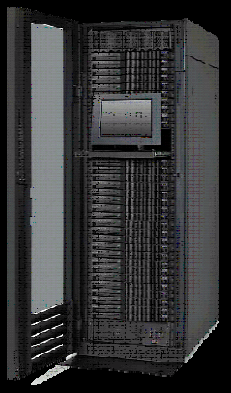
Source: www.top500.org

Deep Computing Systems

Deep Computing Portfolio

- ***Systems***
  - Servers, blades, clusters, workstations
    - POWER, Intel, AMD Opteron
    - Linux, Unix, Windows
    - High performance interconnects
  - Special-purpose systems and accelerators
    - Blue Gene, Cell
    - Gov't & research partnerships (e.g., ASC, Roadrunner)
  - Storage systems and virtualization
- ***Software & Tools***
  - System & data management (e.g., CSM, GPFS)
  - Compilers, schedulers, libraries, tools
  - Grid & on demand middleware
- ***Solutions & Services***
  - Deep Computing Capacity on Demand
  - Deep Computing Visualization
  - Industry-specific solutions
  - Engineering & technology services
  - IBM Global Financing

High Performance Switch



Clusters (1350)

P5 575

BladeCenter

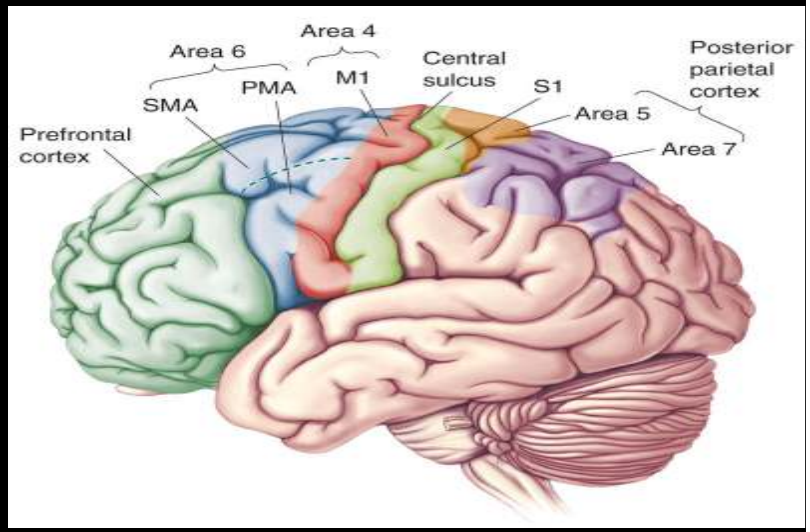
IntelliStation

Blue Gene

Collaborative Research Projects

Blue Brain

The first objective of the **Blue Brain** is to create a cellular level, software replica of the Neocortical Column for real-time simulations.



Blue Brain will search for:

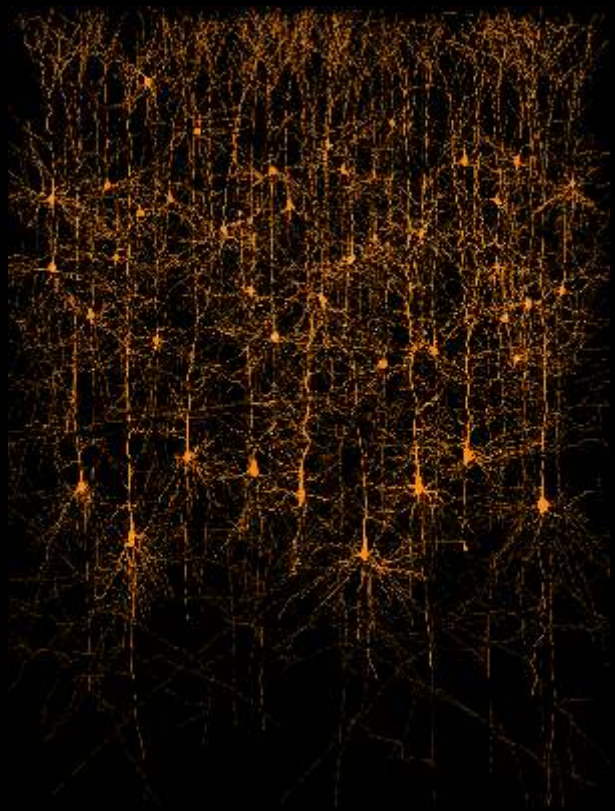
- New insights into how human beings think and remember.
- How specific defects in our circuitry may contribute to autism, schizophrenia and Parkinson's.

With Blue Brain, research that used to require several years of lab work can be done in a matter of days or minutes – using **Blue Gene**.

A project of Ecole Polytechnique Fédérale de Lausanne.

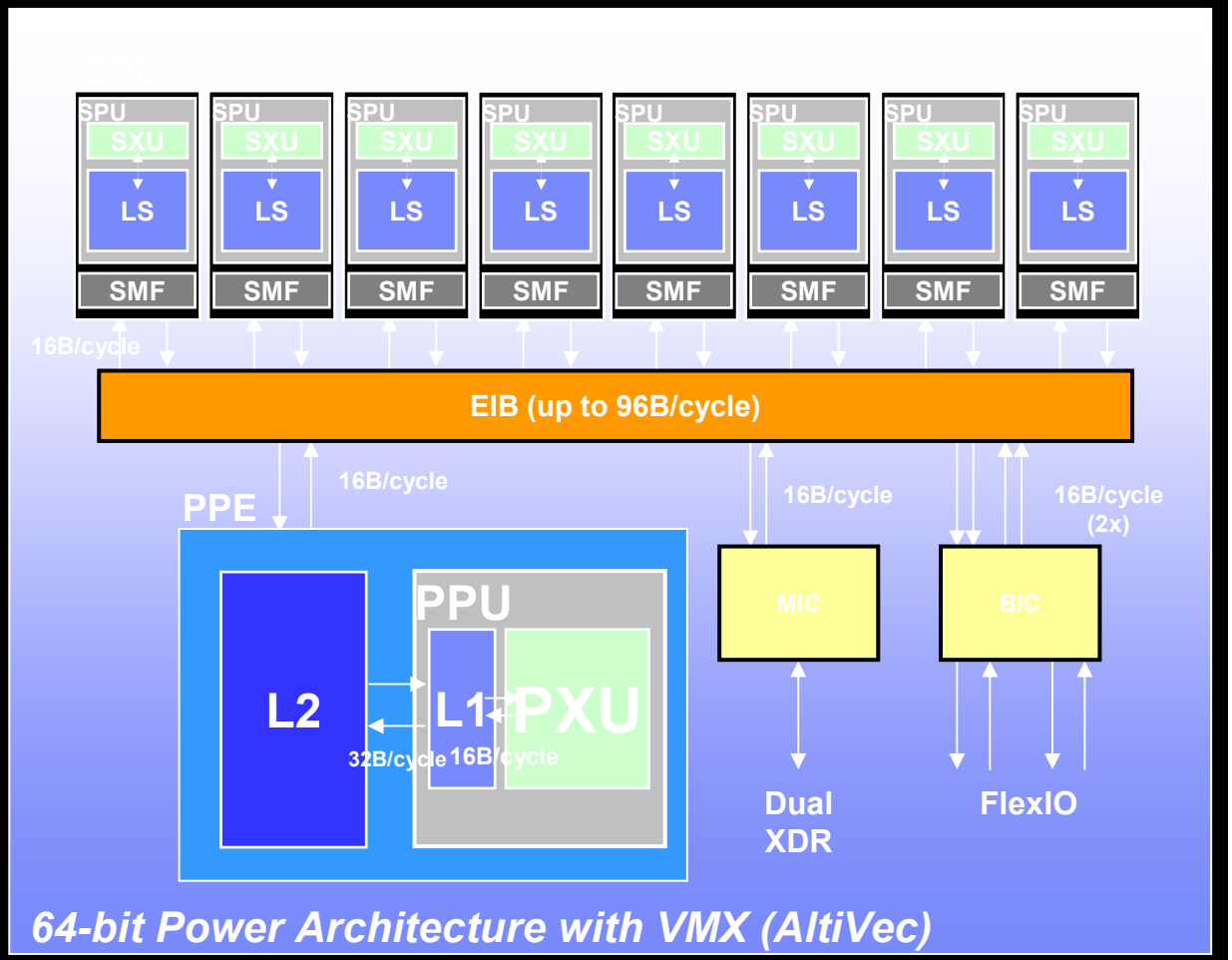
Blue Brain research phases

- software replica of a column of the neocortex
  - 85% of brains total mass
  - Required for language, learning, memory and complex thought
  - the essential first step to simulating the whole brain
  - provides the link between genetic, molecular and cognitive levels of brain function.
- subsequent phases will be to expand the simulation to include circuitry from other brain regions and eventually the whole brain.

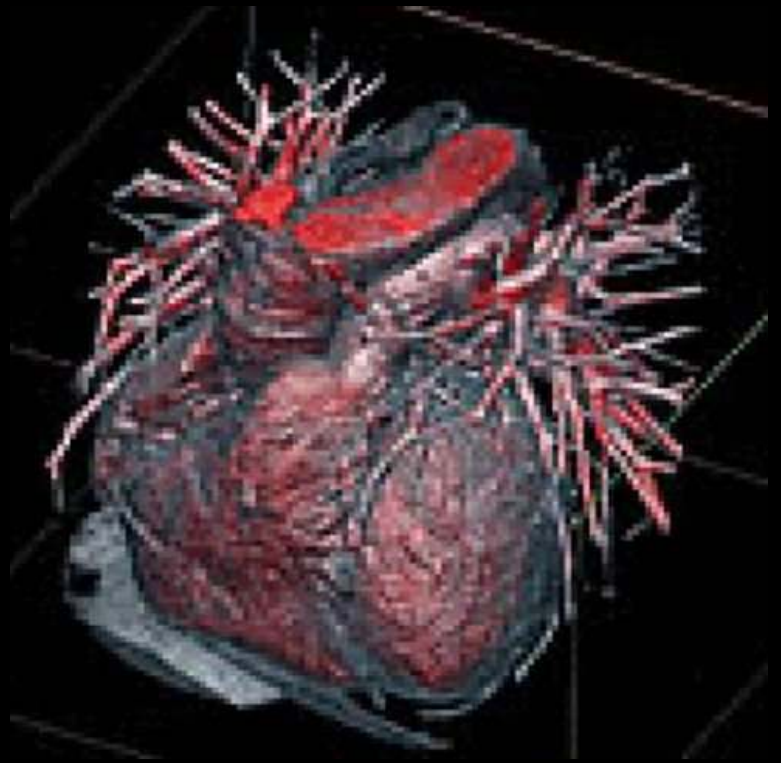


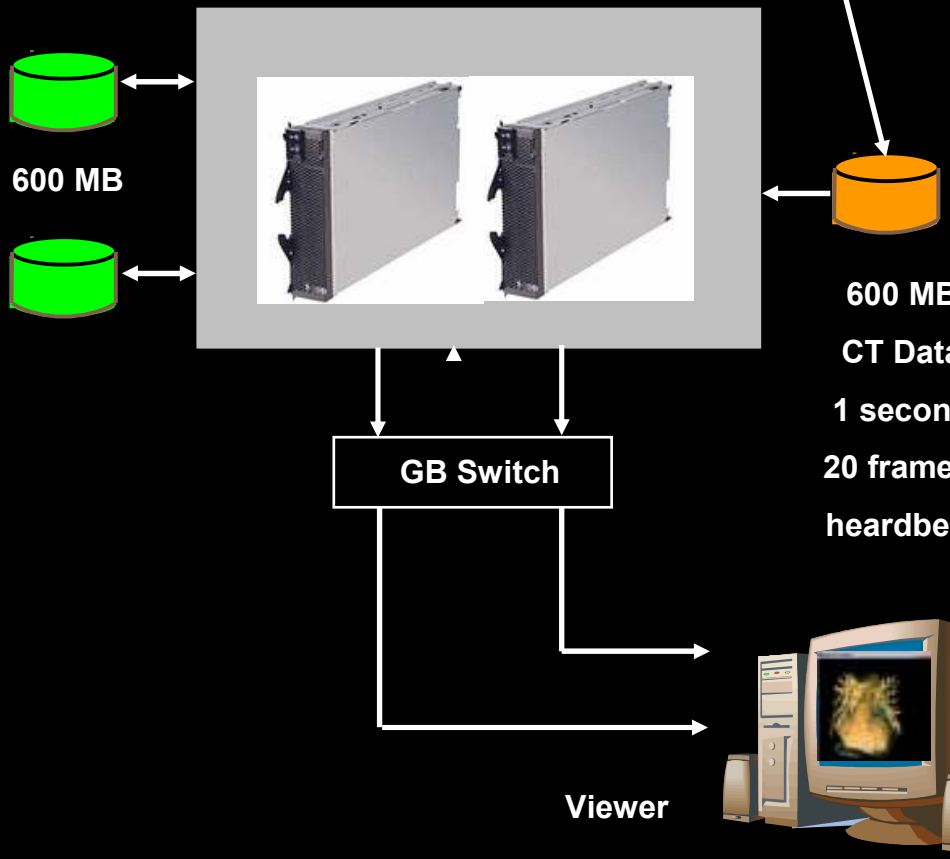
Cell BE Multi-Core System-on-Chip

- Power Processor Element
  - *Control tasks*
- Synergistic Processor Element
  - *Data-intensive tasks*
- Memory Interface Controller
  - *Rambus XDR memory*
- Bus Interface Controller
  - *Rambus FlexIO*
- Element Interconnect Bus
  - *Data movement*



Research Collaboration: Heart Modelling

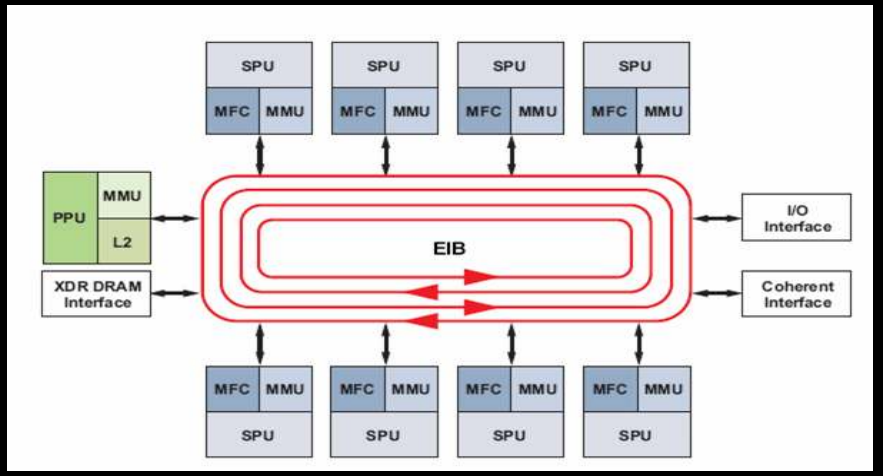




Multi-Core Acceleration Capability: Cell Broadband Engine™

Cell Broadband Engine

- **"Supercomputer & Network on a Chip"**
- 1 PPE + 8 Synergistic Processing Elements (SPE) cores
- Element Interface Bus (EIB) @ 300+GB/s
- 1 Cell BE Blade = 2 Cell BE Chips = 16 SPEs



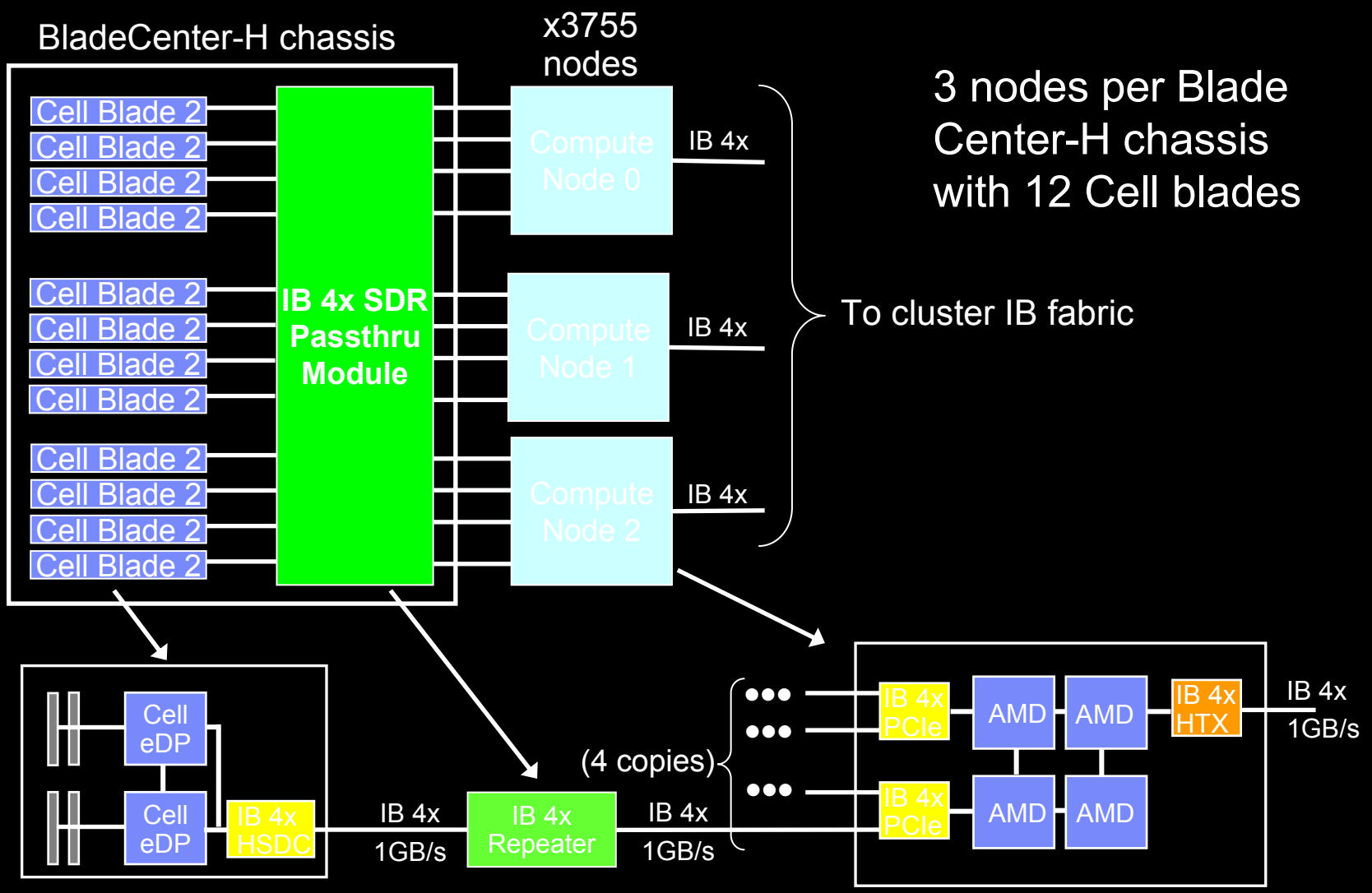
Revolutionary Hybrid Supercomputer at Los Alamos National Laboratory Will Harness Cell BE Chips and AMD Opteron™ Processor Technology

x86 Linux Master Cluster

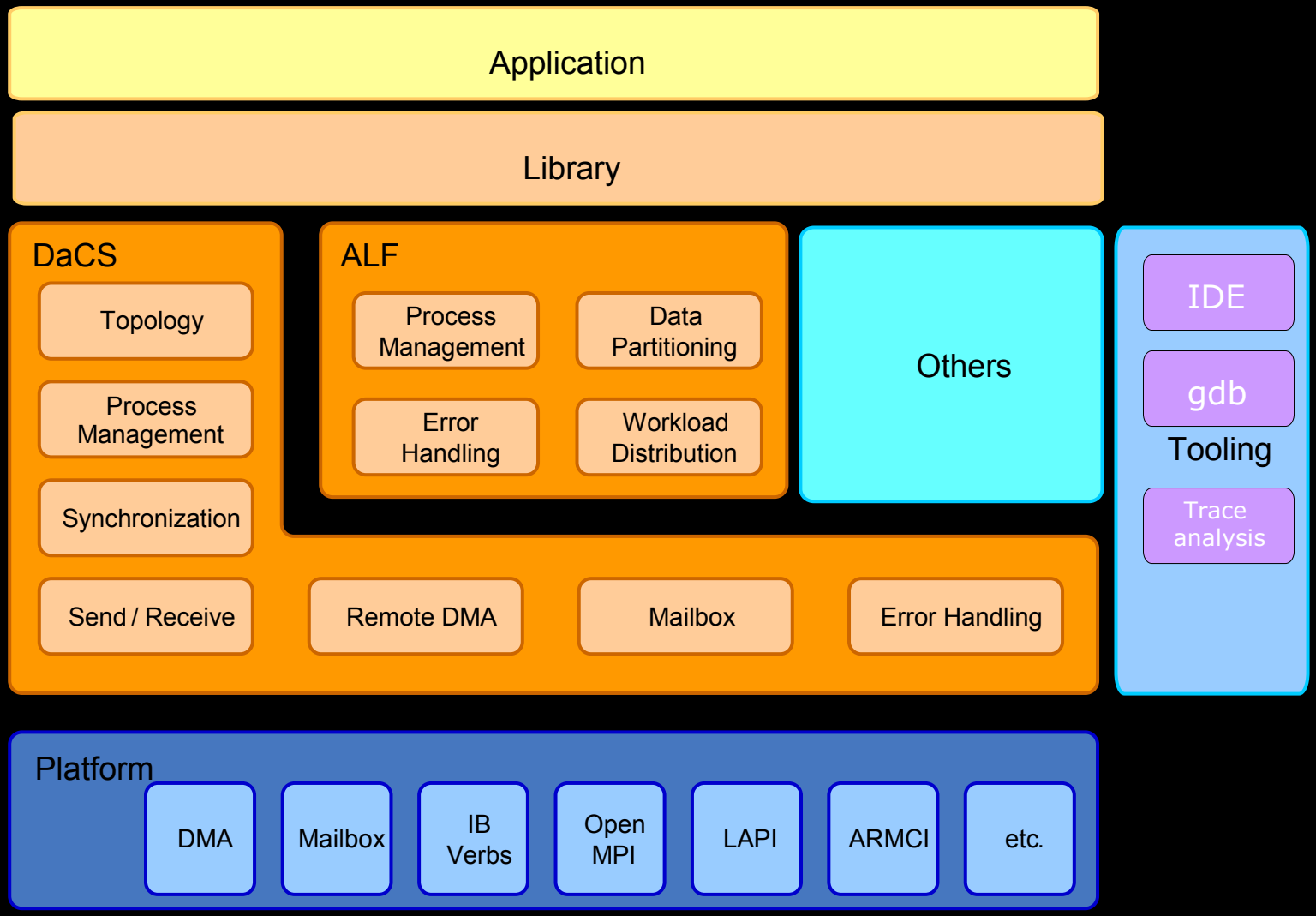
Cell BE Accelerator Cluster

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc.
Development collaboration between Sony, Toshiba and IBM.

Cell Blade Attachment



Accelerator Library Framework & Data and Communication Synchronization Library



IBM Research Collaborations

- IBM-HPCX Life Sciences outreach programme
  - Co-funded post-doctoral research positions porting & optimisation for new Life Sciences projects onto HPCX
  - Increased use of HPCX by BBSRC-funded projects from 0.6% in 2004 to 2.6% in 2005 (target >1%)
- University of Wales Swansea – Blue-C & Inst. Of Life Science
  - Clinical Genomics DB2 platform for Chernobyl Tissue Bank
  - Health Informatics – data mining of anonymised patient records
  - HPC workshops: parallelisation and porting of codes

IBM software research

- Innovations in GPFS
  - Policy managed disk-tape migration
  - More platforms (Windows, Solaris)
  - Infiniband – uDAPL RDMA data transfers
  - Scaling, management and performance improvements
- Low Latency Ethernet
  - 3.0 µs MPI latency
  - 465 MB/s unidirectional, 905MB/s bidirectional
  - ‘Reasonable’ message lengths

Collaborations: Computational Biology

Systems Biology

- **EPFL**
  Blue Brain
- **Yale, Ontario Cancer Center**
  RNA interference
- **University of Pennsylvania**
  Genetic epidemiology of breast cancer
- **Columbia University**
  DREAM project on reverse engineering
- **University of Groeningen**
  Mass spectroscopy of endometrial cancer
- **University of Illinois at Chicago**
  Sarcomere modeling
- **Institute for Advanced Study**
  Biology of p53
- **Translational Genomics Institute and MPI**
  Gene expression analysis
- **Institute for Systems Biology**
  MPSS analysis of macrophage stimulation

Blue Gene

- **EPFL**
  Blue Brain
- **Columbia, UIUC, Stanford**
  Protein Folding
- **NIH, Wabash, Baylor, UNC, Stony Brook, U Arizona**
  Protein – membrane systems
- **Mayo Clinic**
  Drug docking

Pharmaceutical Discovery

- **Mayo Clinic**
  Drug docking
- **Scripps Florida**
  Protein modeling and interactions
- **Singapore GIS, Cambridge University**
  RNA interference
- **University of Queensland, JGI-DOE**
  Microbial genome analysis

Information Based Medicine

- **National Geographic**
  Genographic Project
- **Hospital for Special Surgery (Cornell/Weil Med Center)**
  HealthMiner
- **Neurospin**
  Imaging

Collaborations: Emerging Applications

Cell

- **Saarland University**
  OpenRT ray tracer
- **Trinity College, Dublin**
  Next generation character animation
- **Lawrence Livermore National Lab**
  Speech recognition
- **Pacific Northwestern National Lab**
  Video analytics, Science applications
- **Univ of Illinois, MIT, Stanford**
  Programming models for emerging apps
- **University of Maryland**
  Bioinformatics
- **UPC Barcelona**
  Pattern matching in strings

Science, Blue Gene

- **Columbia, NIH, Wabash College, Baylor University, UNC, SUNY Stony Brook, U of Arizona, Stanford, University of Illinois, UC San Francisco**
  Blue Matter, Protein folding
- **EPF-Lausanne**
  Simulation of Neocortex and post-CMOS devices
- **Juelich, MIT, Boston University**
  QCD
- **Juelich & ETH-Zurich**
  Bionanotechnology
- **Lawrence Livermore National Lab**
  Simulation of hydrogen-storage materials
- **SanRaffaele**
  Discovery of anti-inflammatory drugs
- **ETH-Zurich**
  Finance modeling
- **UC Berkeley**
  Biochemical reactions
- **San Diego Supercomputing Center**
- **Argonne National Laboratory**

Collaborations: Software / Visualization

Blue Gene

- **Argonne National Laboratory**
  MPI, parallel file systems
- **Technical University of Vienna, CMU**
  FFT library
- **University of Edinburgh**
  LAPACK, ScaLAPACK
- **Northwestern University**
  MPI-IO
- **UPC Barcelona**
  Performance tools, MPI
- **PNNL, Ohio State University**
  Global Arrays
- **University of Illinois**
  Higher level programming models (HTA/Matlab)
- **Purdue University**
  UPC

ACTC

- **Juelich**
  KOJAK performance tool
- **University of Maryland**
  Memory simulators
- **NERSC**
  Memory simulation

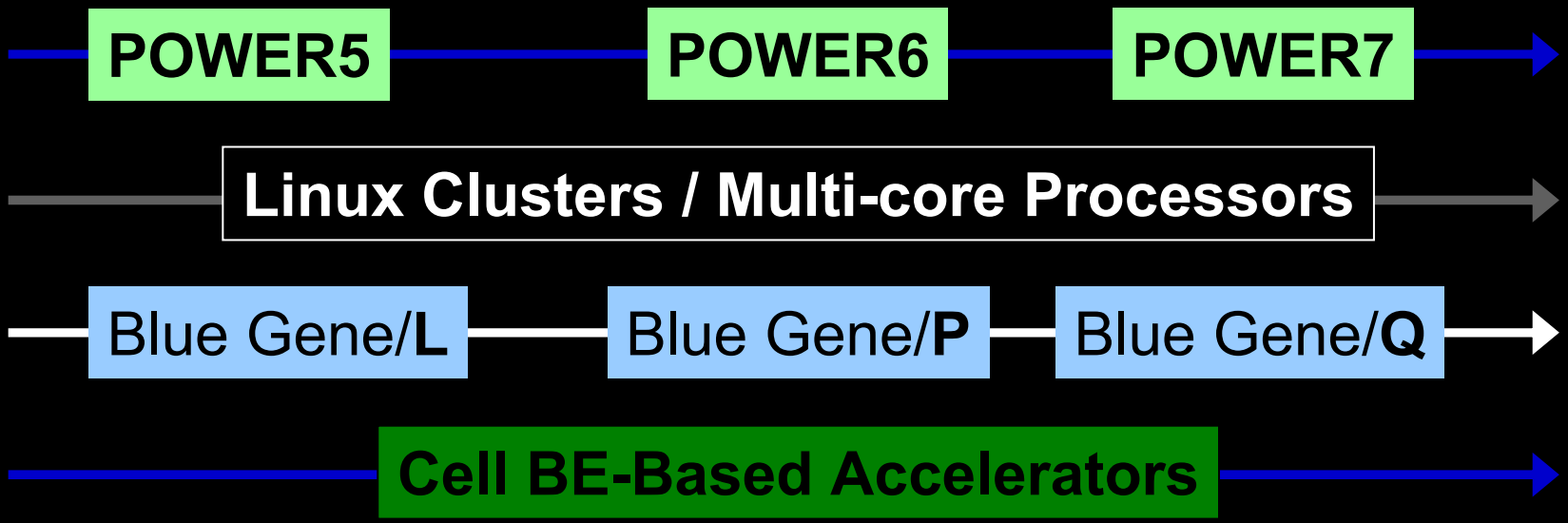
Deep Computing Visualization

- **University of Utah**
  Graphics & Visualization
- **Columbia University**
  Robotics & Visualization
- **NYU, Dassault Systemes**
  Subdivision-surface algorithms
- **UFMG / Brazil Deep Computing Competency Center**
- **Memorial University of Newfoundland, Landmark Center for Visualization**

Deep Computing directions

Deep Computing Conceptual Roadmap

- Petaflop systems with ultra scalability
  - Hybrid architectures
  - System level accelerators
  - Mature software stack
- Rich application infrastructure



HPC Conceptual Roadmap

- **PERCS** is IBM's HPCS project – PERCS *Details on next slide.*
- **POWER 7** is:
  - The system for the PERCS technologies.
  - The target system for IBM's phase 3 bid for HPCS.
- HPCS requirements:
  1. Performance
  2. Programmability – including development time
  3. Portability;
  4. Robustness
- *All in a general-purpose, commercially viable offering*
- All four of these requirements can be best satisfied by the POWER series

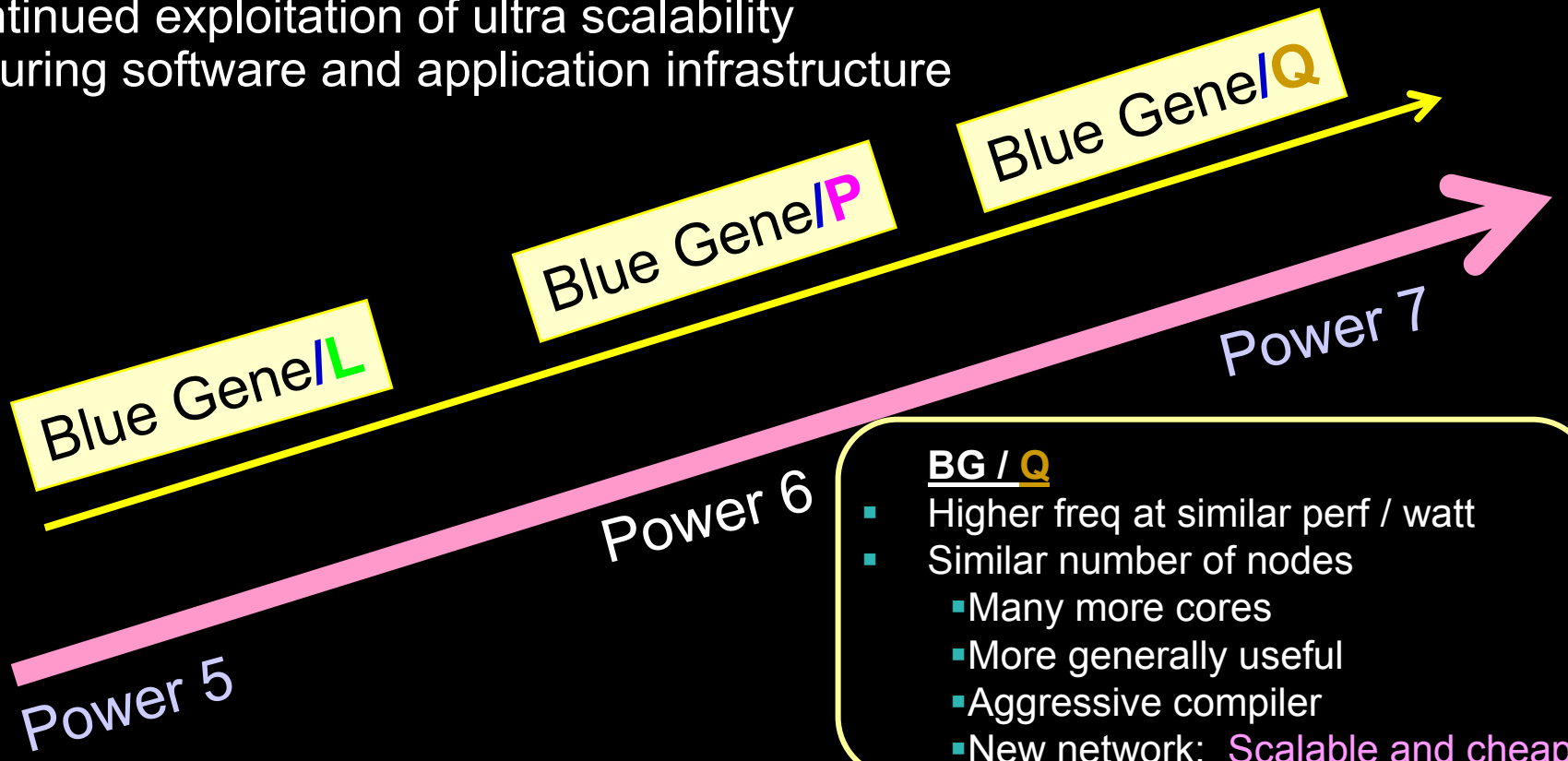
POWER also satisfies the requirements of being a general-purpose system with an established eco system and support structure, making it more suitable to the productivity targets of HPCS

PERCS – Productive Easy-to-use Reliable Computing System

- **PERCS** is innovative design for an economically successful server line –
  - Touching every aspect of the system from technology to application.
- The **PERCS** software environment is rich in tools and productivity enhancements.
- **PERCS** performance targets:
  - 10x to 1000x on HPC Challenge benchmarks
  - 60x for storage
- **PERCS** innovations:
  - Highly configurable processor chip.
  - Switchless architecture for reduced latency, cost and power.
  - Ultra-dense packaging for improved performance and reduced cost. Optics on the QCM.
  - Rich tool environment for enhanced productivity (PTP, HPCT, CPO, Rationale)
  - X10 programming language
  - Ultra-scale file system delivering 60X performance
  - Hardware support for PGAS, RDMA, collectives
  - Extensive features for improved memory performance
  - And much more

HPC conceptual roadmap

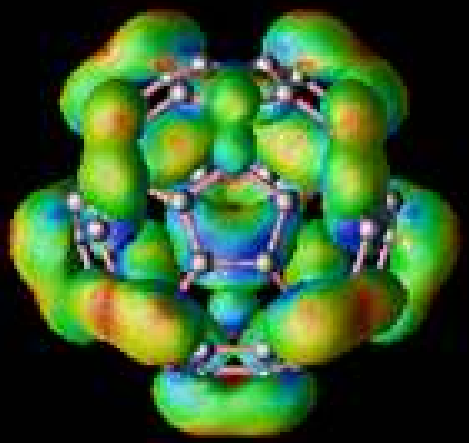
- **Blue Gene** Road Map
  - BG/**P** will reach **1 PetaFlop** peak between now and 2010
    Extensions / improvements of BG/**L**
  - BG/**Q** aimed at **10 PetaFlops** between 2010 and 2012
- Continued exploitation of ultra scalability
- Maturing software and application infrastructure



BG / Q
- Higher freq at similar perf / watt
- Similar number of nodes
  - Many more cores
  - More generally useful
  - Aggressive compiler
  - New network: Scalable and cheap

In the end, it's not about the technology. It's what you do with it that counts. IBM will:

- Continue to innovate across the whole systems stack to deliver leadership performance and usability.
- Help solve problems that are currently intractable or not cost-effective.
- Work with customers to accelerate discovery in science, engineering, and business .



Thank you

Questions?

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.

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), 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, DS4000, DS6000, DS8000, e-business(logo), e-business on demand, eServer, Express Middleware, Express Portfolio, Express Servers, Express Servers and Storage, GigaProcessor, HACMP, HACMP/6000, IBM TotalStorage Proven, IBMLink, 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.

Special Notices (Cont.)

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc.
Rambus is a registered trademark of Rambus, Inc.
XDR is a trademark of Rambus, Inc.
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 trademarks of Microsoft Corporation in the United States, other countries or both.
Intel, Intel Xeon, Itanium and Pentium are trademarks or registered 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.