IBM

Addressing HPC Challanges Through Innovation

Maciej Remiszewski

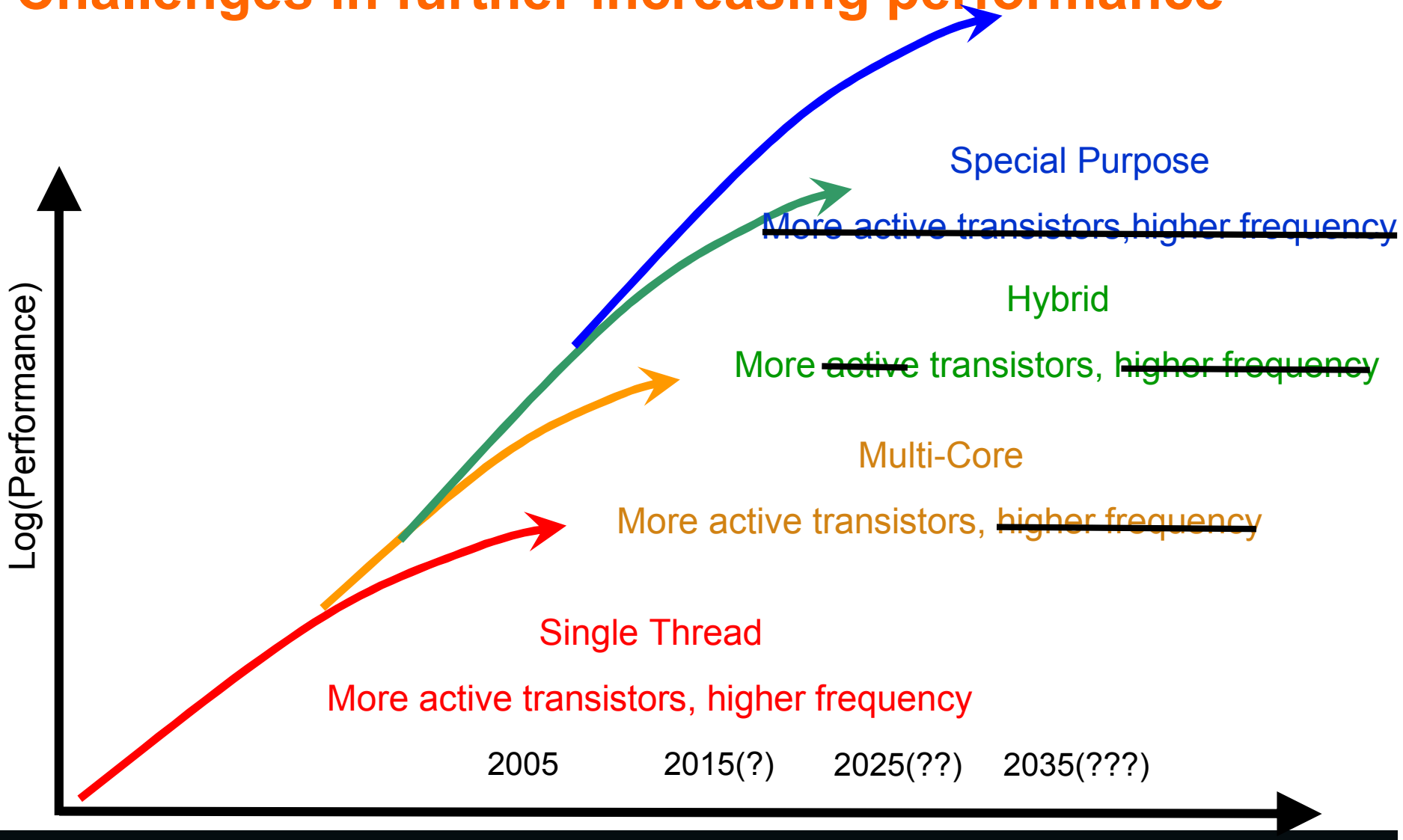
STOP
STUDYING THE ROAD AHEAD.
START
MAKING PROGRESS.

AGENDA

- Current and future trends in HPC
  - ...why do we need more performance & capacity?
- IBM Deep Computing technologies
  - ...addressing challenges through innovation
- The Nautilus Project
  - ...building the greenest supercomputer in the World

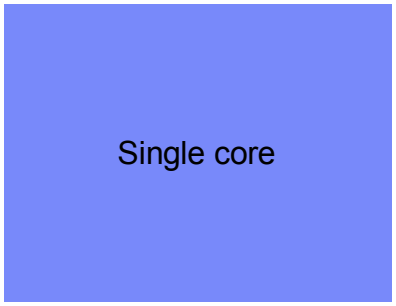
HPC application areas

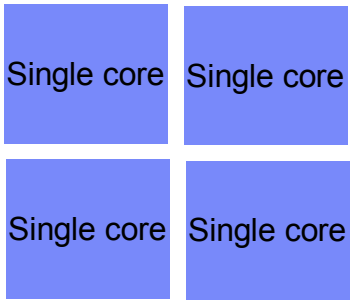
Challenges in further increasing performance

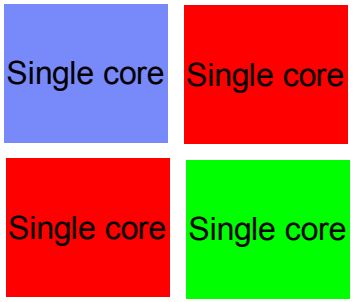


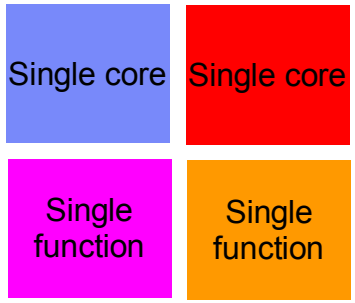
Current and future trends in HPC

How to maintain portability, programmability, architecture?



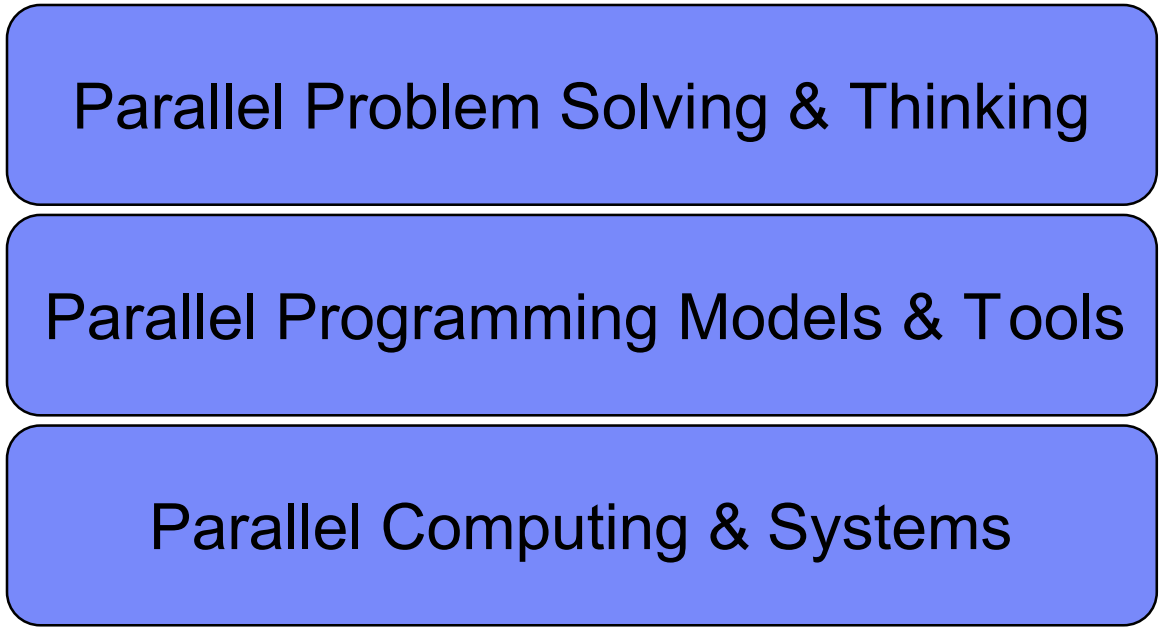






Are we really ready for massive parallelism?

- How can we best utilize exisiting and future technology?



- It seems that we have bigger worries than just technology...

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.

Deep computing combines several techniques to solve extremely complex problems in this sea of digital data:

Advanced mathematics

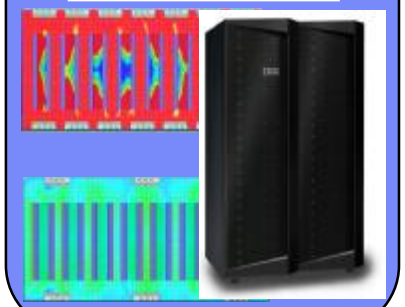
Domain-specific knowledge

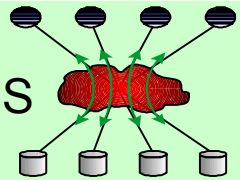
- Specialized software
- Powerful hardware

The full picture...

Software	Tools	Programming models	Applications

x86



Infrastructure	Network 	Storage 	GPFS 

x86 Systems / IBM BladeCenter H

HC10 Workstation blade (Intel Core2 Duo)

HS12 1CPU Intel Xeon (24GB RAM)

HS21 2CPU Intel Xeon (16GB RAM)

HS21XM 2CPU Intel Xeon (32GB RAM)

LS22 2CPU AMD Opteron (32GB RAM)

LS42 4CPU AMD Opteron (64GB RAM)

JS12 1CPU POWER6 (32GB RAM)

JS22 2CPU POWER6 (64GB RAM)

QS21 2CPU Cell BE (2GB XDRAM)

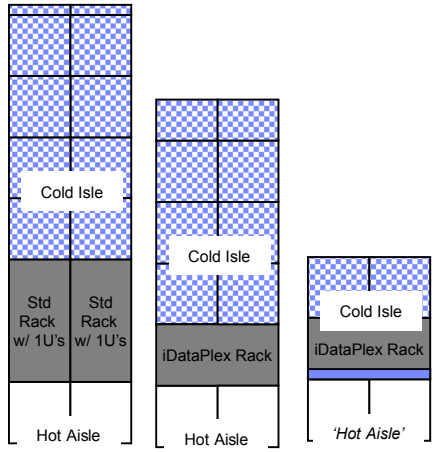
QS22 2CPU PowerXCell (32GB RAM)

PN41 (DPI)

x86 Systems / iDataPlex

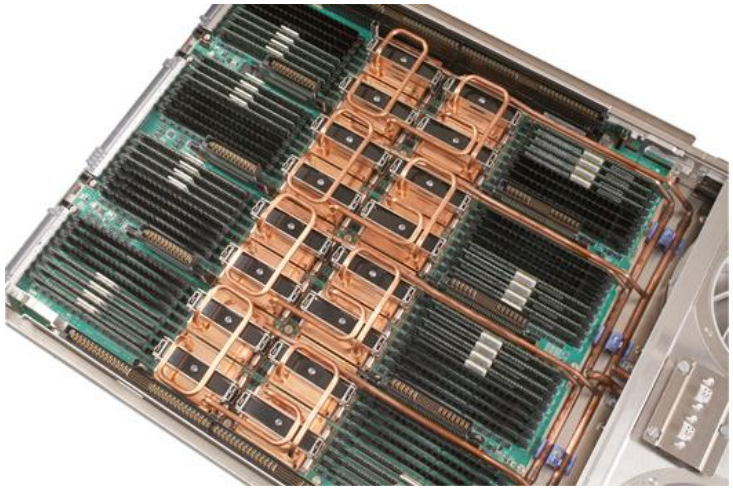
Key iDataPlex concepts:

- higher density packaging
- infrastructure integration
- no HW redundancy
- water cooled rear door
- alternative rack placement
- pre-integrated at plant



Power Systems / POWER6 p575

THE NEW POWER EQUATION



✓ *Water Cooled for energy efficiency*

Node Size

- 2U, 24” X 51” deep, full width drawer
- 32 X 4.7GHz POWER6 cores per node
- 256 GB memory per node

Rack Frame Size

- 42U, 24” rack frame
- 14 nodes per rack frame
- 448 cores per rack frame
- 3.5 TB memory per rack frame

Environment

- AIX & Linux
- Water cooled
- Data Center environment

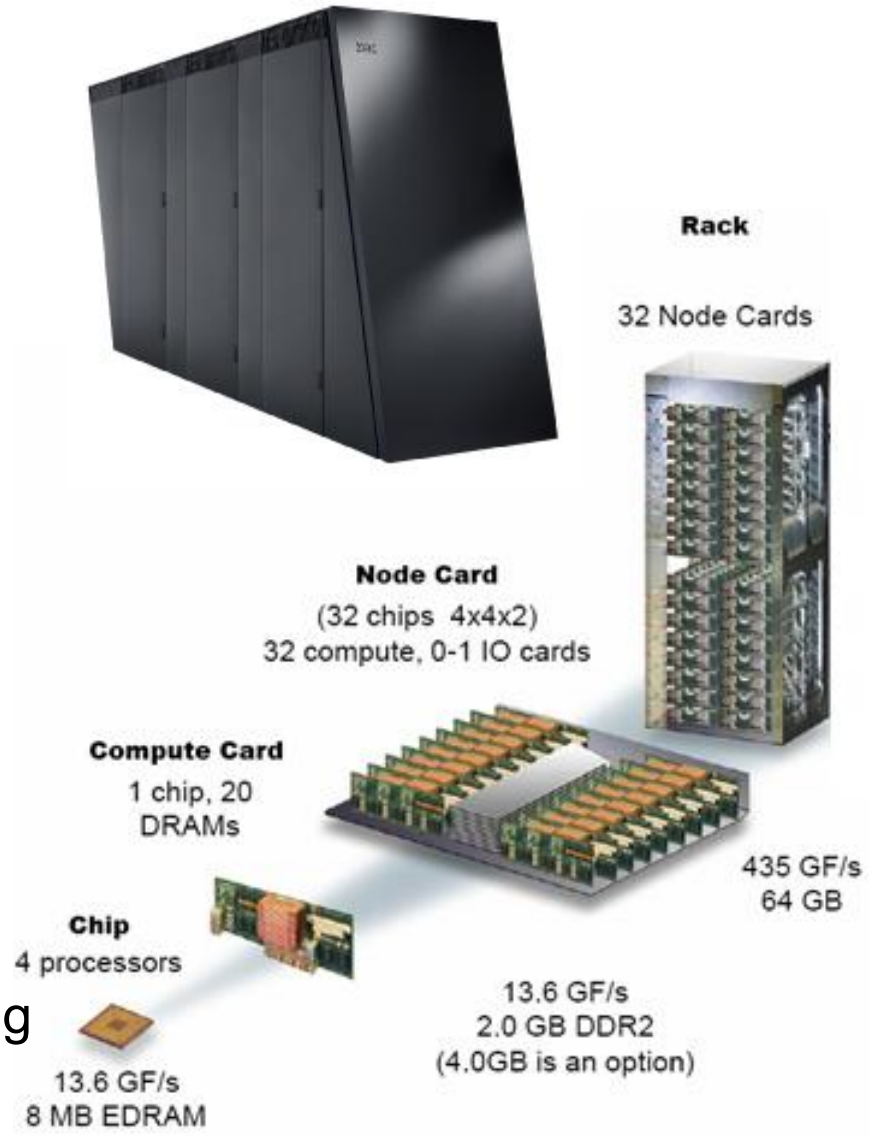
Strengths

- 5X performance improvement over P5+
- 3X improved energy efficiency over P5+

BlueGene/P

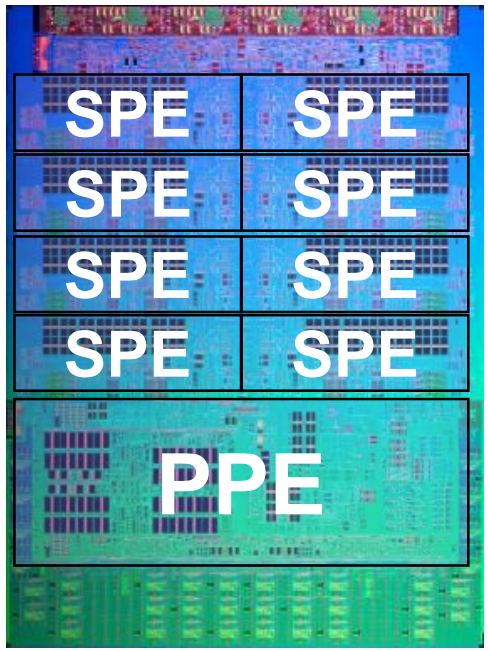
Key BlueGene concepts:

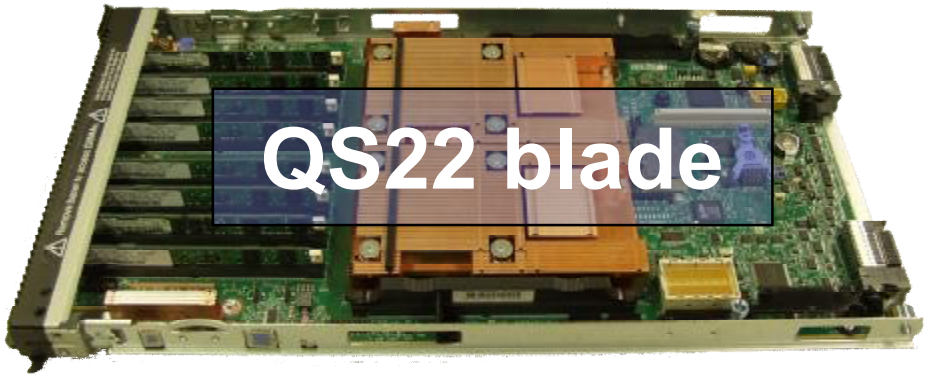
- 1024 compute nodes per BG/P rack
- Each node features:
  - 4 PowerPC 450 CPUs @ 850 MHz
  - 2 or 4 GB RAM per node
  - integrated specialized HPC interconnect
- 8-64 IO nodes per rack
  - for 10GbE file system access
- Performance, Efficiency and Reliability
- Ballanced system for application scaling



PowerXCell / The Cell Processor

- Hybrid microprocessor with 9 cores
  - PPE (control processor) + 8 SPEs (performance optimized)
- Based on a 64-bit Architecture
  - Path for OS, legacy apps, and SW development
- Streaming architecture
  - 155+ concurrent transactions to memory per processor
- Efficient architecture
  - 230.4 GFLOPS (SP) / 108.8 GFLOPS (DP) for 92 Watts
- Real-time architecture
  - Resource allocation, and
  - replacement management
- Security-enabled architecture
  - SPEs individually programmable as secure processors





Roadrunner Breaks Petaflop Record (June '08)

In total, the supercomputer at Los Alamos connects 6,948 dual-core AMD Opteron™ chips and 12,960 PowerXCell 8i processors.

Tri-blade configuration (2xQS22 & 1 LS21)

3,456 tri-blade units (each 400 GFlops).

10,000 connections (both InfiniBand® and Gigabit Ethernet) require 57 miles of fiber optic cable.

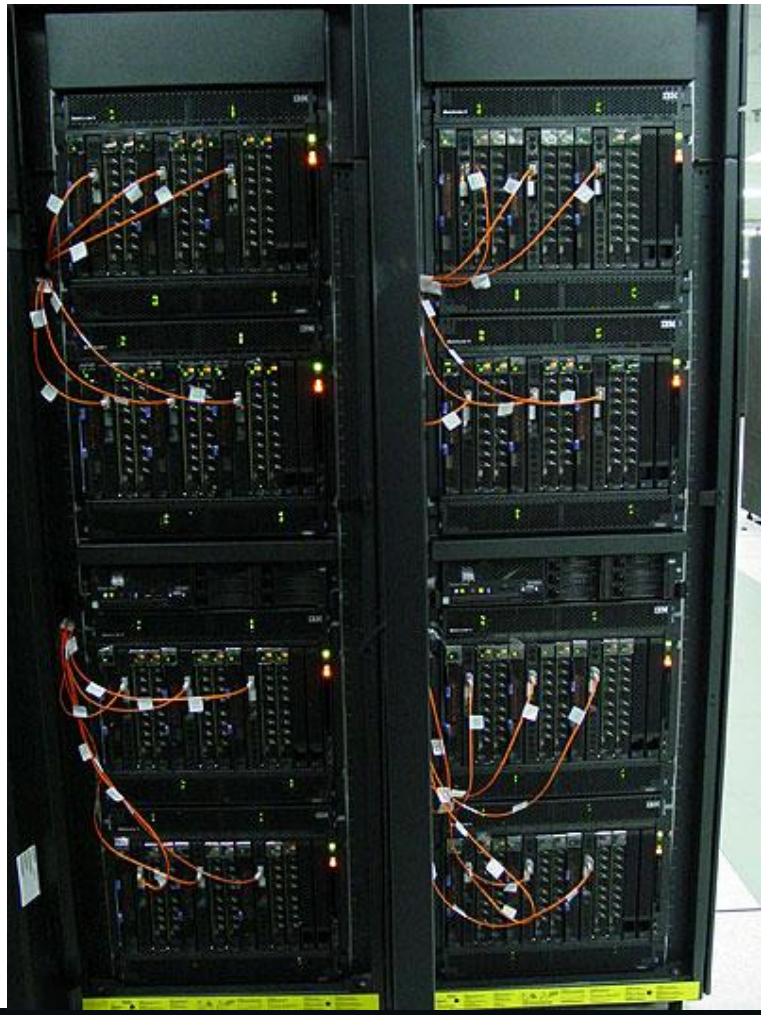
The system has 80 terabytes of memory

Weighs 500,000 pounds

288 racks

6,000 square feet.

...still No.1 on top500!



IBM Storage Systems used in HPC

DS5000 series (ie. DS5300)

- high-performance general purpose storage system for SATA and FC drives
- flexible size and configuration

DCS9900

- designed for HPC & streaming applications
- ultra-high drive capacity with SATA (RAID6)
- fixed configurations (300, 600 or 1200 drives)

IBM General Parallel File System

High performance file system

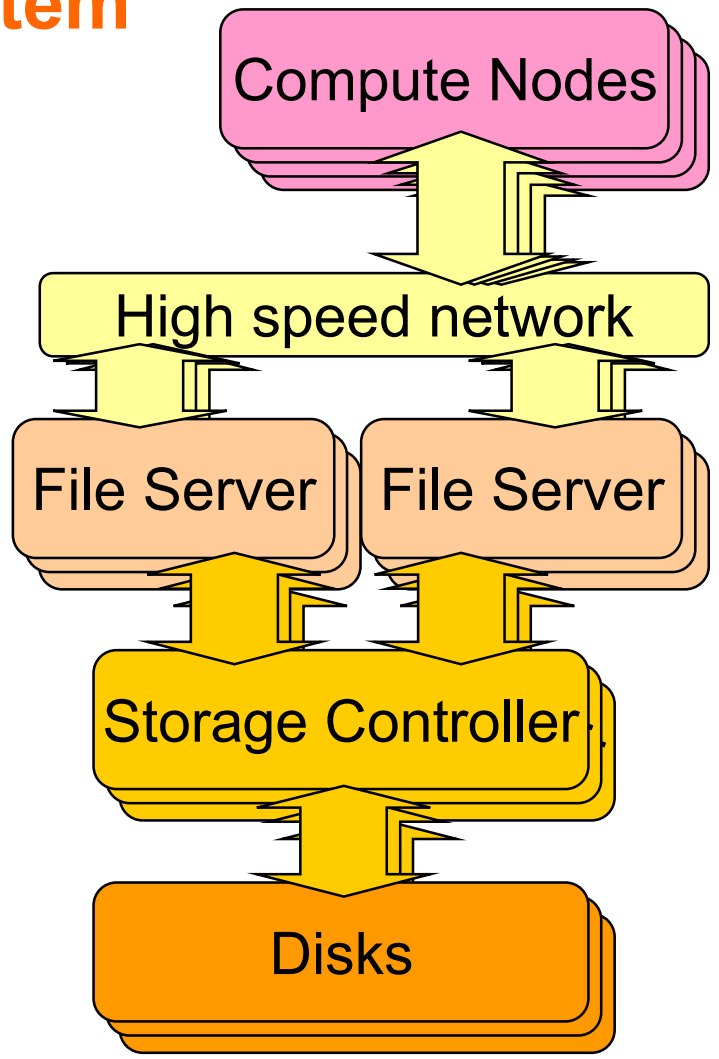
- thousands of nodes
- Petabytes of capacity
- 10s-100s GB/s of IO bandwidth

Building block approach

- balanced for capacity & bandwidth
- maximum performance of all components

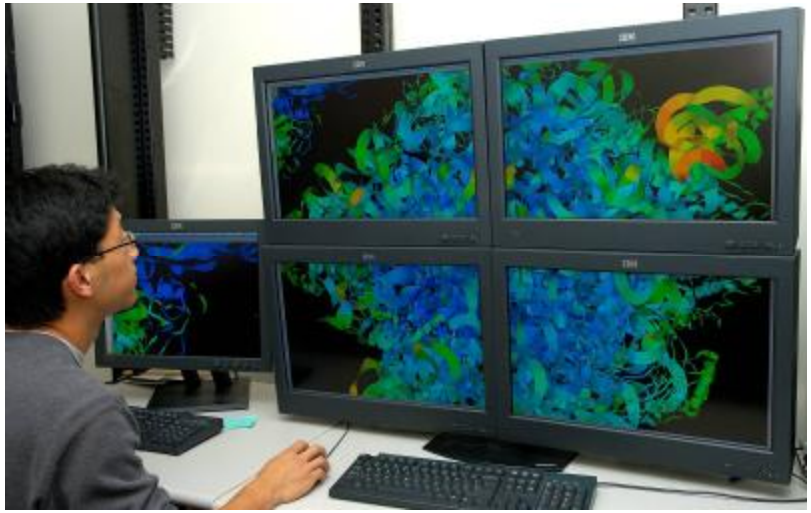
Management at file system level

- storage pools, policies, data replication...



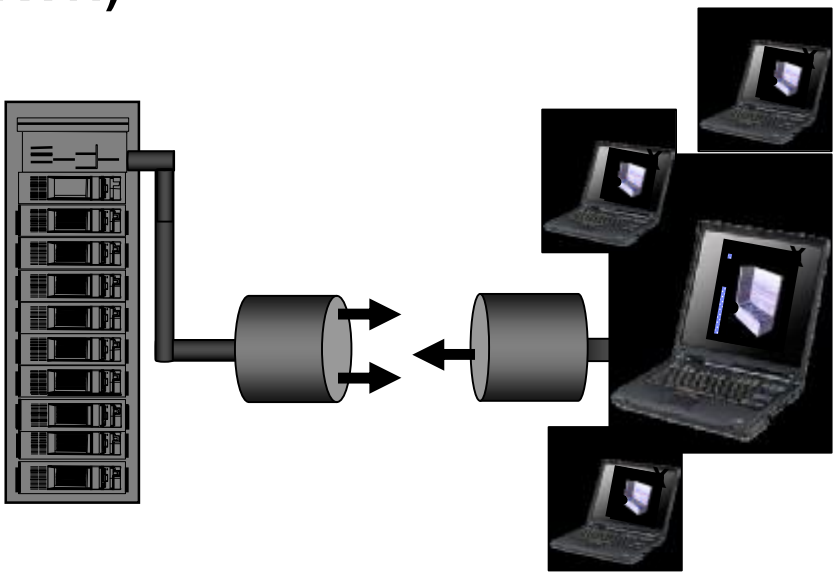
Deep Computing Visualization (DCV)

Scalable Parallel Visual Networking (SPVN)



Improve performance & screen dimensions

Remote Visual Networking (RVN)



Work remotely & collaborate

What is Nautilus?

- A collaboration between IBM and University of Warsaw
  - Interdisciplinary Centre for Mathematical
  - and Computational Modelling (ICM)
- A development system for hybrid computing
  - QS22 blades (PowerXCell8i processors)
  - DDR Infiniband – fully non-blocking
  - LS21 blades (AMD Opteron) will be added for a Roadrunner-like setup
- A two rack, 36 kW supercomputer ranking:
  - No. 221 on the current top500 list with 18.57 TFlops
  - No. 1 on the current green500 list with 536 MFlops/Watt
- **Nautilus is the most energy efficient supercomputer in the World!**

November 2008 edition of the Green500 list

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)	TOP500 Rank*
1	536.24	Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw	BladeCenter QS22 Cluster, PowerXCell 8i 4.0 Ghz, Infiniband	34.63	221
2	530.33	Repsol YPF	BladeCenter QS22 Cluster, PowerXCell 8i 3.2 Ghz, Infiniband	26.38	430
2	530.33	Repsol YPF	BladeCenter QS22 Cluster, PowerXCell 8i 3.2 Ghz, Infiniband	26.38	431
2	530.33	Repsol YPF	BladeCenter QS22 Cluster, PowerXCell 8i 3.2 Ghz, Infiniband	26.38	432
5	458.33	DOE/NNSA/LANL	BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz , Infiniband	138	41
5	458.33	IBM Poughkeepsie Benchmarking Center	BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz , Infiniband	138	42
7	444.94	DOE/NNSA/LANL	BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz , Voltaire Infiniband	2483.47	1
8	371.67	ASTRON/University Groningen	Blue Gene/P Solution	94.5	76
9	371.67	IBM - Rochester	Blue Gene/P Solution	126	56
9	371.67	RZG/Max-Planck-Gesellschaft MPI/IPP	Blue Gene/P Solution	126	57

The Nautilus Project

interdisciplinary centre for mathematical and computational modelling

536 MFlops/Watt running Linpack

STOP STUDYING THE ROAD AHEAD. START MAKING PROGRESS.

Mission Statement of the JCCC (http://cell.icm.edu.pl)

Maintaining up-to-date knowledge and skills related to the CBEA and hybrid computing

Supporting ICM users and researches in porting and tuning applications on the CBEA

Organizing CBEA and hybrid computing programming classes at ICM

Identifying new application areas for Cell and assisting in application enablement

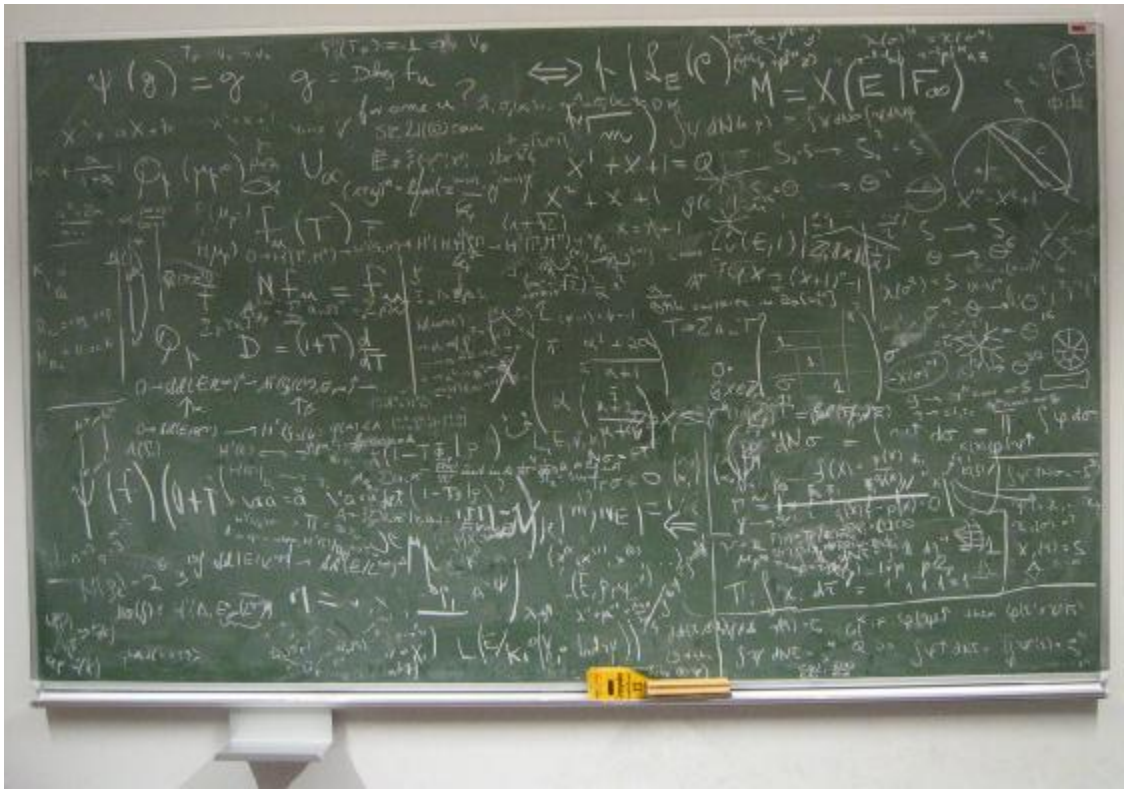
Maintaining relationships with developers at other institutions involved with Cell

Identifying potential for involvement in test & development activities related to the CBEA

Evaluating opportunities of commercial engagement in the business and industrial sector

Coordinating work of other local teams involved in CBEA and hybrid computing

Questions?



IBM

Thank you for your attention

;-)

STOP
STUDYING THE ROAD AHEAD.
START
MAKING PROGRESS.