IBM Blue Gene Updates

George Chiu

June 2005: 25th TOP10 of the TOP500 list

Year	Computer	Measured Tflop/s	Theoretical Peak Tflop/s	Number of Processors	% peak	Flops/Hz	Processor MHz
2005	IBM BlueGene/L DD2, LLNL	136.80	183.50	65536	75	4	700
2005	IBM BlueGene/L DD2, IBM Yorktown	91.29	114.688	40960	80	4	700
2004	Columbia, SGI Altrix 3700, Itanium 2 for NASA	51.87	60.96	10160	84	4	1500
2002	Earth Simulator, NEC	35.86	40.96	5120	87.5	16	500
2005	Mare Nostrum, IBM JS20 PowerPC970	27.91	42.144	4800	66	4	2200
2005	IBM BlueGene/L DD2, ASTRON	27.45	34.4064	12288	80	4	700
2004	Thunder, Itanium 2, LLNL	19.94	22.938	4096	87	4	1400
2005	IBM BlueGene/L DD2, EPFL	18.2	22.9376	8192	79	4	700
2005	IBM BlueGene/L DD2, AIST	18.2	22.9376	8192	79	4	700
2005	Cray Redstorm, Sandia	15.250	20.000	5000	76	2	2000
2005	IBM ASC Purple	60.49	72.048	9480	84	4	1900

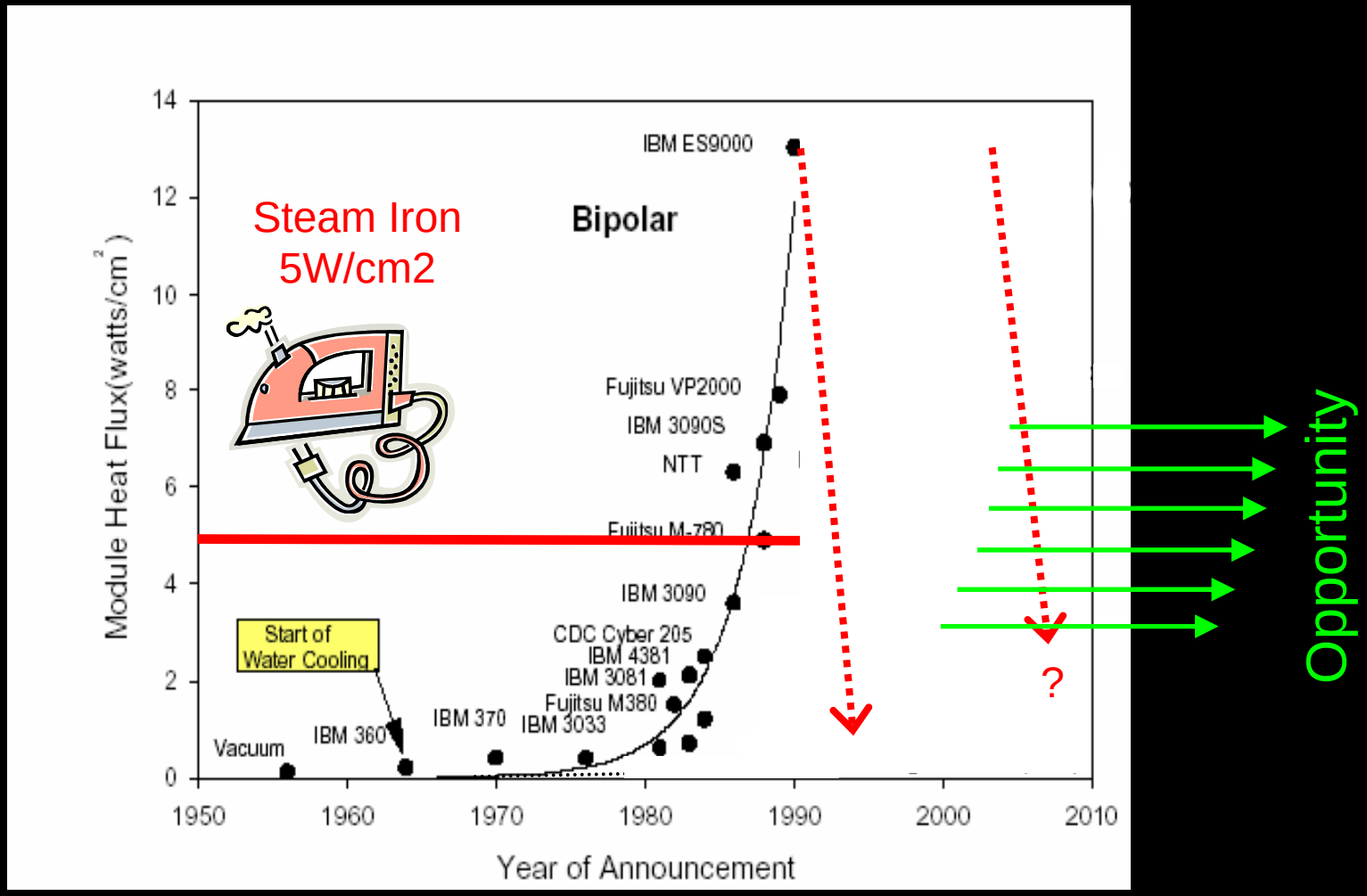
IBM is the clear leader in the TOP500 list with 51.8% of systems and 57.9% of installed performance.

Comparing Systems

	ASCI White	ASCI Q	Earth Simulator	Blue Gene/L
Machine Peak (TF/s)	12.3	20.48	40.96	367
Total Mem. (TBytes)	8	22	10	32
Footprint (sq ft)	10,000	14,000	34,000	2,500
Power (MW)	1	1.9	6-8.5	1.5
Cost ($M)	100	165	400	100
# Nodes	512	2048	640	65,536
MHz	375	1250	500	700

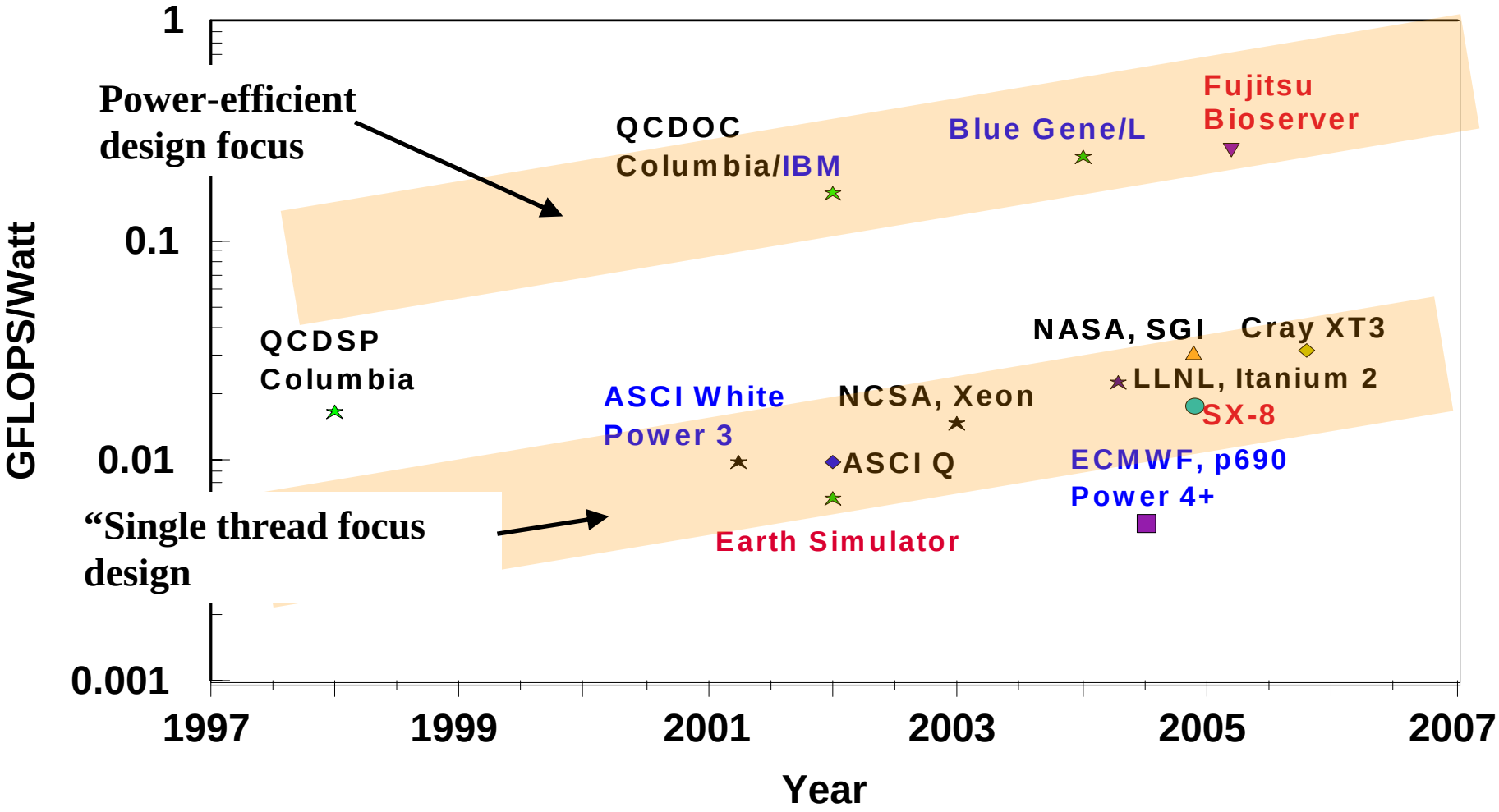
System power is again THE problem.

Single thread focus has resulted in power inefficient design



© 2004 IBM Corporation

Supercomputer Power Efficiencies



Future Scaling without innovation

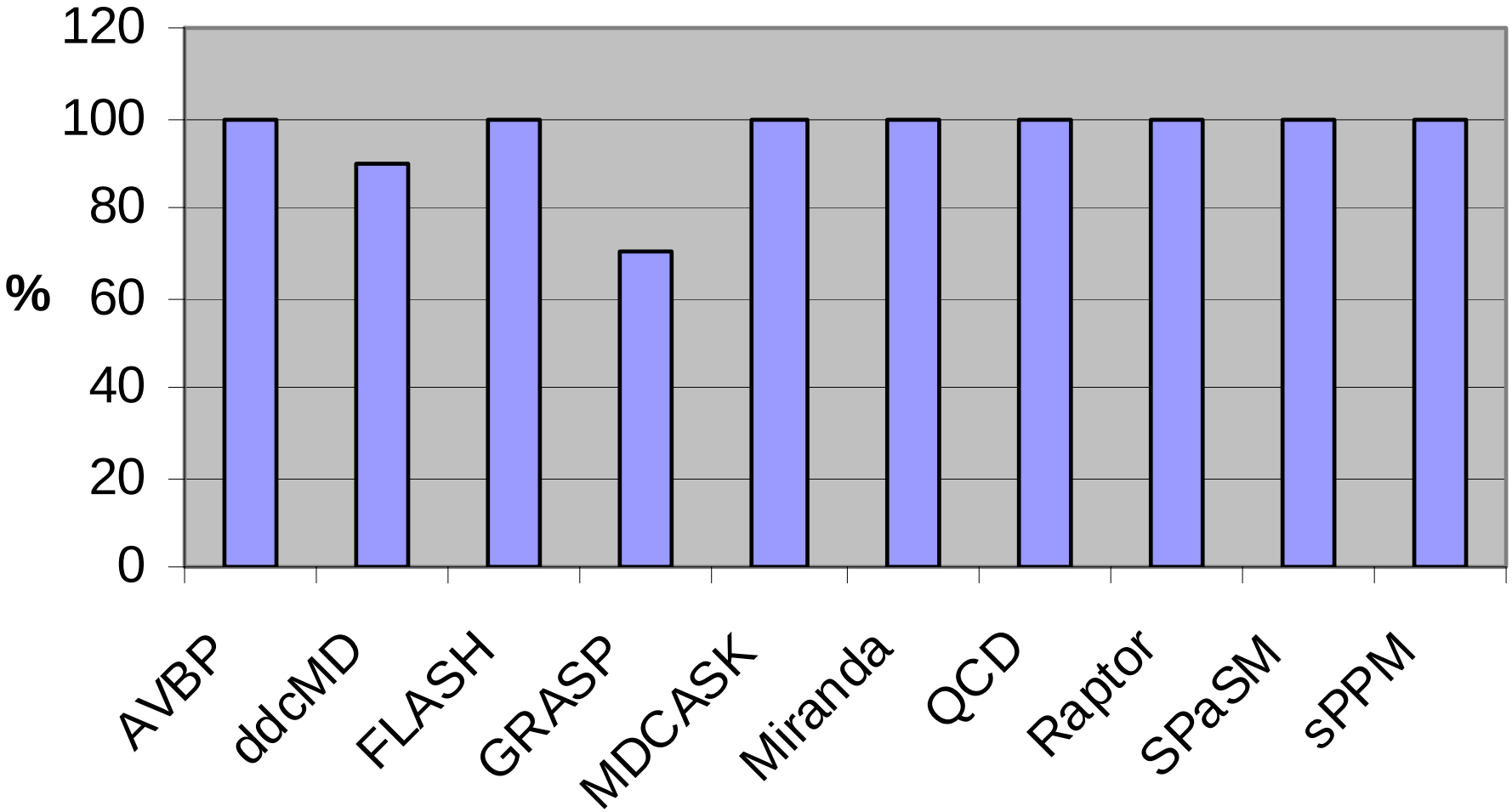
If we scale current peak performance numbers for various architectures and allowing system peak doubling every 18 months. Trouble ahead

	Projected Year	Blue Gene/L	Earth Simulator	MareNostrum
250 TF	2005	1.0 MWatt	100 MWatt	5 MWatt
1 PF	2008	2.5 MWatt	200 MWatt	15 MWatt
10 PF	2013	25 MWatt	2000 MWatt	150 MWatt
100 PF	2020	250 MWatt	20000 MWatt	1500 MWatt

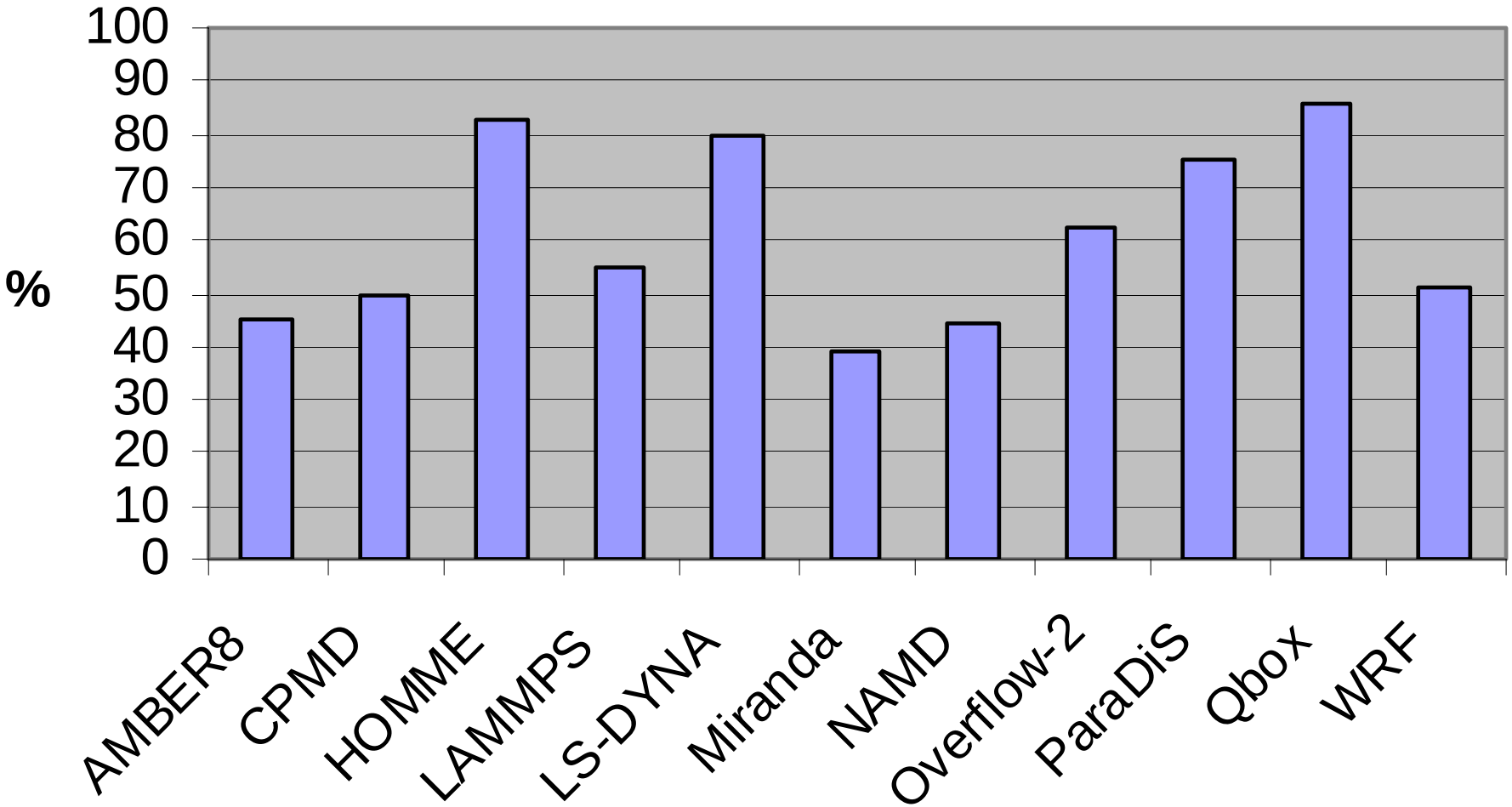
Applications

- N-body simulation
  - Classical molecular dynamics: AMBER8, Blue Matter, ddcMD, DL_POLY, GRASP, GROMACS, LAMMPS, LJ, MDCASK, NAMD, PRIME (Schrodinger), SPaSM
  - Quantum chemistry: CHARMm, CPMD, FEQMD, GAMESS-UK, GAMESS-US, Gaussian, NWChem, Qbox,
  - Plasma Physics: TBLE, GTC
  - Stellar dynamics of galaxies: Enzo
- Complex multiphysics code
  - Climatology: CCSM, HOMME
  - Computational Fluid Dynamics: FLUENT, Miranda, Overflow, POP (Ocean), Raptor, SAGE, sPPM, STAR-CD,
  - Astronomy: Accretion, planetary formation and evolution, stellar evolution, FLASH (supernova), Radiotelescope (ASTRON)
  - Electromagnetics: FDTD code
  - Finite element analysis, Car Crash: LS-Dyna, NASTRAN, PAM-Crash (ESI), HPCMW (RIST)
  - Radiative transport: 2-D SPHOT, 3-D UMT2000
  - Neutron transport :Sweep3D
  - Weather: MM5, IFS (ECMWF), WRF
- Life Sciences and Biotechnology: mpiBLAST, Smith-Waterman
- CAD/CAE: AVBP
- Crystallography with X-Ray Diffraction: Shake & Bake
- Drug Screening: OpenEye Scientific Software, Tripos, MOE (CCG: Chemical Computing Group)
- Finance: NIWS (Nissei)
- Nanophysics and Materials: LSMS
- Quantum Chromodynamics: IBM, Boston University, KEK, Jülich

Weak Scaling Parallel Efficiency



Strong Scaling Parallel Efficiency



Blue Gene – A System Approach that will continue

- Integration
  - Use system-on-a-chip (processors, EDRAM, communication networks) to achieve high integration, low cost.
  - Communication networks and processor using the same technology – performance improvement hand in hand.
  - Communication networks and processor are on the same development cycle.
- Low power and High Packaging Density
  - Embedded processors
  - More processors/rack, so that Blue Gene rack performance exceeds that based on high power processors.
  - High density, smaller footprint, more processors/ft^2.
  - High efficiency, more flops/W.
- System design
  - Powerful networks scaling out for High Performance Computing
  - Balanced memory and network (short latency, high bandwidth, high efficiency, low cost)
  - 1K I/O nodes and 64K compute nodes: Divide and conquer
- Leader in Cost Performance

Japan Supercomputer Effort

- Japan's National pride is at stake. They aim to regain the number 1 spot.

- Since November 2004, the Japanese Ministry for Education, Culture, Sports, Science and Technology (MEXT) planned a new supercomputer effort to beat US.

- NEC, Hitachi, the University of Tokyo, Kyushu University, and the Institute of Physical and Chemical Research will receive about $1B to design and build the follow-on to Earth Simulator. The first phase will start in 3Q05, and the second phase in 3Q06.
    - Our guess is that 1PF is the 2008 target.
    - Official goal is 3-4 PF, in 2011, real target is 10 PF.

- Concentrated in one design point.

Closing Points

- Blue Gene represents an innovative way to scale to multi-teraflops capability
  - Low power => efficient packaging => floor space consumption
  - Massive scalability
  - Balanced design in terms of memory subsystem and communication subsystem
  - Better than COTS clusters by virtue of density, scalability and innovative interconnect design
  - Bottom up hardware architecture
  - Top Down software design
- Blue Gene is applicable to a wide range of Deep Computing workloads
- Blue Gene R&D continues to protect the investment of users

Mission Impossible?

- Build a sports car like a Porsche
- Engine power doubles every 2-3 year
- Running full throttle for at least a week without maintenance
- Energy efficient like a Toyota Prius
- Adapts to various driving conditions: race track, wet surface, snow days, LA freeway…
- Price kept low like a Kia
- Development $ shrinking
- Stiff foreign competition
- Must be made in America