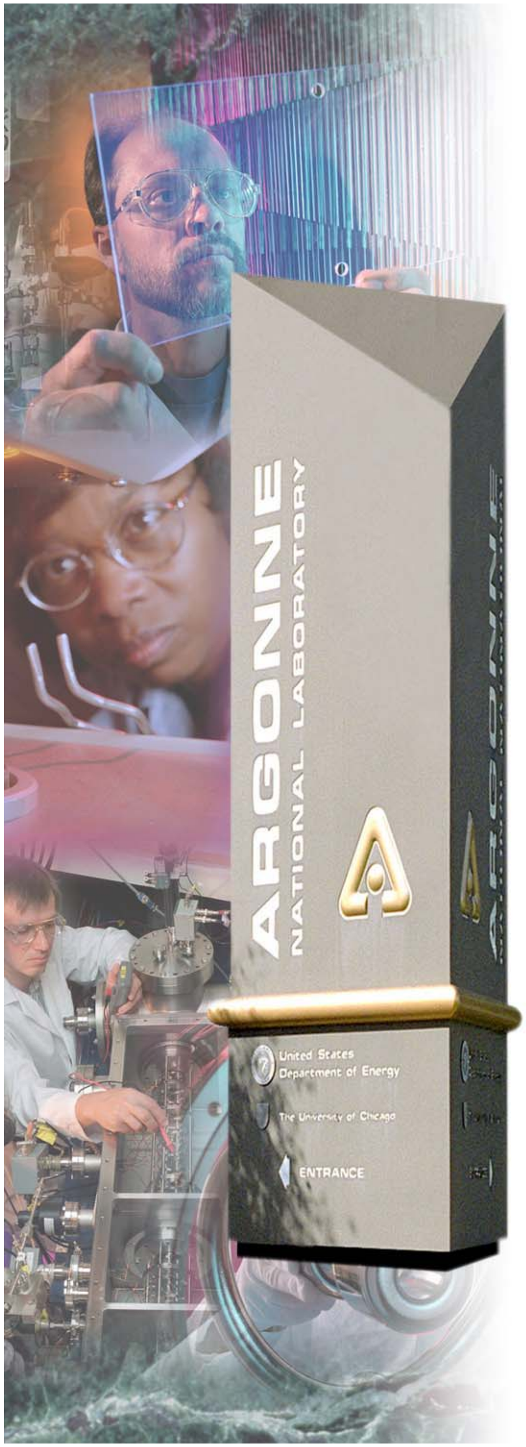



# *A Science-Based Case for Large Scale Computation*

***William D. Gropp***
***Mathematics and Computer Science***
***www.mcs.anl.gov/~gropp***

***Argonne National Laboratory***

Office of Science
U.S. Department of Energy

*A U.S. Department of Energy*
*Office of Science Laboratory*
*Operated by The University of Chicago*

## *Building a Science-Based Case*

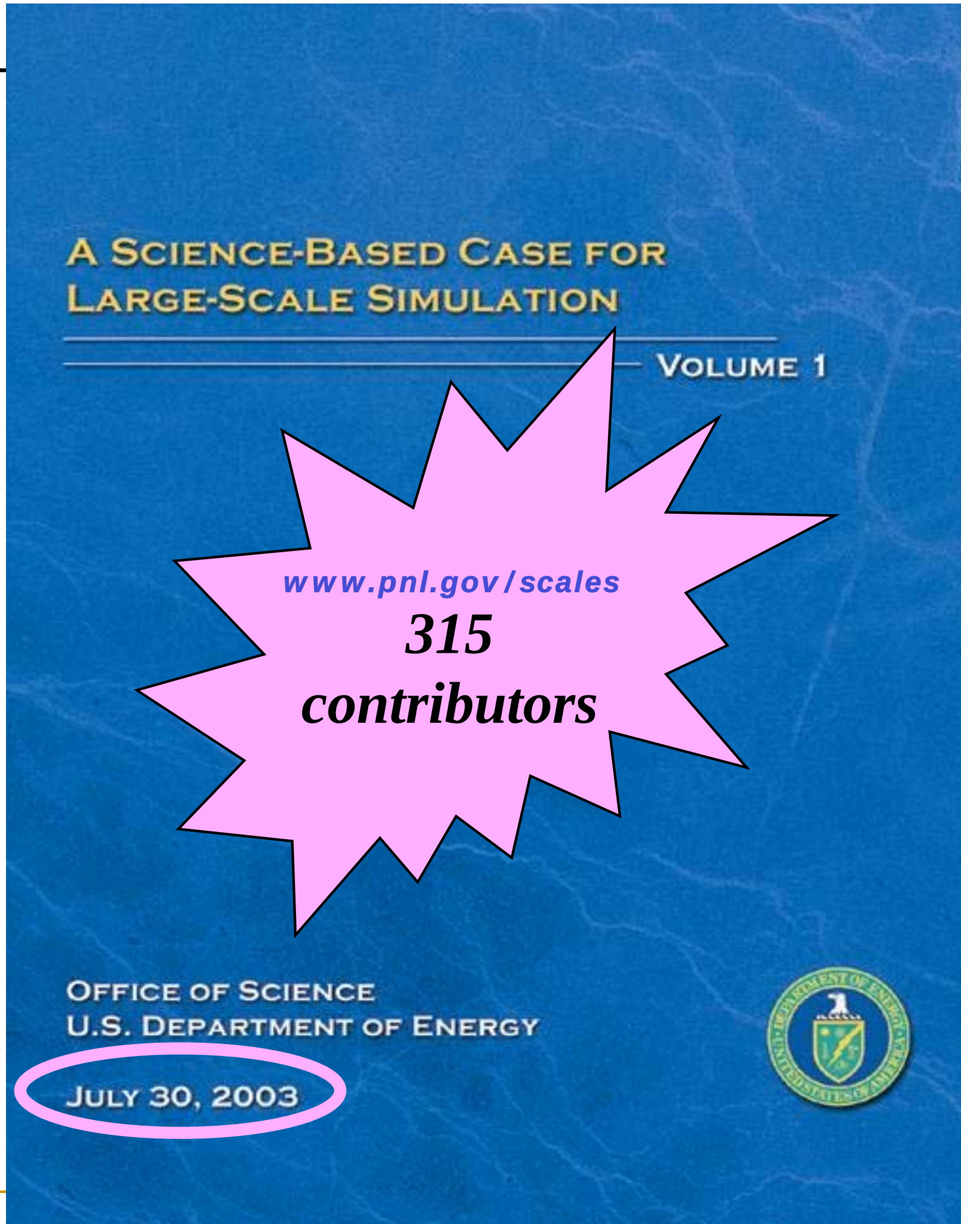
- **Process and Context**
- **Large Scale Computation**
  - Current and future
  - Its not just the hardware
- **Why Build a Case**
- **How to Build a Case**
- **Opportunities and Examples**
- **Final Thoughts and Cautions**

# Process

- **SCaLeS workshop, June 2003 in Washington D.C.**
  - Produced two reports
    - *Volume 1 in July 2003 (76 pages)*
    - *Volume 2 in September 2004 (289 pages)*
- **Many related reports and workshops**
- **Acknowledgements**
  - Many Slides, particularly on SCaLeS, thanks to David Keyes
  - SCaLeS Reports
    - *Chief editor David Keyes*
    - *Co-Editors Thom Dunning (applications), Philip Colella (mathematics), myself (computer science)*

# *Context: recent reports promote simulation*

- **Cyberinfrastructure (NSF, 2003)**
  - new research environments through cyberinfrastructure
- **Facilities for the Future of Science (DOE, 2003)**
  - "ultrascale simulation facility" ranked #2 in priority (behind ITER)
- **High End Computing Revitalization Task Force (Interagency, 2004)**
  - strategic planning on platforms
- **SCaLeS report, Vol 1 (DOE, 2003) & Vol 2 (DOE, 2004)**
  - implications of large-scale simulation for basic scientific research
- **Capability Computing Needs (DOE, 2004)**
  - Profiles of leading edge DOE codes in 11 application domains
- **Future of Supercomputing (NAS, 2005)**
  - broad discussion of the future of supercomputing
- **PITAC (Interagency, 2005)**
  - challenges in software and in interdisciplinary training
- **Simulation-based Engineering Science (NSF, 2005)**
  - opportunities in dynamic, data-driven simulation and engineering design




- Chapter 1. *Introduction*
- Chapter 2. *Scientific Discovery through Advanced Computing: a Successful Pilot Program*
- Chapter 3. *Anatomy of a Large-scale Simulation*
- Chapter 4. *Opportunities at the Scientific Horizon*
- Chapter 5. *Enabling Mathematics and Computer Science Tools*
- Chapter 6. *Recommendations and Discussion*

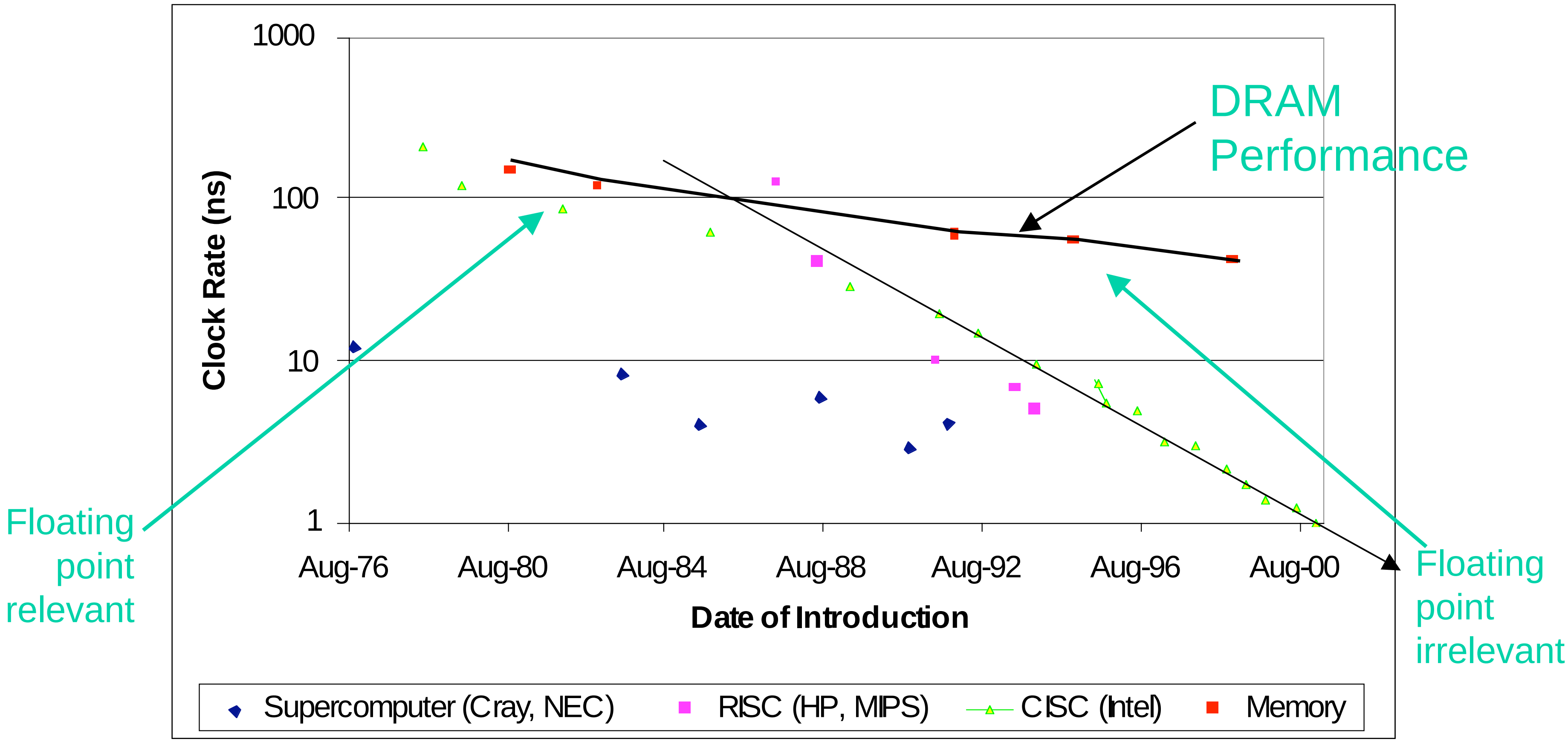
**Volume 2 (2004):**

- **11 chapters on applications**
- **8 chapters on mathematical methods**
- **8 chapters on computer science and infrastructure**

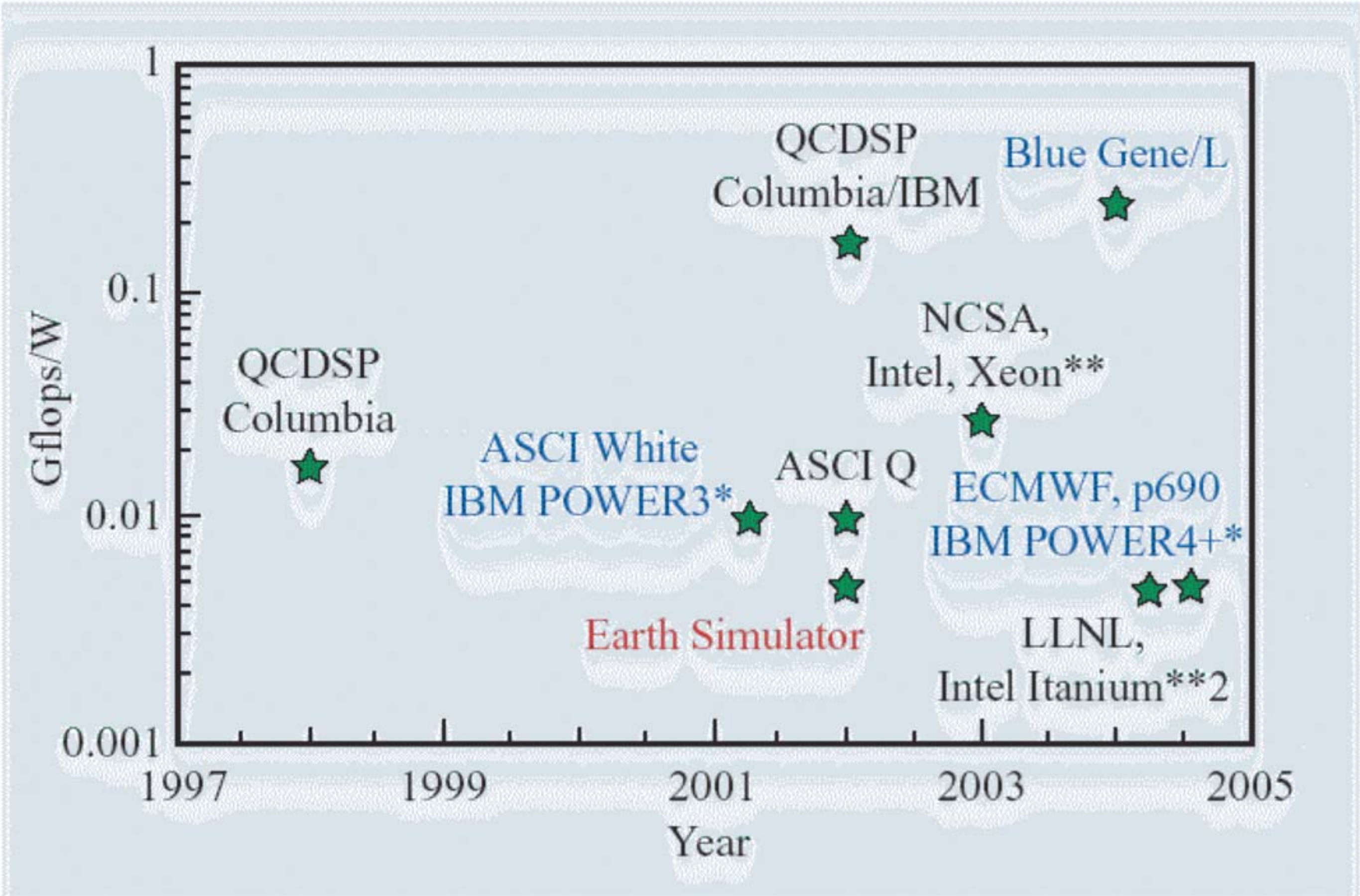
# *Large Scale Computation*

- **Computing is in routine use in science and engineering**
  - Widespread use of small to medium-scale computing
- **Not all computation is “Large Scale”**
  - Capacity — Meeting the raw need for cycles
    - *Can be provided with large numbers of loosely connected systems*
    - *Human Genome, much bioinformatics is widescale task farm*
  - Capability — Meeting the raw need for cycles, memory, I/O, etc in solving a *single instance* of a problem
    - *Requires tighter (faster) coordination between systems*
    - *Memory (problem size), not just cycles*
- **Harder to make is the case for large-scale computing**
  - Lifetime of an individual high-end system is short (3-5 years)
  - Capability computers requires more resources spent on interconnects, memory systems
    - *Price per peak op is higher for capability systems*
- **What are the trends in large scale computing hardware?**

# *CPU and Memory Performance*

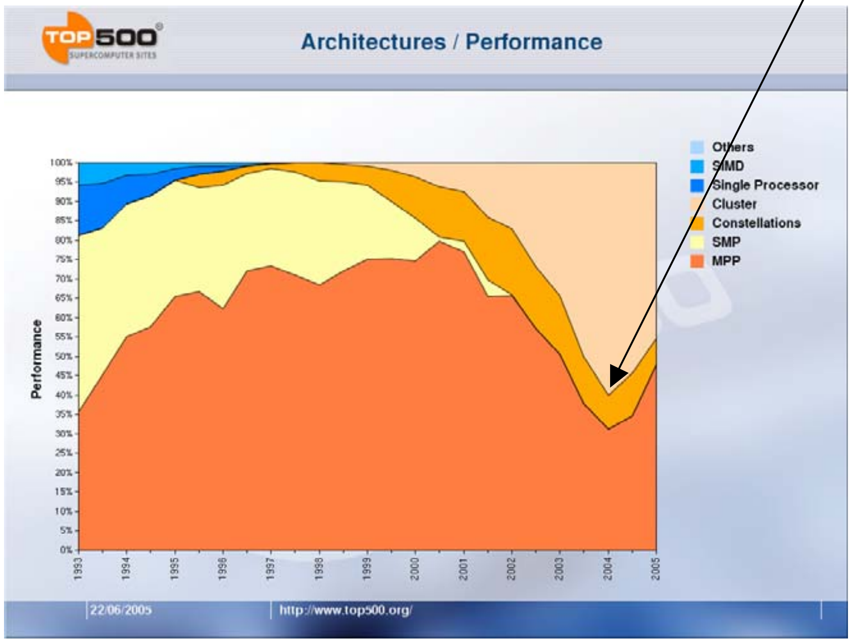


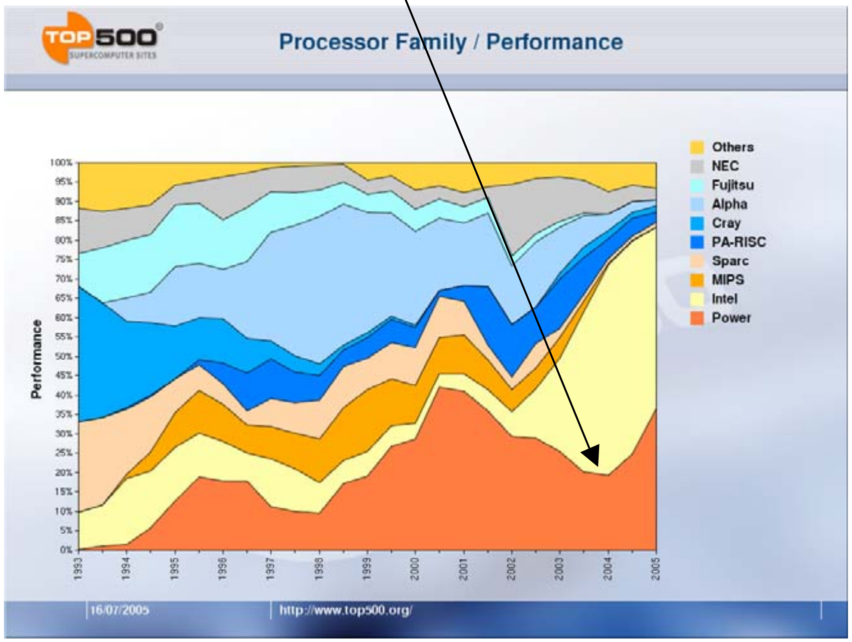
# Power Efficiencies of Supercomputers



# *Clusters Face Competition on the Top500*

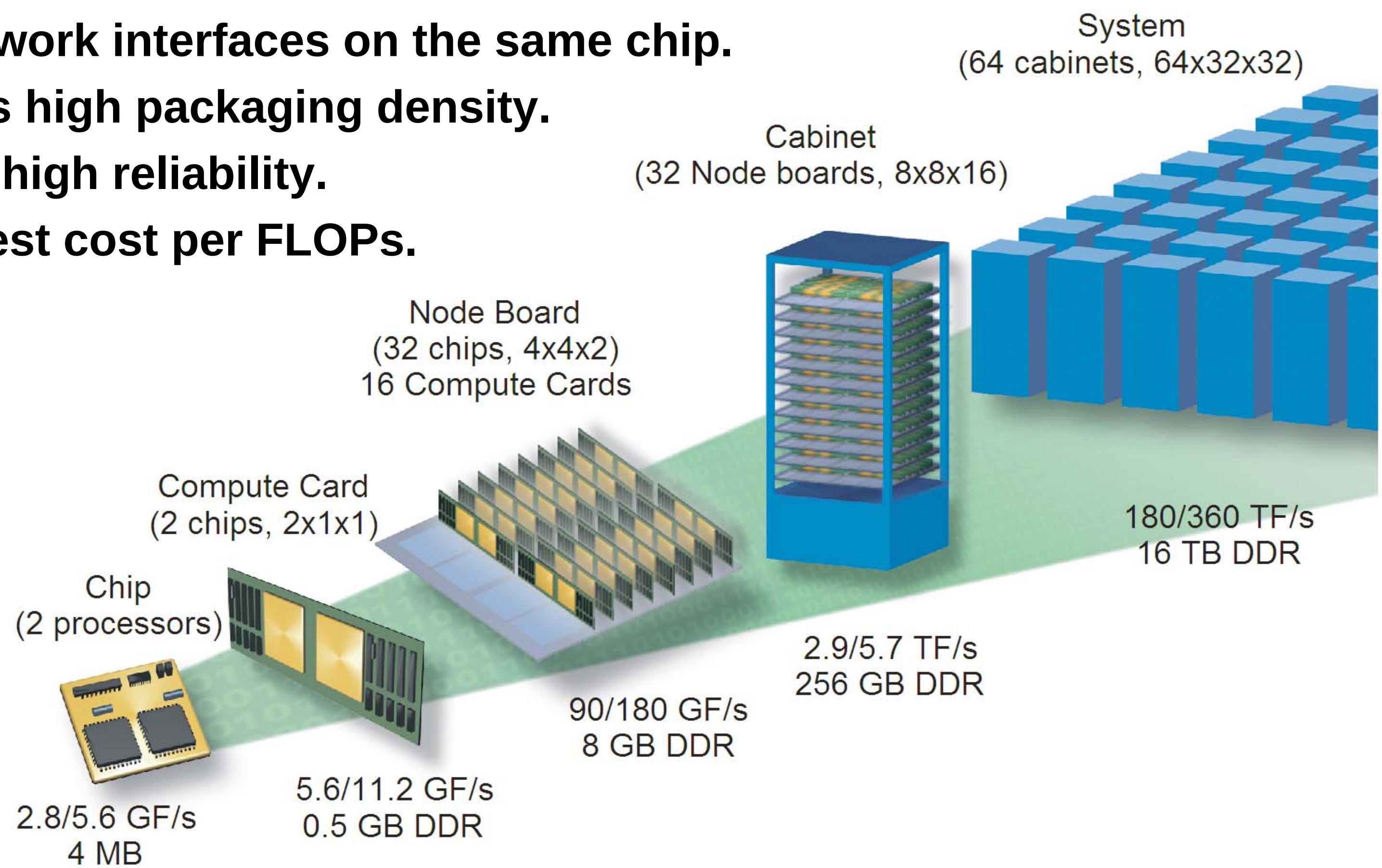
IBM BlueGene Introduced



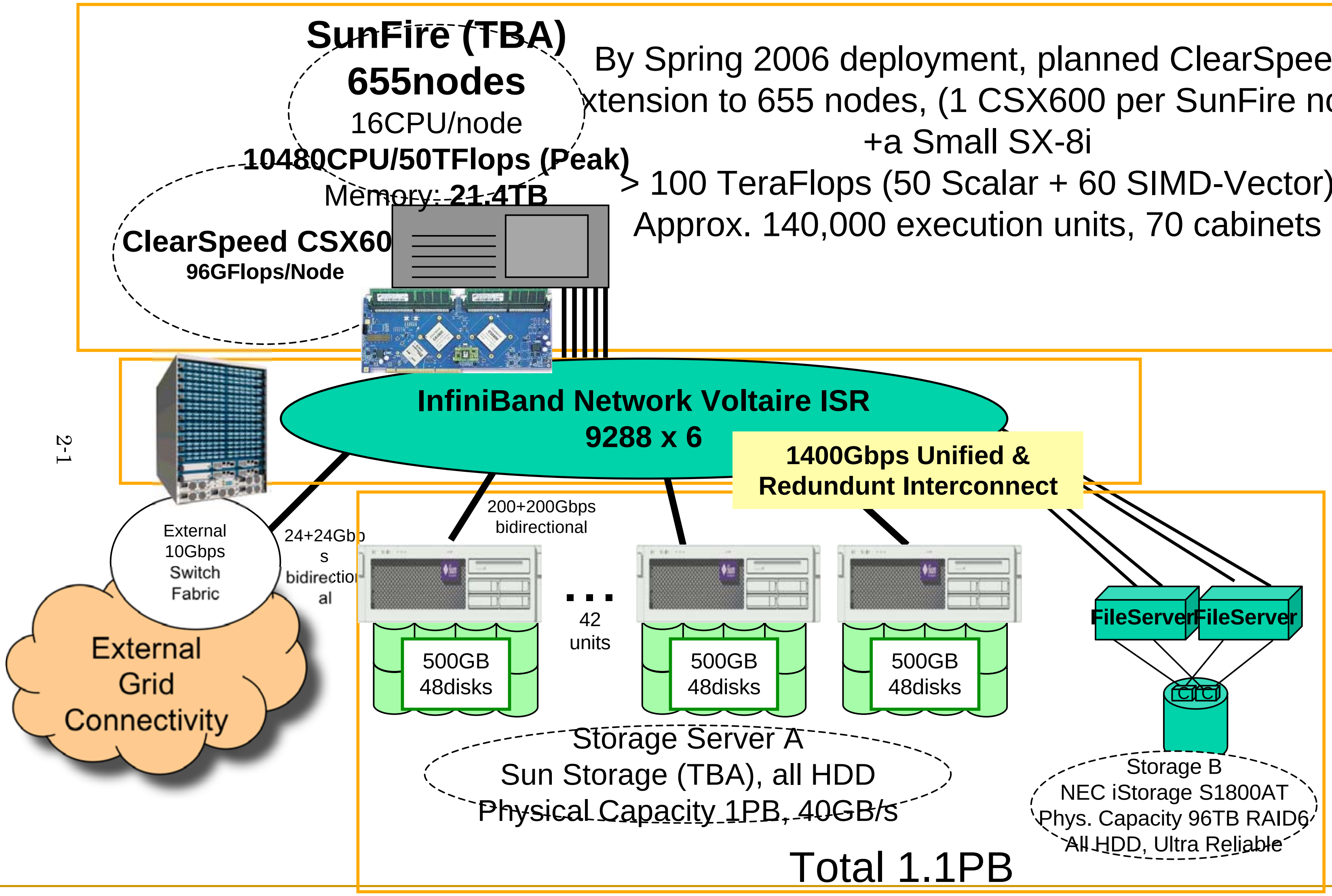



# *The Blue Gene Family of Computers*

- Puts multiple processors on the same chip.
- Puts network interfaces on the same chip.
- Achieves high packaging density.
- Delivers high reliability.
- Has lowest cost per FLOPs.



# NEC/Sun Campus Supercomputing Grid: Core Supercomputer Infrastructure @ Titech GSIC - to become operational late Spring 2006



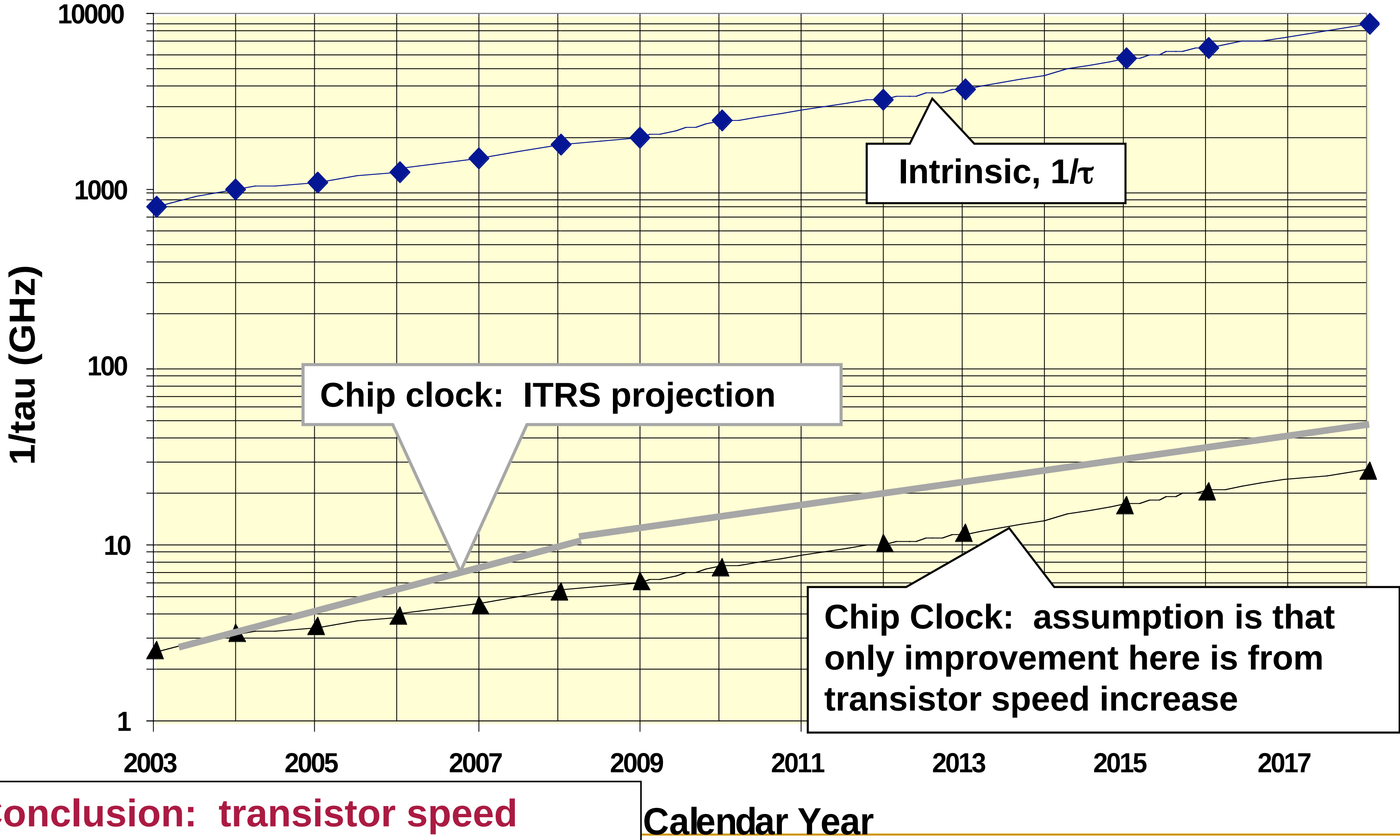

2-1

Thanks to Satoshi Matsuoka, TITech

# *Commodity Systems Will Continue to Scale*

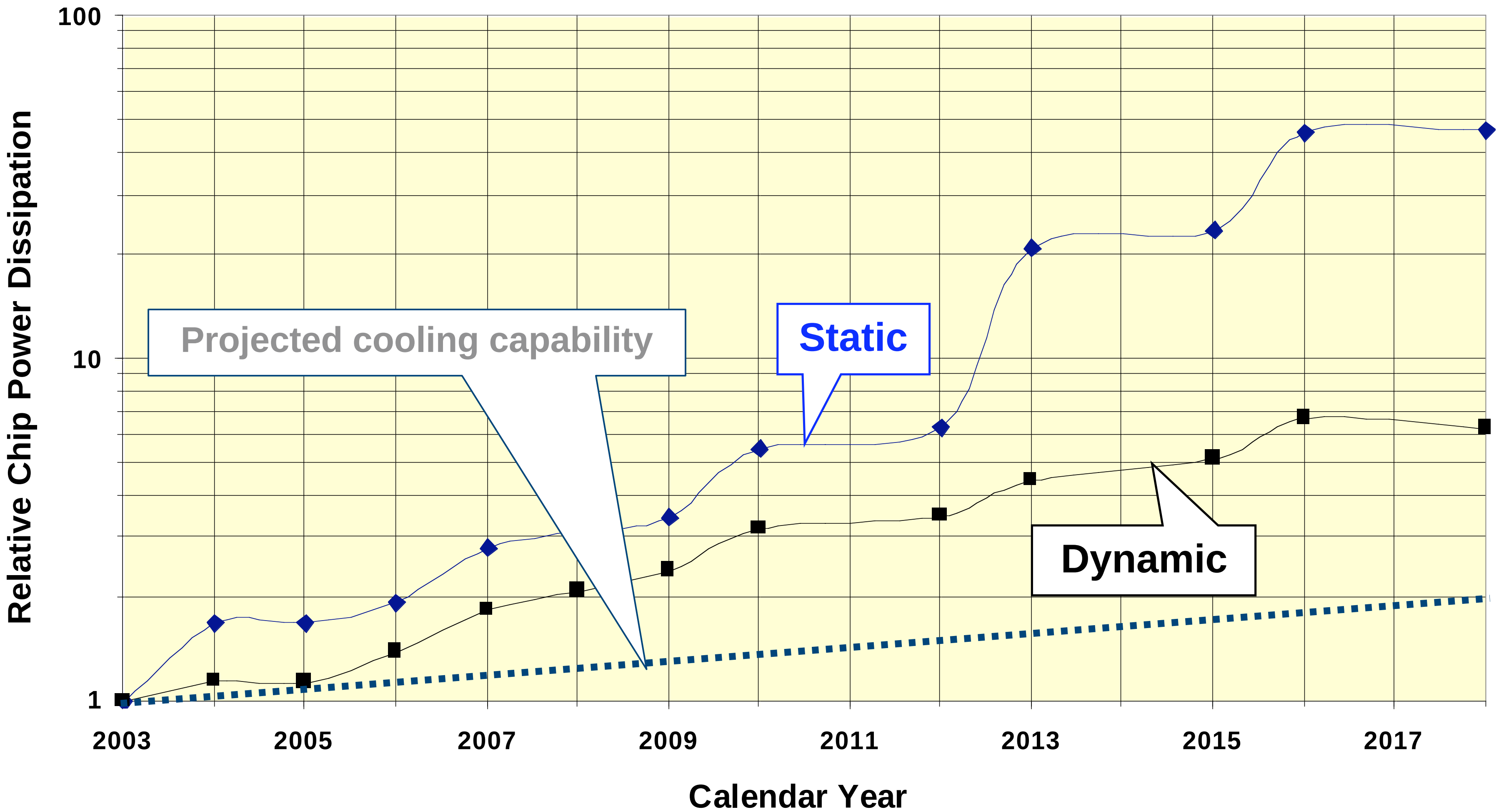
- **ITRS MOSFET Scaling Trends, Challenges, and Key Technology Innovations**
- **Industry-wide effort to map IC technology generations for the next 15 years**
  - Projections are based on modeling, surveys, literature, experts' technical judgment
- **The next three charts taken from a talk by Peter M. Zeitzoff, Sematech, given at the *Workshop on Frontiers of Extreme Computing*, 2005, and based on the 2003 ITRS report**
- **Demonstrates plausibility of continued scaling of CMOS-based devices for at least 10 years**

## *Frequency scaling: Transistor Intrinsic Speed and Chip Clock Frequency for High-Performance Logic. Data from 2003 ITRS.*




**Conclusion: transistor speed improvement is a critical enabler of chip clock frequency improvement**

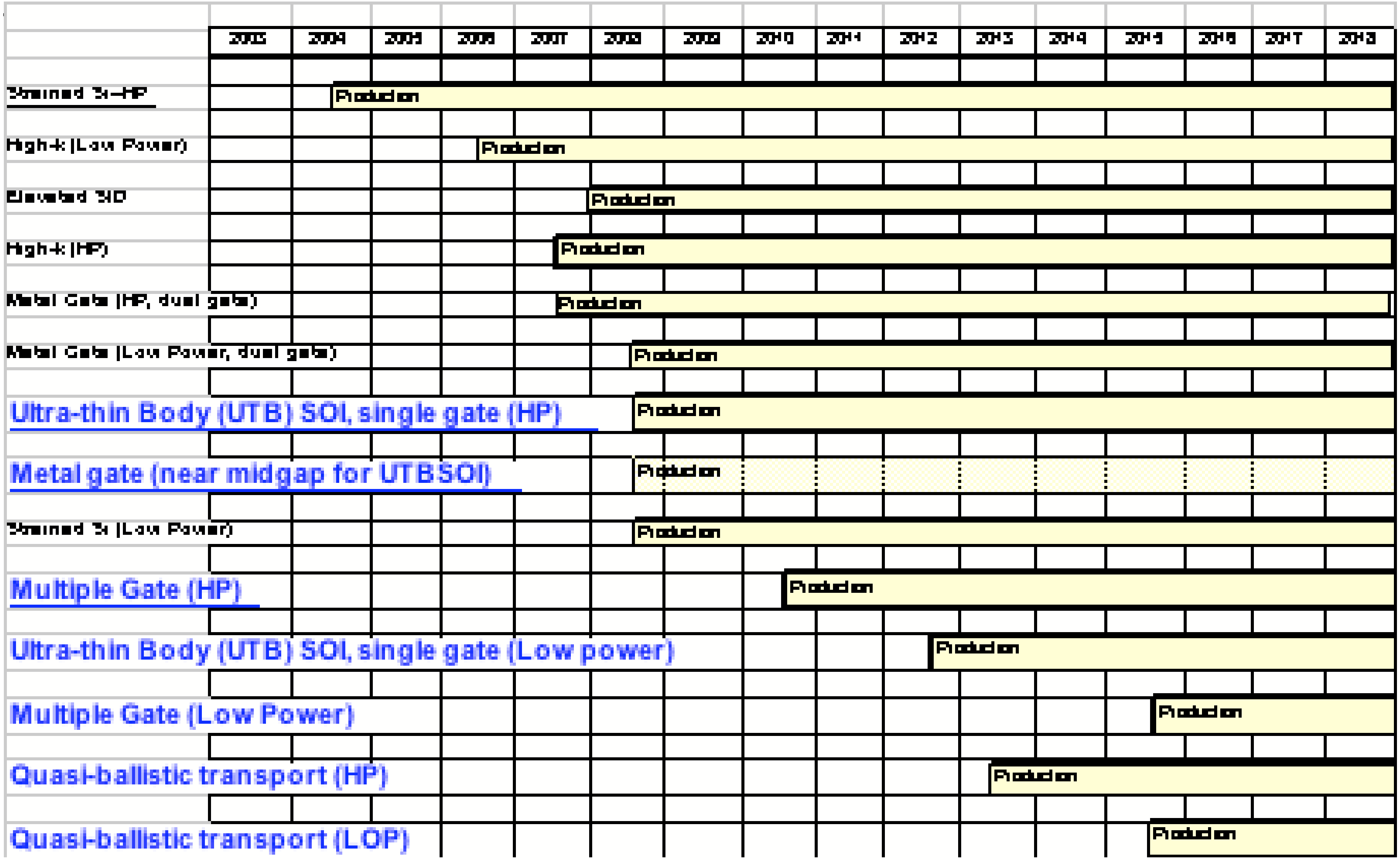
## *Potential Problem with Chip Power Dissipation Scaling: High-Performance Logic, Data from 2003 ITRS*




**Unrealistic assumption, to make a point about $P_{static}$: all transistors are high performance, low $V_t$ type**

## *Timeline of Projected Key Technology Innovations from '03 ITRS, PIDS Section*

**This timeline is from PIDS evaluation for the 2003 ITRS**



# *Components of Large Scale Computing*

- **Large Scale computing is about more than the hardware**
  - Also Mathematics (e.g., modeling and numerical algorithms) and Computer Science (e.g., CS algorithms and programming abstractions)
  - The math and CS are persistent, unlike the hardware
    - *Continue to benefit successive generations of hardware*
- **Advances in hardware have been very predictable:**
  - As we've seen; projections expect the same for at least the next 13 years
- **Advances in math and CS have not been predicted (at least to the detail of the hardware road maps)**
  - Doesn't mean that they're not predictable; the hardware advances have involved new research and innovation
  - We need a "Moore's Law" for algorithms

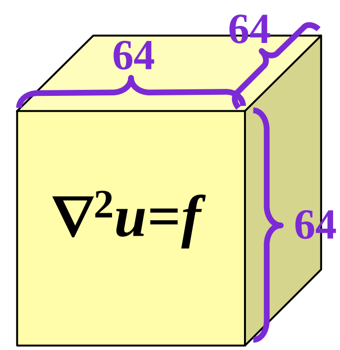
# *Hardware, Mathematics, Computer Science*

- **Applications can make a case for needing 100-1000 fold (and more!) increase in computing performance**
- **Where will this come from?**
  - Improvements in computer hardware (bigger, faster systems)
  - Advances in modeling and mathematics
  - Advances in computing techniques

# *The power of optimal algorithms*

- **Advances in algorithmic efficiency can rival advances in hardware architecture**
- **Consider Poisson's equation on a cube of size $N=n^3$**

| *Year* | *Method* | *Reference* | *Storage* | *Flops* |
|---|---|---|---|---|
| 1947 | GE (banded) | Von Neumann & Goldstine | $n^5$ | $n^7$ |
| 1950 | Optimal SOR | Young | $n^3$ | $n^4 \log n$ |
| 1971 | CG | Reid | $n^3$ | $n^{3.5} \log n$ |
| 1984 | Full MG | Brandt | $n^3$ | $n^3$ |

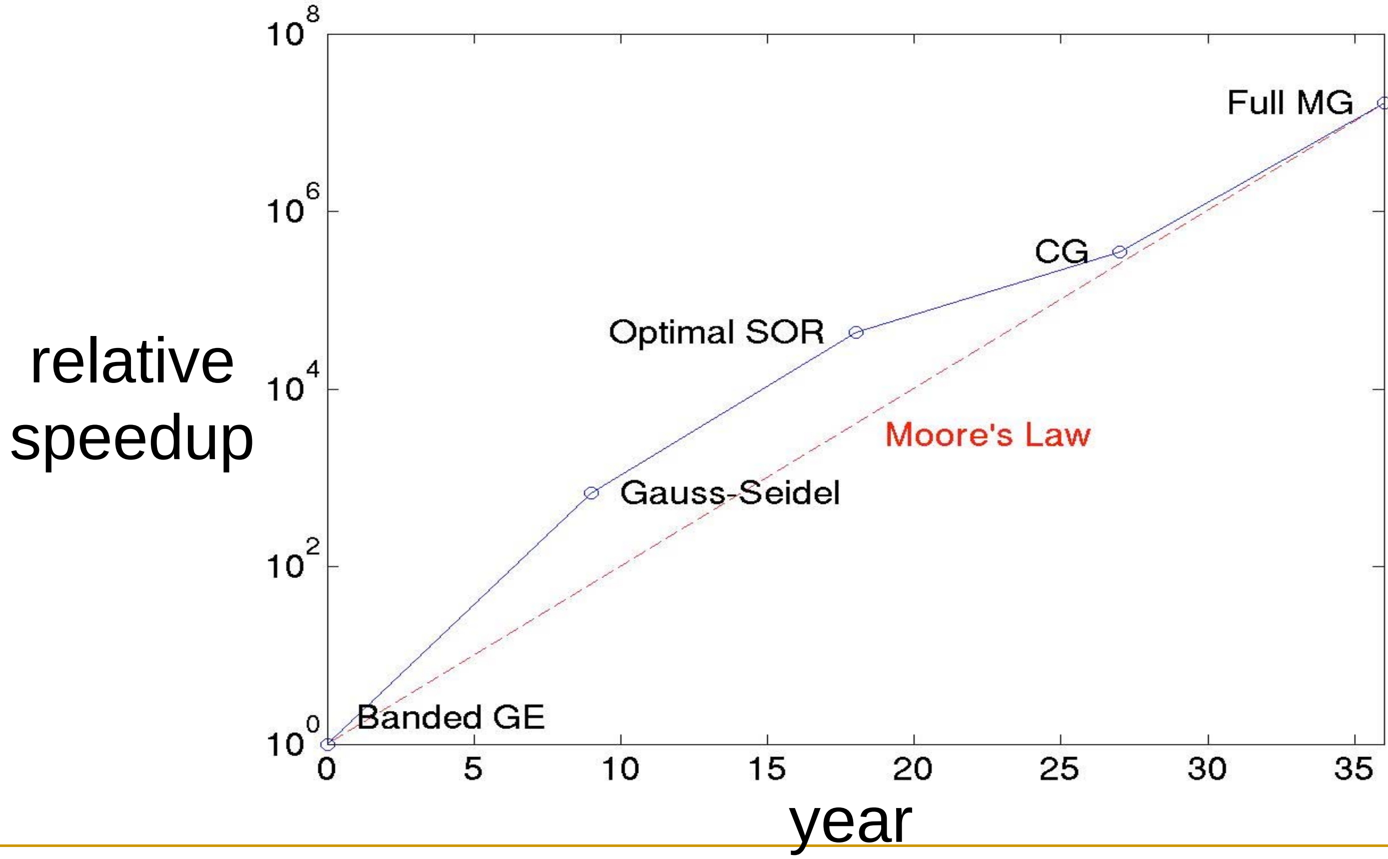



- **If $n$=64, this implies an overall reduction in flops of ~16 million***

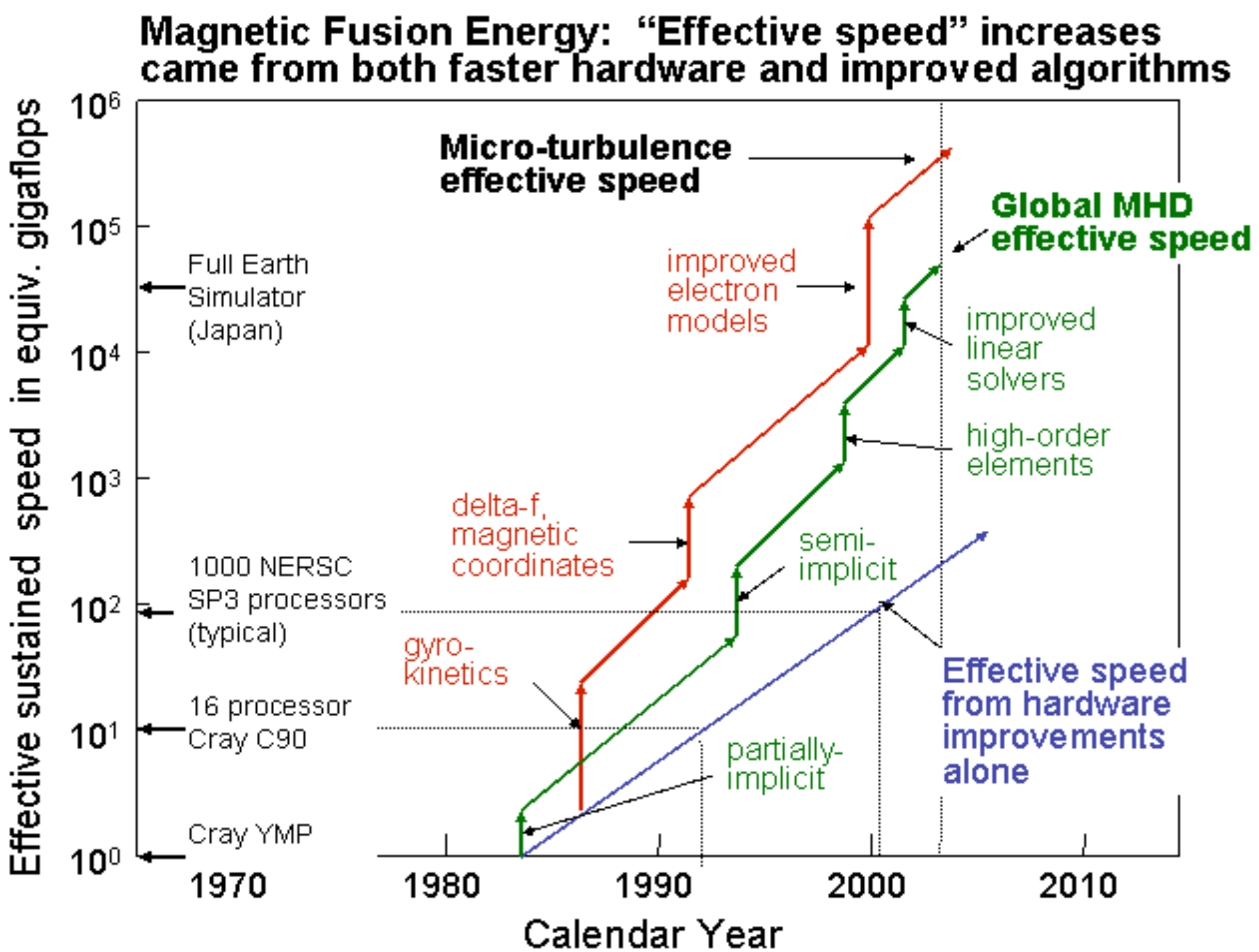
*Six-months is reduced to 1 s

# *Algorithms and Moore's Law*

- This advance took place over a span of about 36 years, or 24 doubling times for Moore's Law
- $2^{24} \approx 16$ million $\Rightarrow$ the same as the factor from algorithms alone!



# *"Moore's Law" for MHD simulations*




"Semi-implicit":

All waves treated implicitly, but still stability-limited by transport

"Partially implicit":

Fastest waves filtered, but still stability-limited by slower waves

Figure from SCaLeS report, Volume 2

# *"Moore's Law" for combustion simulations*

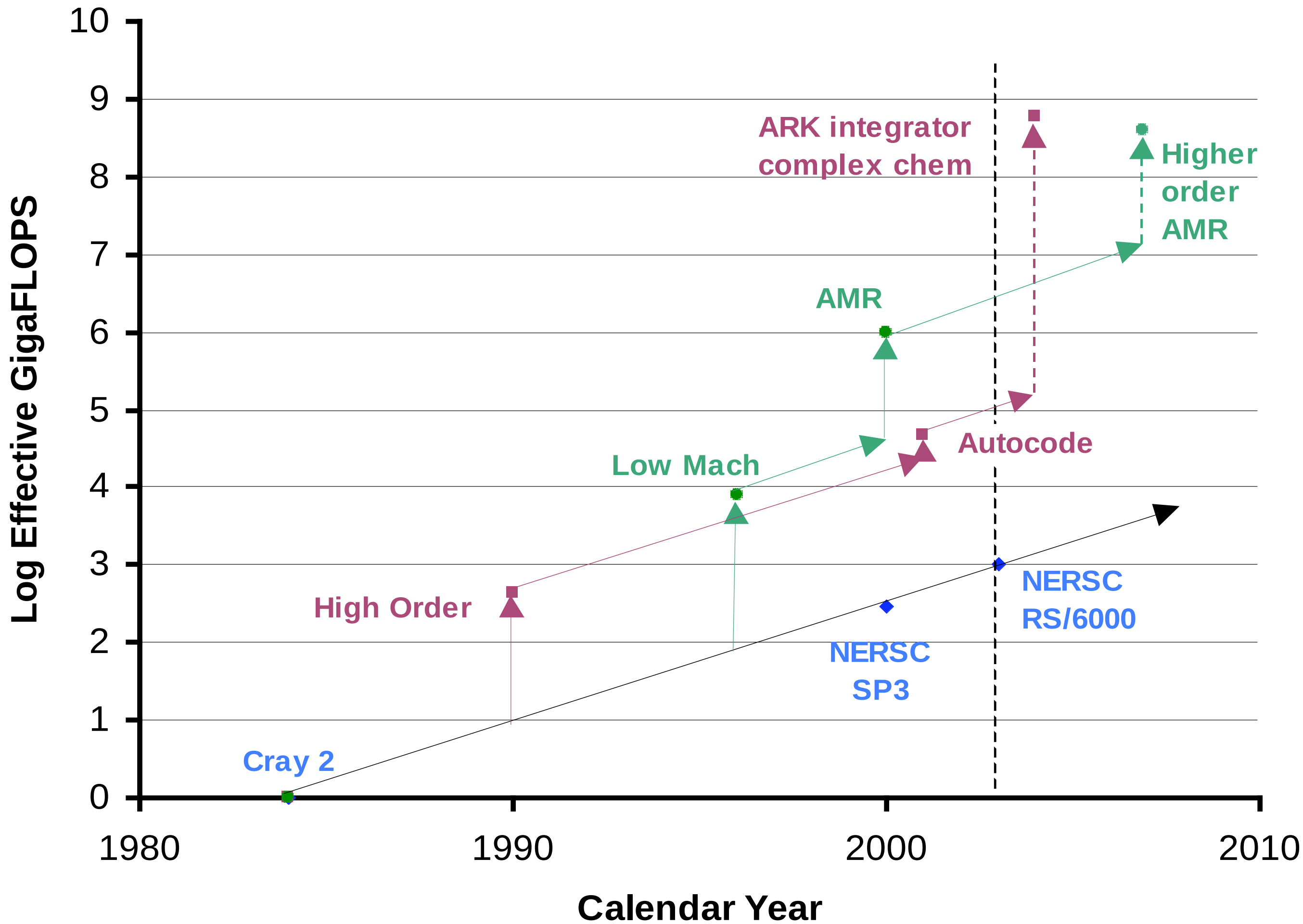



Figure from SCaLeS report, Volume 2

# Gordon Bell Prize "price performance"

| Year | Application | System | $ per Mflops MMflop/s |
|---|---|---|---|
| 1989 | Reservoir modeling | CM-2 | 2,500 |
| 1990 | Electronic structure | IPSC | 1,250 |
| 1992 | Polymer dynamics | cluster | 1,000 |
| 1993 | Image analysis | custom | 154 |
| 1994 | Quant molecular dyn | cluster | 333 |
| 1995 | Comp fluid dynamics | cluster | 278 |
| 1996 | Electronic structure | SGI | 159 |
| 1997 | Gravitation | cluster | 56 |
| 1998 | Quant chromodyn | custom | 12.5 |
| 1999 | Gravitation | custom | 6.9 |
| 2000 | Comp fluid dynamics | cluster | 1.9 |
| 2001 | Structural analysis | cluster | 0.24 |

**Four orders of magnitude in 12 years**

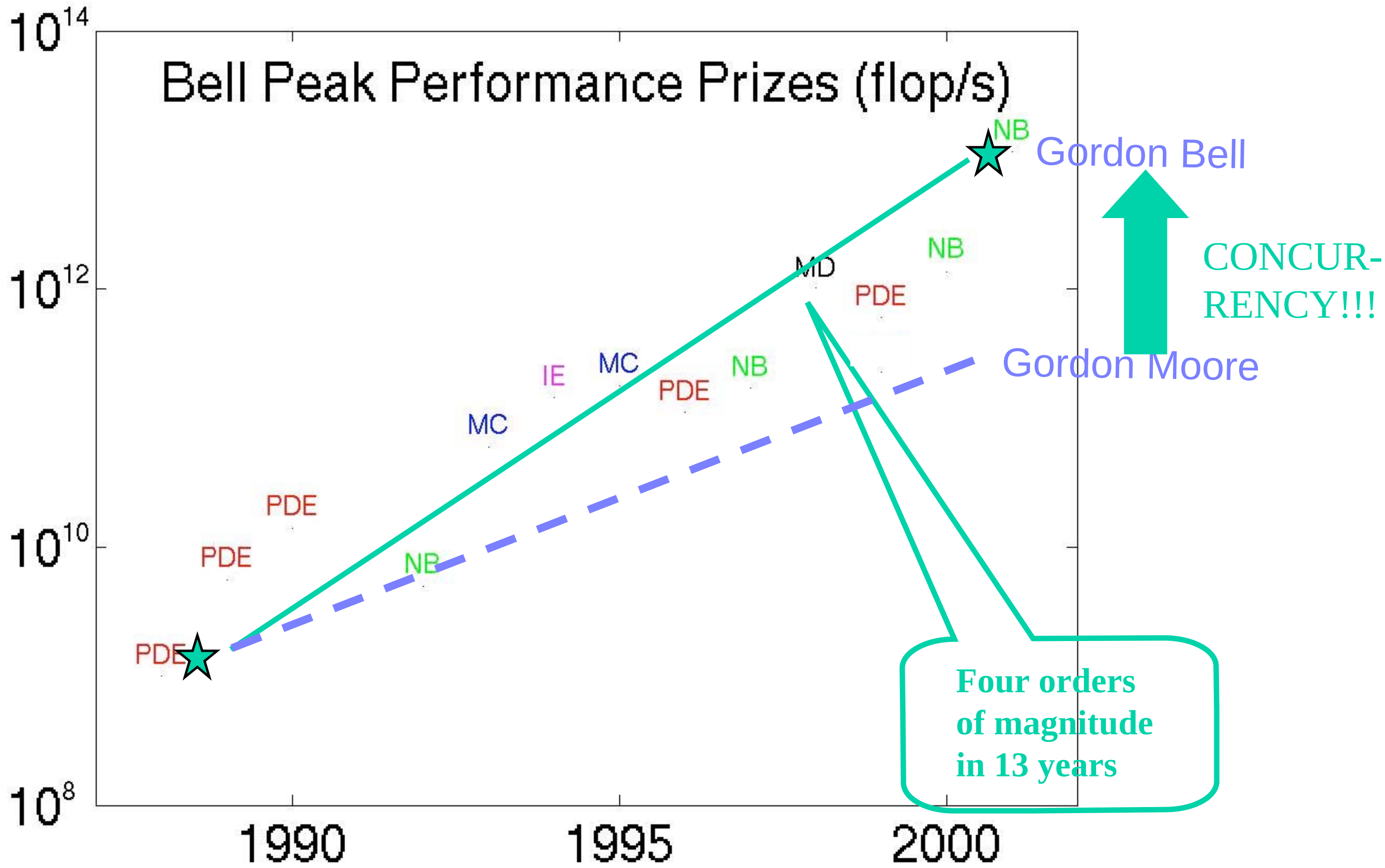
**Price/performance has stagnated and no new such prize has been given since 2001.**

## *Gordon Bell Prize "peak performance"*

| *Year* | *Type* | *Application* | *No. Procs* | *System* | *Gflop/s* |
|---|---|---|---|---|---|
| 1988 | PDE | Structures | 8 | Cray Y-MP | 1.0 |
| 1989 | PDE | Seismic | 2,048 | CM-2 | 5.6 |
| 1990 | PDE | Seismic | 2,048 | CM-2 | 14 |
| 1992 | NB | Gravitation | 512 | Delta | 5.4 |
| 1993 | MC | Boltzmann | 1,024 | CM-5 | 60 |
| 1994 | IE | Structures | 1,904 | Paragon | 143 |
| 1995 | MC | QCD | 128 | NWT | 179 |
| 1996 | PDE | CFD | 160 | NWT | 111 |
| 1997 | NB | Gravitation | 4,096 | ASCI Red | 170 |
| 1998 | MD | Magnetism | 1,536 | T3E-1200 | 1,020 |
| 1999 | PDE | CFD | 5,832 | ASCI BluePac | 627 |
| 2000 | NB | Gravitation | 96 | GRAPE-6 | 1,349 |
| 2001 | NB | Gravitation | 1,024 | GRAPE-6 | 11,550 |
| 2002 | PDE | Climate | 5,120 | Earth Sim | 26,500 |
| 2003 | PDE | Seismic | 1,944 | Earth Sim | 5,000 |
| 2004 | PDE | CFD | 4,096 | Earth Sim | 15,200 |
| 2005 | MD | Solidification | 131,072 | BGL | 101,700 |

**Four orders of magnitude in 13 years**

# *Gordon Bell Prize outpaces Moore's Law*



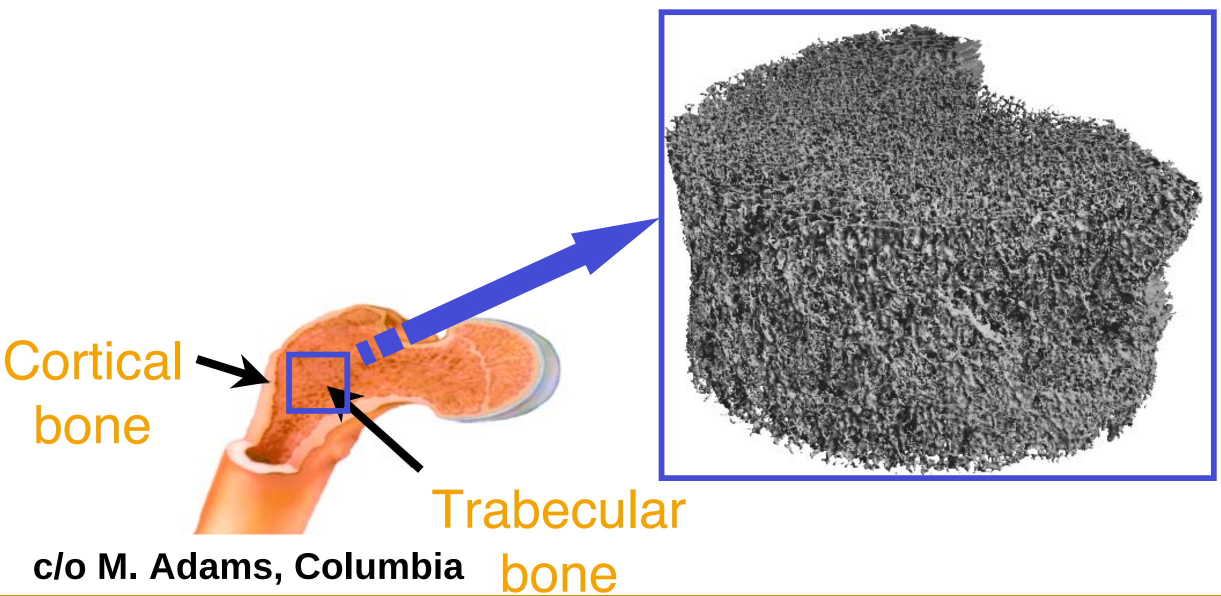
# *Heretofore difficult apps are now parallelized*

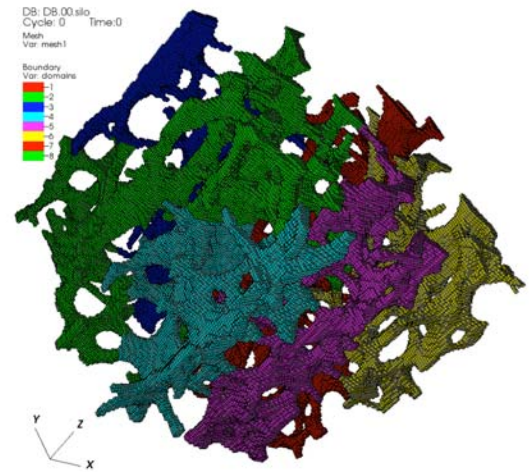
- **Unstructured grids**
- **Implicit, as well as explicit, methods**
- **Massive spatial resolution**
- **Thousand-fold concurrency**
- **Strong scaling within modest ranges**
- **Weak scaling without obvious limits**

**See, e.g., Gordon Bell "special" prizes in recent years …**

# *2004 Gordon Bell "special" prize*

- **2004 Bell Prize in "special category" went to an implicit, unstructured grid bone mechanics simulation**
  - 0.5 Tflop/s sustained on 4 thousand procs of IBM's ASCI White
  - 0.5 billion degrees of freedom
  - large-deformation analysis
  - employed in NIH bone research at Berkeley



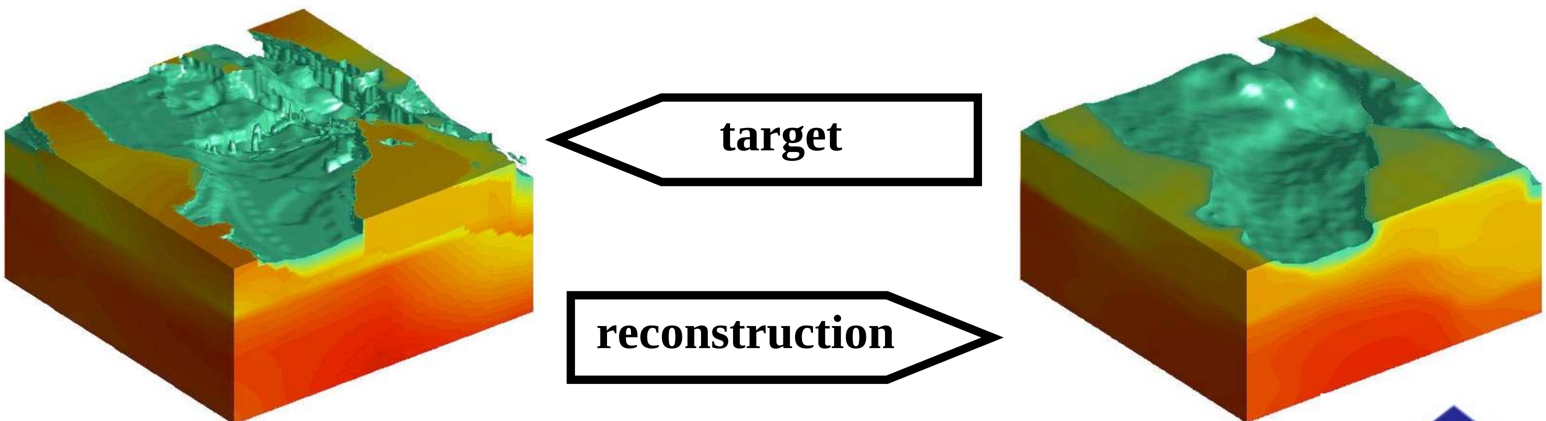



TOPS
Terascale Optimal PDE Simulations

**c/o M. Adams, Columbia**

# *2003 Gordon Bell "special" prize*

- **2003 Bell Prize in "special category" went to unstructured grid geological parameter estimation problem**
  - 1 Tflop/s sustained on 2 thousand processors of HP's "Lemieux"
  - each explicit forward PDE solve: 17 million degrees of freedom
  - seismic inverse problem: 70 billion degrees of freedom
  - employed in NSF seismic research at CMU



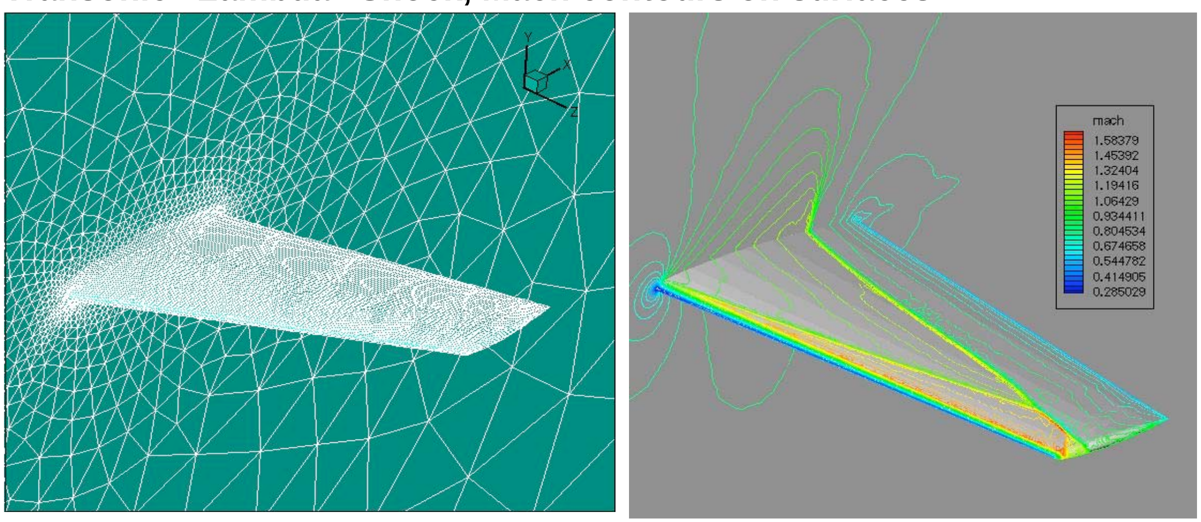

TOPS
Terascale Optimal PDE Simulations

**c/o O. Ghattas, UT Austin**

# *1999 Gordon Bell "special" prize*

- **1999 Bell Prize in "special category" went to implicit, unstructured grid aerodynamics problems**
  - 0.23 Tflop/s sustained on 3 thousand processors of Intel's ASCI Red
  - 11 million degrees of freedom
  - incompressible and compressible Euler flow
  - employed in NASA analysis/design missions

**Transonic "Lambda" Shock, Mach contours on surfaces**




TOPS
Terascale Optimal PDE Simulations

# *Why build a case?*

- **Competing needs**
  - Where can funding provide the most impact?
  - For example, build large computer or build experimental facility
- **Best way to solve problems**
  - Computing provides a general purpose infrastructure
    - *Economies of Scale*
      - Each large scale system can benefit many areas of science and engineering
- **Opportunity to accelerate science**
  - Virtual experiments
  - What ifs

# *The Role(s) of Computation in Advancing Science*

- **Assist with Complexity; answer questions**
  - Modeling and simulation of complex processes
  - Difficulty in case – no definite timeline
- **Engineering of advanced instruments**
  - Easier case — once decision is made to build an instrument, computation can improve quality and/or reduce code to build and operate
    - *Accelerator (ILC design)*
  - Harder case — Go/No Go decision may depend on likelihood of success (ITER)
    - *Fusion Reactor (ITER) – US rejoins international consortium in 2003*
      *"It [ITER design runs on supercomputers] gave us confidence that ITER will succeed" Ray Orbach*
- **Applications**
  - Many possibilities
  - Computing a well-established tool
  - Difficulty in case — Use of Large-Scale computing, separate from embarrassingly parallel parameter studies
- **Note this is often a closed loop — understanding in applications feeds back into basic science**

# *Can simulation produce more than "insight"?*

**"The purpose of computing is *insight*, not numbers."**

**— R. W. Hamming (1961)**

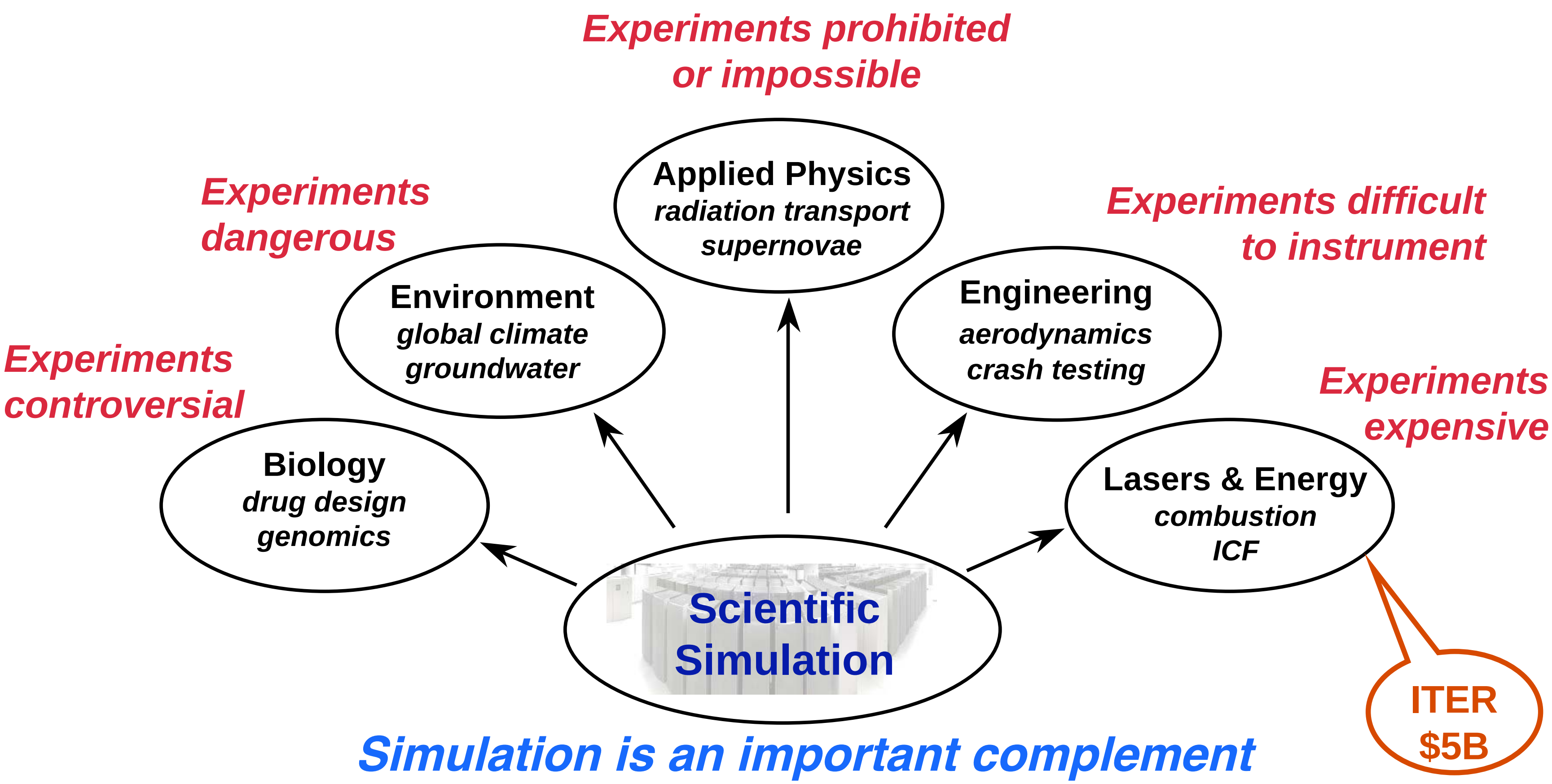
**"The computer literally is providing a new window through which we can observe the natural world in exquisite detail."**

**— J. S. Langer (1998)**

**"What changed were simulations that showed that the new ITER design will, in fact, be capable of achieving and sustaining burning plasma."**

**— R. L. Orbach (2003, in Congressional testimony about why the U.S. is rejoining the International Thermonuclear Energy Reactor (ITER) consortium)**

# *Large scale simulation can be an alternative or adjunct to experimentation*




***Simulation is an important complement to experiment in many areas***

# *DOE Office of Science Priorities: Near Term*


Near-Term

| Priority | Program | Facility |
|---|---|---|
| 1 | FES | ITER |
| 2 | ASCR | UltraScale Scientific Computing Capability |
| Tie for 3 | HEP | Joint Dark Energy Mission |
| | BES | Linac Coherent Light Source |
| | BER | Protein Production and Tags |
| | NP | Rare Isotope Accelerator |
| Tie for 7 | BER | Characterization and Imaging |
| | NP | CEBAF Upgrade |
| | ASCR | Esnet Upgrade |
| | ASCR | NERSC Upgrade |
| | BES | Transmission Electron Achromatic Microscope |
| 12 | HEP | BTeV |

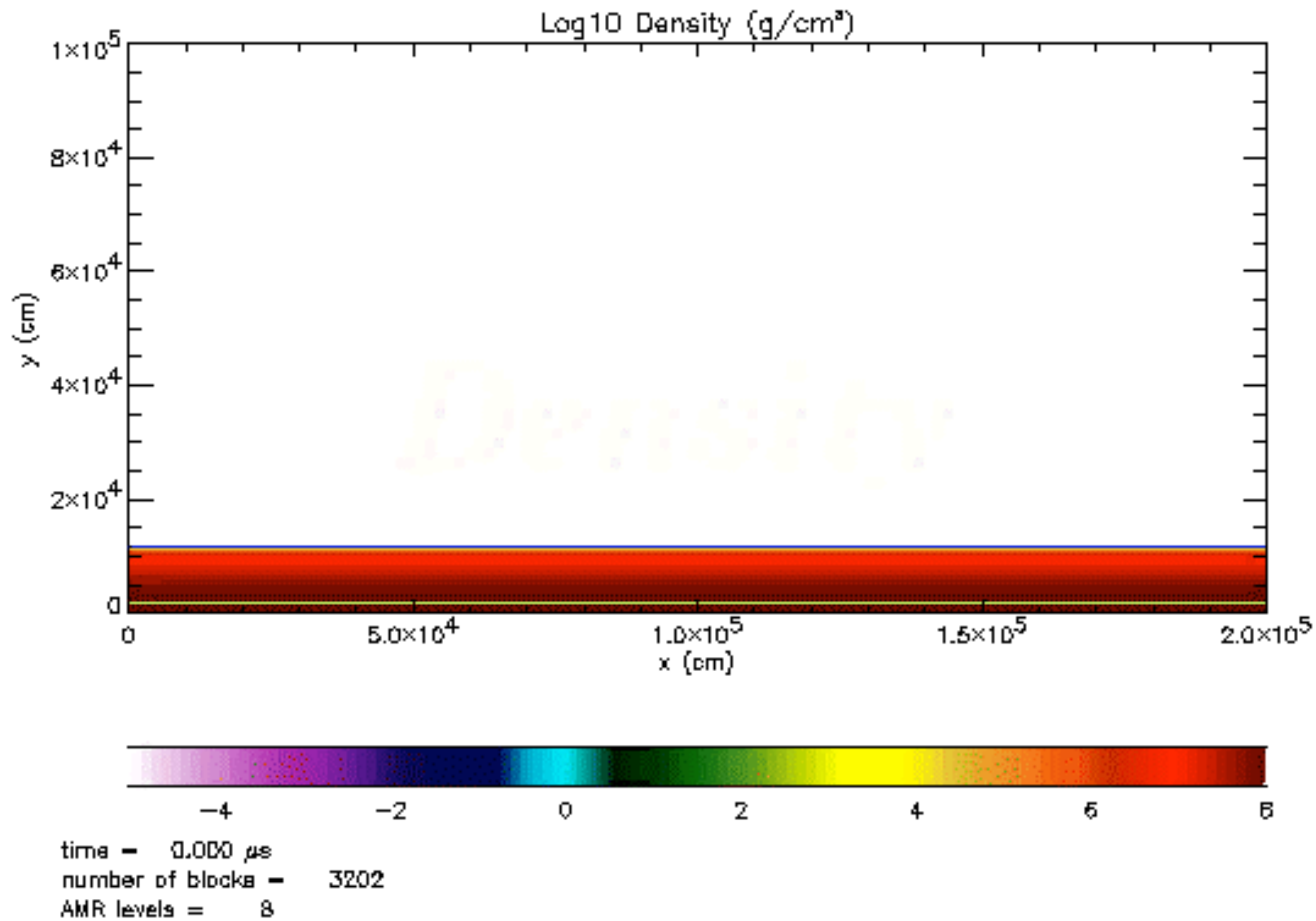
# *How to Build the Case*

- **Where can *capability* computing be used?**
- **How *not* to build a case**
  - Avoid “because we can” arguments
- **Why now?**
  - Demonstrated success
  - Capability of systems
- **Understand the interrelated roles of hardware, mathematics, computer science, and software**
  - SCaLeS report emphasized the contribution of algorithms to overall performance
- **Build on successes**

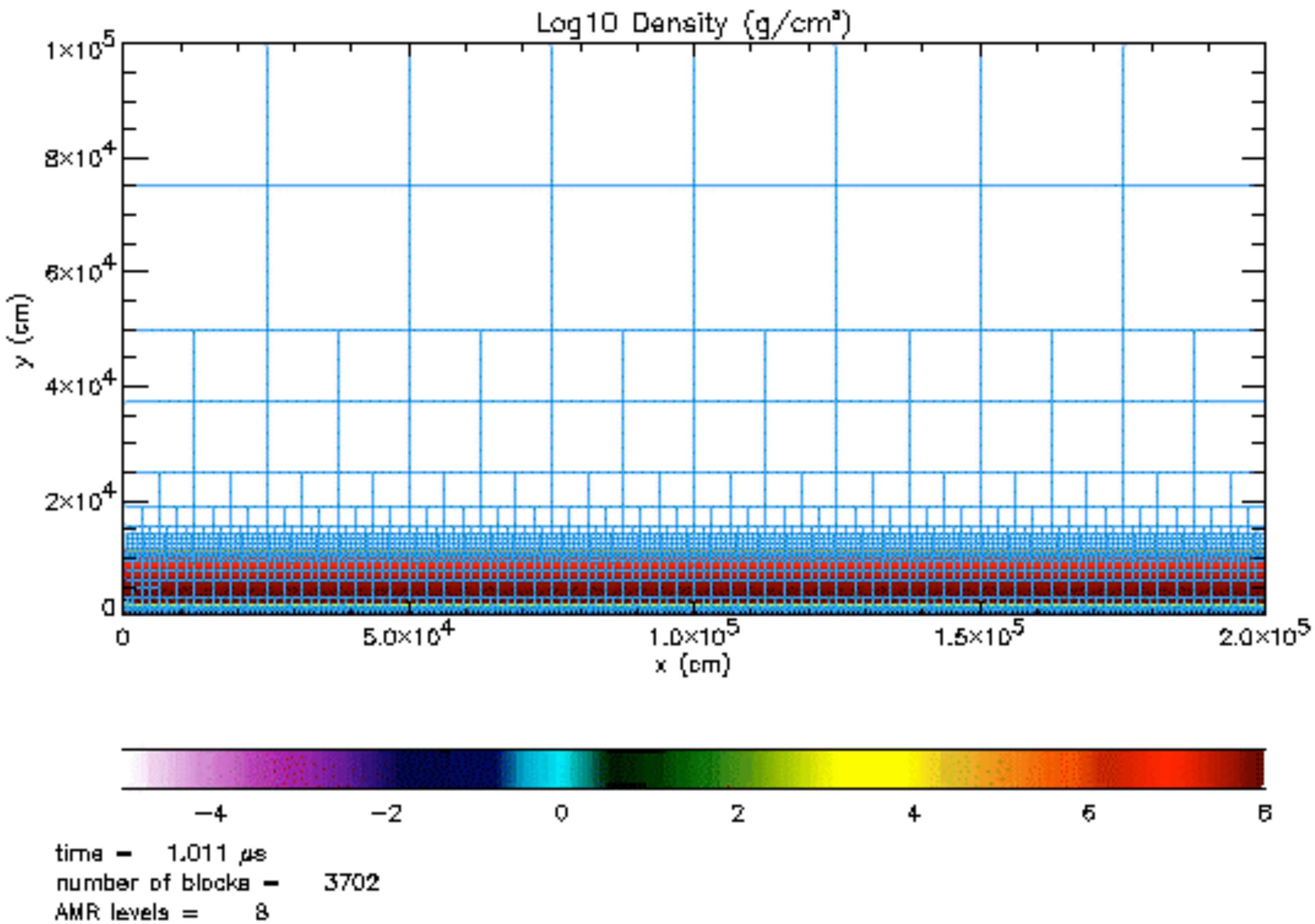
# *Some Questions Cannot be Answer Without Large Scale Computation*

- **Astrophysics is a great example**
- **Experiments are difficult**
  - Only one universe currently accessible
- **Processes are complex**
  - Gravity, Electromagnetics, Nuclear burning, wide range of time and spatial scales, …
  - Great progress has been made by combining
    - *Largest scale computing*
    - *Scalable, parallel algorithms*
    - *Good numerics, including adaptive algorithms*
  - One example is the FLASH center (http://flash.uchicago.edu/website/home/)

# X-Ray Burst on the Surface of a Neutron Star



# *Adaptive Mesh Refinement*



# *Examples from the DOE Office of Science*

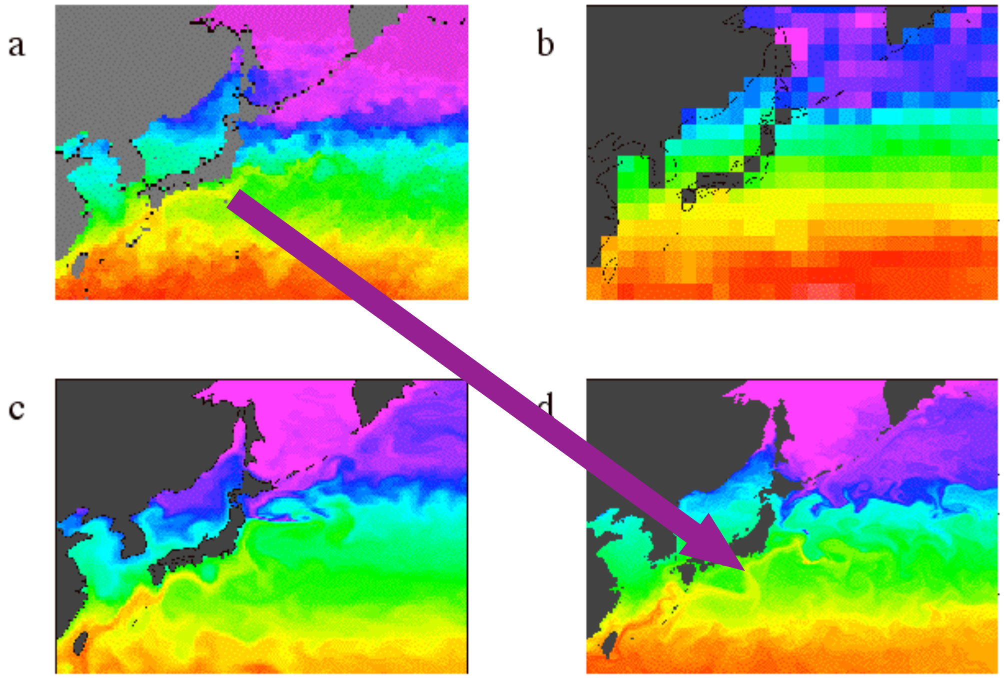
- **The Office of Science (www.er.doe.gov) is the single largest supporter of basic research in the physical sciences in the United States, providing more than 40 percent of total funding.**
- **The Office of Science manages fundamental research programs in basic energy sciences, biological and environmental sciences, and computational science.**
  - Also is the U.S. Government's largest single funder of materials and chemical sciences.
  - Supports research in climate change, geophysics, genomics, life sciences, and science education.
- **As part of the SCaLeS Workshop, scientists in these areas were asked "what could you do with 100-1000x current computational capability?"**

# *What would scientists do with 100-1000x?*
# *Example: predict future climates*

- **Resolution**
  - refine atmospheric resolution from 160 to 40 km
  - refine oceanic resolution from 105 to 15km
- **New "physics"**
  - atmospheric chemistry
  - carbon cycle
  - dynamic terrestrial vegetation (nitrogen and sulfur cycles and land-use and land-cover changes)
- **Improved representation of subgrid processes**
  - clouds
  - atmospheric radiative transfer

# *What would scientists do with 100-1000x? Example: predict future climates*

***Resolution of Kuroshio Current:*** Simulations at various resolutions have demonstrated that, because equatorial meso-scale eddies have diameters ~10-200 km, the grid spacing must be < 10 km to adequately resolve the eddy spectrum. This is illustrated in four images of the sea-surface temperature. Figure (a) shows a snapshot from satellite observations, while the three other figures are snapshots from simulations at resolutions of (b) 2°, (c) 0.28°, and (d) 0.1°.

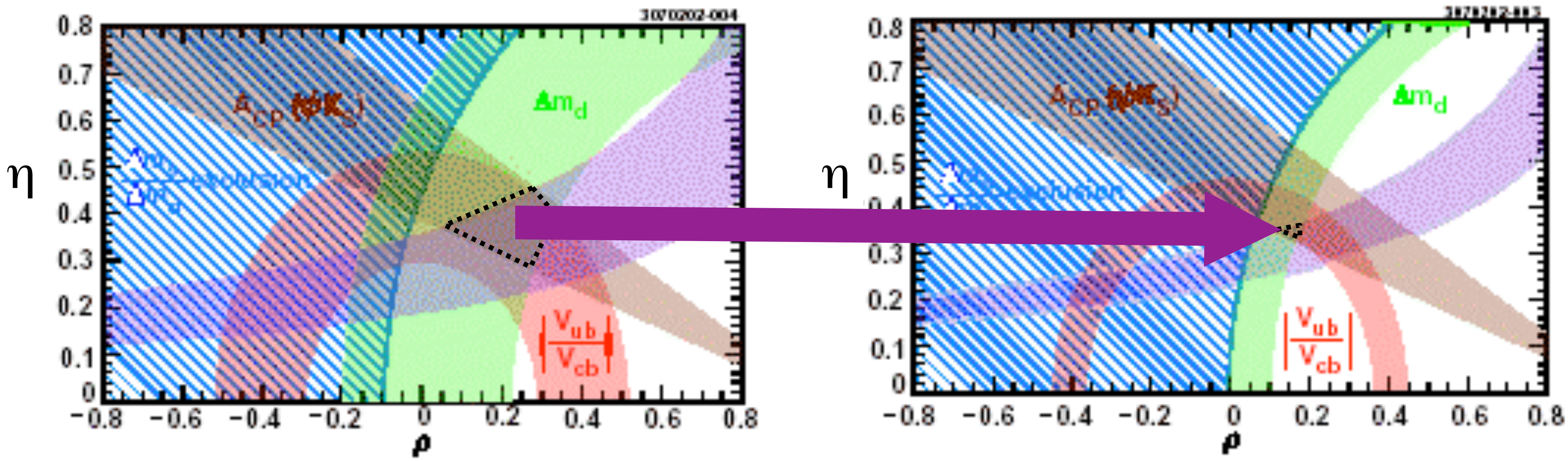


# *What would scientists do with 100-1000x? Example: probe structure of particles*

- **Resolution**
  - take current 4D quantum chromodynamics models from 32 ×32×32×16 to 128×128×128×64
- **New physics**
  - “unquench” the lattice approximation: enable study of the gluon structure of the nucleon, in addition to its quark structure
  - obtain chiral symmetry by solving on a 5D lattice in the domain wall Fermion formulation
  - allow precision calculation of the spectroscopy of strongly interacting particles with unconventional quantum numbers, guiding experimental searches for states with novel quark and gluon structure

# *What would scientists do with 100-1000x? Example: probe structure of particles*

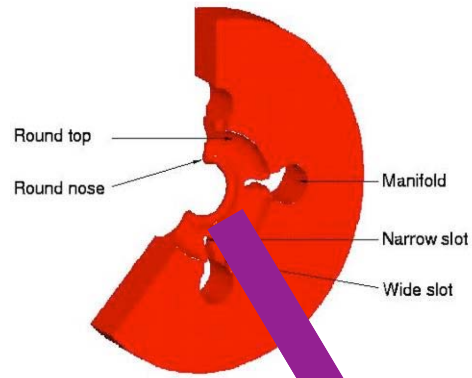
***Constraints on the Standard Model parameters ρ and η.*** For the Standard Model to be correct, these parameters from the Cabibbo-Kobayashi-Maskawa (CKM) matrix must be restricted to the region of overlap of the solidly colored bands. The figure on the left shows the constraints as they exist today. The figure on the right shows the constraints as they would exist with no improvement in the experimental errors, but with lattice gauge theory uncertainties reduced to 3%.



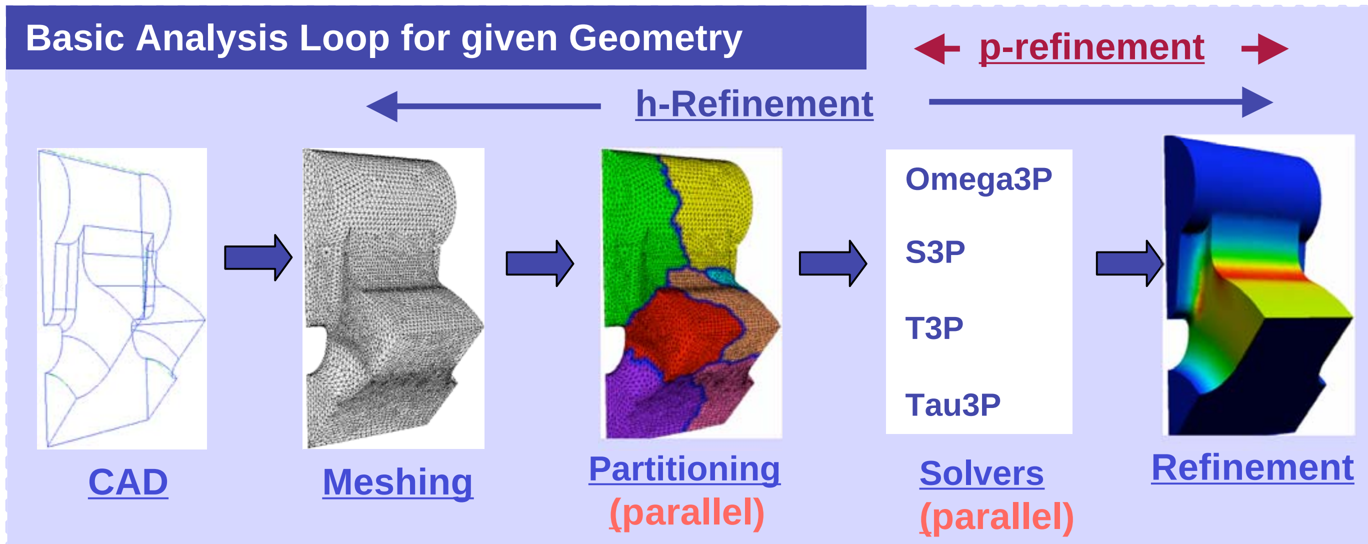
# *What would scientists do with 100-1000x? Example: design accelerators*

- **Resolution**
  - complex geometry (long assemblies of damped detuned structure (DDS) cells, each one slightly different than its axial neighbor) requires unstructured meshes with hundreds of millions of degrees of freedom
  - Maxwell eigensystems for interior elements of the spectrum must be solved in the complex cavity formed by the union of the DDS cells
- **Novel capability**
  - PDE-based mathematical optimization will replace expensive and slow trial and error prototyping approach
  - each inner loop of optimization requires numerous eigensystem analyses

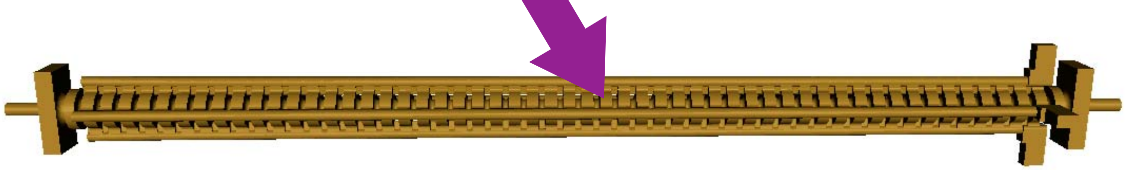
# *What would scientists do with 100-1000x? Example: design accelerators*




DDS CEL

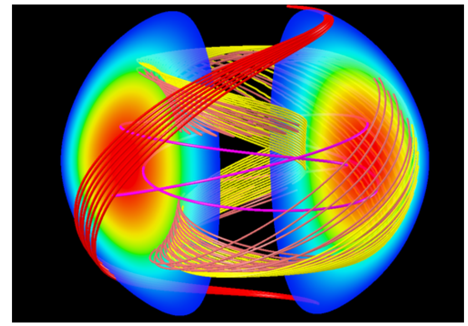


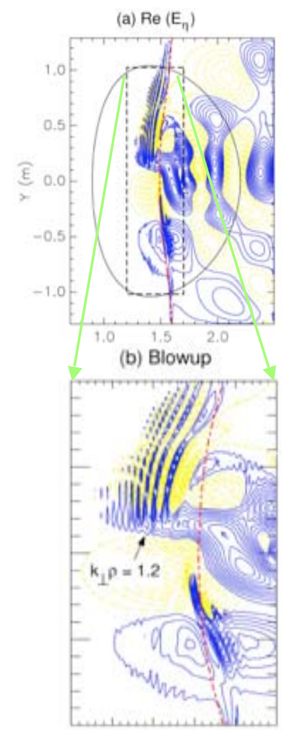

***Next generation accelerators have complex cavities.*** Shape optimization is required to improve performance and reduce operating cost.



# *What would scientists do with 100-1000x? Example: design and control tokamaks*

- **Resolution**
  - refine meshes and approach physical Lundquist numbers
- **Multiphysics**
  - combine MHD, PIC, and RF codes in a single, consistent simulation
  - resolve plasma edge
- **Design and control**
  - optimize performance of experimental reactor ITER and follow-on production devices
  - detect onset of instabilities and modify before catastrophic energy releases from the magnetic field





# *What would scientists do with 100-1000x? Example: control combustion*

- **Resolution**
  - evolve 3D time-dependent large-eddy simulation (LES) codes to direct Navier-Stokes (DNS)
  - multi-billions of mesh zones required
- **New “physics”**
  - explore coupling between chemistry and acoustics (currently filtered out)
  - explore sooting mechanisms to capture radiation effects
  - capture autoignition with realistic fuels
- **Integrate with experiments**
  - pioneer simulation-controlled experiments to look for predicted effects in the laboratory

# *What would scientists do with 100-1000x? Example: control combustion*

Experimental PIV measurement

Instantaneous flame front imaged by density of inert marker

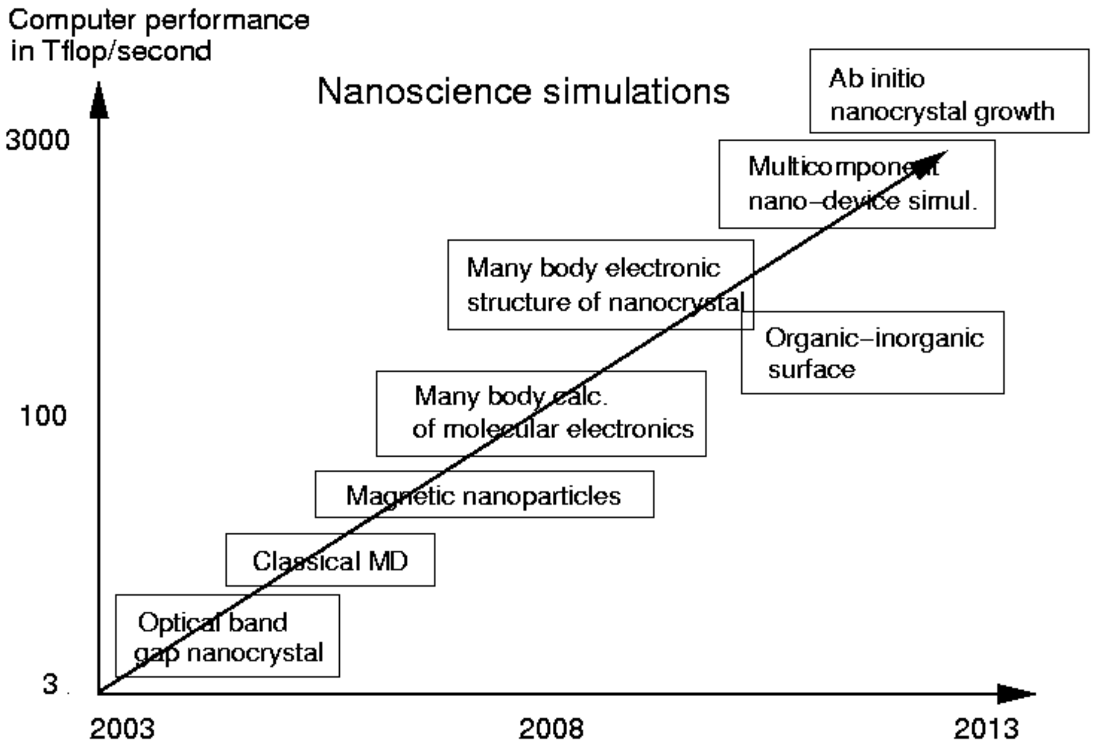
Simulation

Instantaneous flame front imaged by fuel concentration

Images c/o R. Cheng (left), J. Bell (right), LBNL, and NERSC

**2003 SIAM/ACM Prize in CS&E (J. Bell & P. Colella)**

## *What would scientists do with 100-1000x?*
## *Example: understand and predict complex nanostructures*

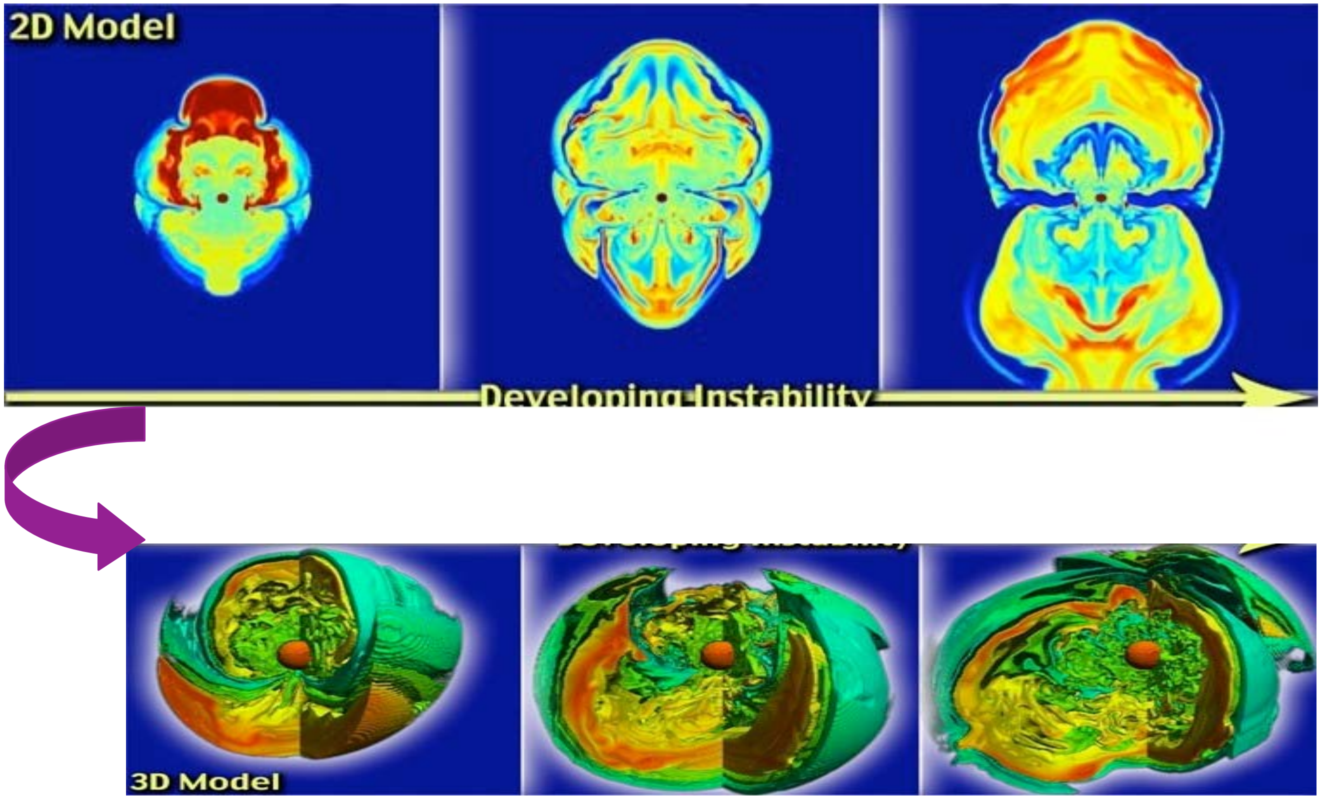



*One projection for the type of problems that might be addressable in computational nanoscience in the future as tera- and peta-scale computational capabilities become available. From the SCaLeS Report, volume 2.*

# *What would scientists do with 100-1000x? Example: probe supernovae*

- **Resolution**
  - current Boltzmann neutrino transport models are vastly under-resolved
  - need at least $512^3$ spatially, at least 8 polar and 8 azimuthal, and at least 24 energy groups energy groups per each of six neutrino types
  - to discriminate between competing mechanisms, must conserve energy to within 0.1% over millions of time steps
- **Full dimensionality**
  - current models capable of multigroup neutrino radiation are lower-dimensional; full 3D models are required

# *What would scientists do with 100-1000x? Example: probe supernovae*




***Stationary accretion shock instability defines shape of supernovae and direction of emitted radiation.*** Lower dimensional models produce insight; full dimensional models are ultimately capable of providing radiation signatures that can be compared with observations.

**c/o A. Mezzacappa, ORNL**

# *Large Scale Computing in the DOE Office of Science*

- **Ultrascale Computing spawned the National Leadership Computing Facility (NLCF)**
  - ORNL + ANL partnership won competitive bid
  - Installed a mix of vector (Cray X1) and distributed memory (Cray XT3, IBM BG/L) systems
- **Leap in achieved performance**
  - IBM BG/L demonstrates application scalability to 128K processors
    - *Solidification application achieves sustained 101.7 TF*
    - *BG/L Consortium http://www.mcs.anl.gov/bgconsortium*
  - Many installed BG/L systems, including
    - *ASTRON/LOFAR http://www.lofar.org/p/systems.htm*
- **Ultrascale computing is a DOE Facilities priority**
- **Challenges**
  - Avoiding a focus on only hardware
    - *Algorithmic advances less certain but more cost effective*
  - Inertia of existing application codes
- **Continuing support for integrated approach, multidisciplinary approach**
  - SciDAC, upcoming program renewal SciDAC2
  - SciDAC 2005 proceedings: http://www.iop.org/EJ/toc/1742-6596/16/1

# *Features of DOE's SciDAC initiative*

- **Affirmation of importance of simulation**
  - for new scientific discovery, not just for "fitting" experiments
- **Recognition that leading-edge simulation is interdisciplinary**
  - physicists and chemists not supported to write their own software infrastructure; deliverables intertwined with those of math & CS experts
- **Commitment to distributed hierarchical memory computers**
  - new code must target this architecture type
- **Commitment to maintenance of software infrastructure (*rare* to find this ☺)**
- **Requirement of lab-university collaborations**
  - complementary strengths in simulation
  - 13 laboratories and 50 universities in first round of projects

# *High Productivity Computing Systems*

- **DARPA project to attack “productivity crisis”**
  - Combined hardware and software solution
  - Intent is to improve product, not to produced a one-of-a-kind or limited production system
- **First round involved 5 vendors and partners**
- **Second (current) routine involves 3 vendors (Cray, IBM, Sun) and partners**
- **Third round will involve 1-2 vendors**
- **Components include**
  - New hardware technology (particularly to overcome problems caused by memory hierarchies)
  - New programming languages
    - *Chapel (Cray), X-10 (IBM), Fortress (Sun)*
    - *Expectation is at most one will emerge; may instead guide changes to existing languages*

# *Summary*

- Simulation will become increasingly cost-effective relative to experiment, while never fully replacing experiment
- Simulation may define today's limit to progress in areas that are already theoretically well modeled
- Simulation aids model refinement in areas not already well modeled (via interplay with theory)
- Advanced simulation makes scientists and engineers more productive
- Simulation depends on synergistic advances in large-scale computing hardware, applied mathematics, and computer science