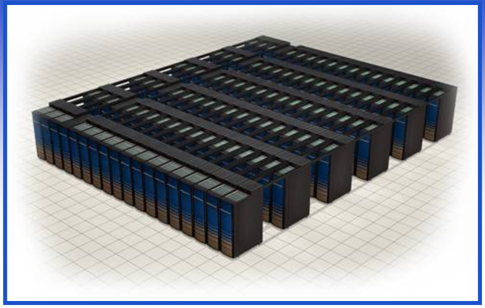






What Supercomputers Do, and What Supercomputers Still Can't Do

Horst D. Simon
Associate Laboratory Director, Computing Sciences
Lawrence Berkeley National Laboratory

CSGF Fellows Meeting
Washington, D.C.
June 21, 2007

BERKELEY LAB

NERSC

Office of Science
U.S. DEPARTMENT OF ENERGY

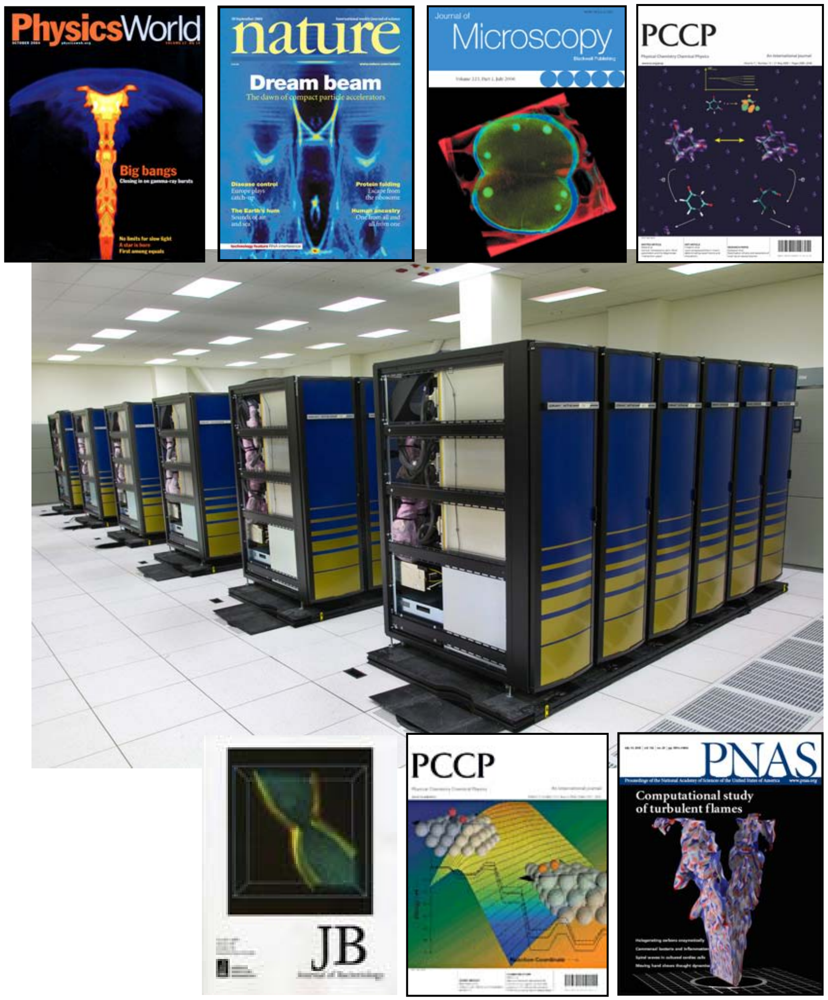
Overview

- Introducing NERSC and Computing Sciences at Berkeley Lab
- Current Trends in Supercomputing (High-End Computing)
- What Supercomputers Do
- What Supercomputers Still Can't Do

Computing Sciences at Berkeley Lab (LBNL)

- Mostly funded by DOE, Office of Science, annual budget $100M, more than 200 staff
  - NERSC Center
  - ESnet
  - Computational Research (including SciDAC projects)
- Supports open, unclassified, basic research
- Close collaborations between UC Berkeley and LBNL in applied mathematics, computer science, and computational science

Science-Driven Computing



National Energy Research Scientific Computer Center (NERSC)

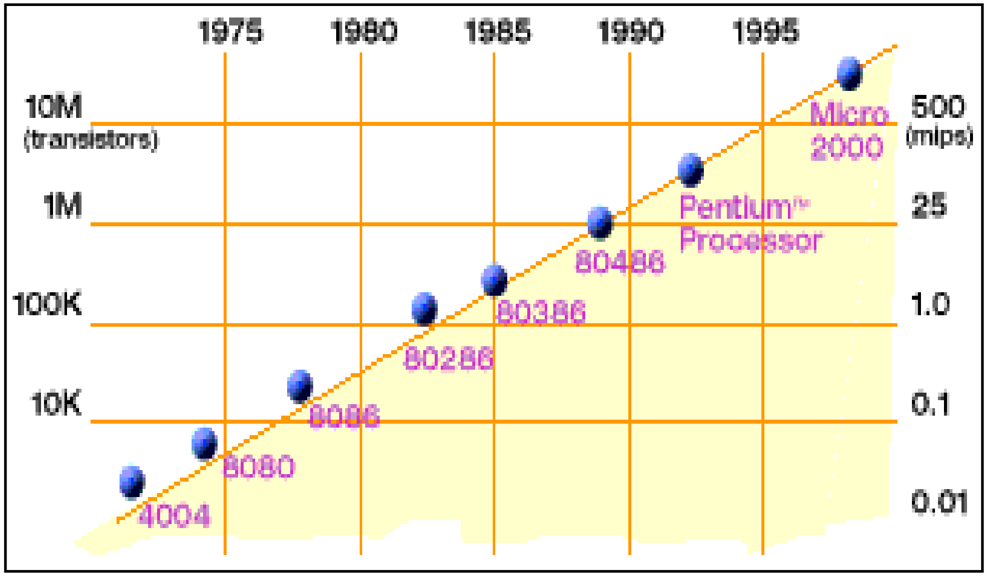
- **computational facility for open science**
- **international user community**
- **2500 users**
- **300 projects**

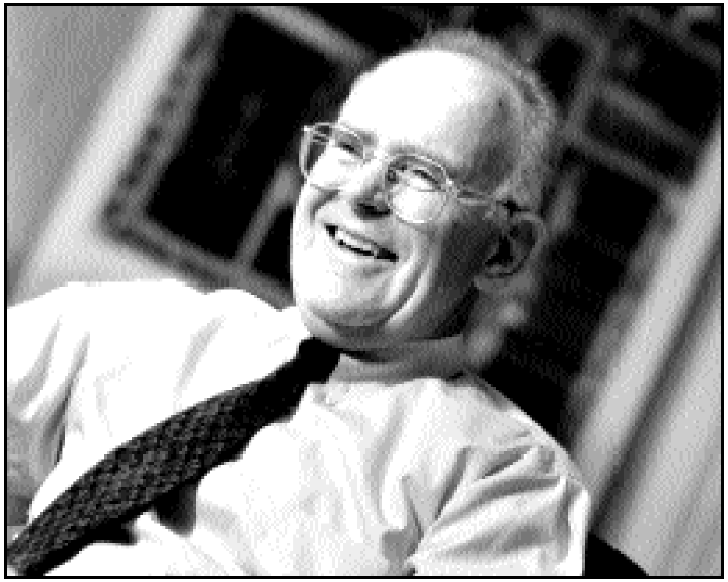
NERSC is enabling new science

Overview

- **Introducing NERSC and Computing Sciences at Berkeley Lab**
- **Current Trends in Supercomputing (High-End Computing)**
- **What Supercomputers Do**
- **What Supercomputers Can't Do**

Technology Trends: Microprocessor Capability



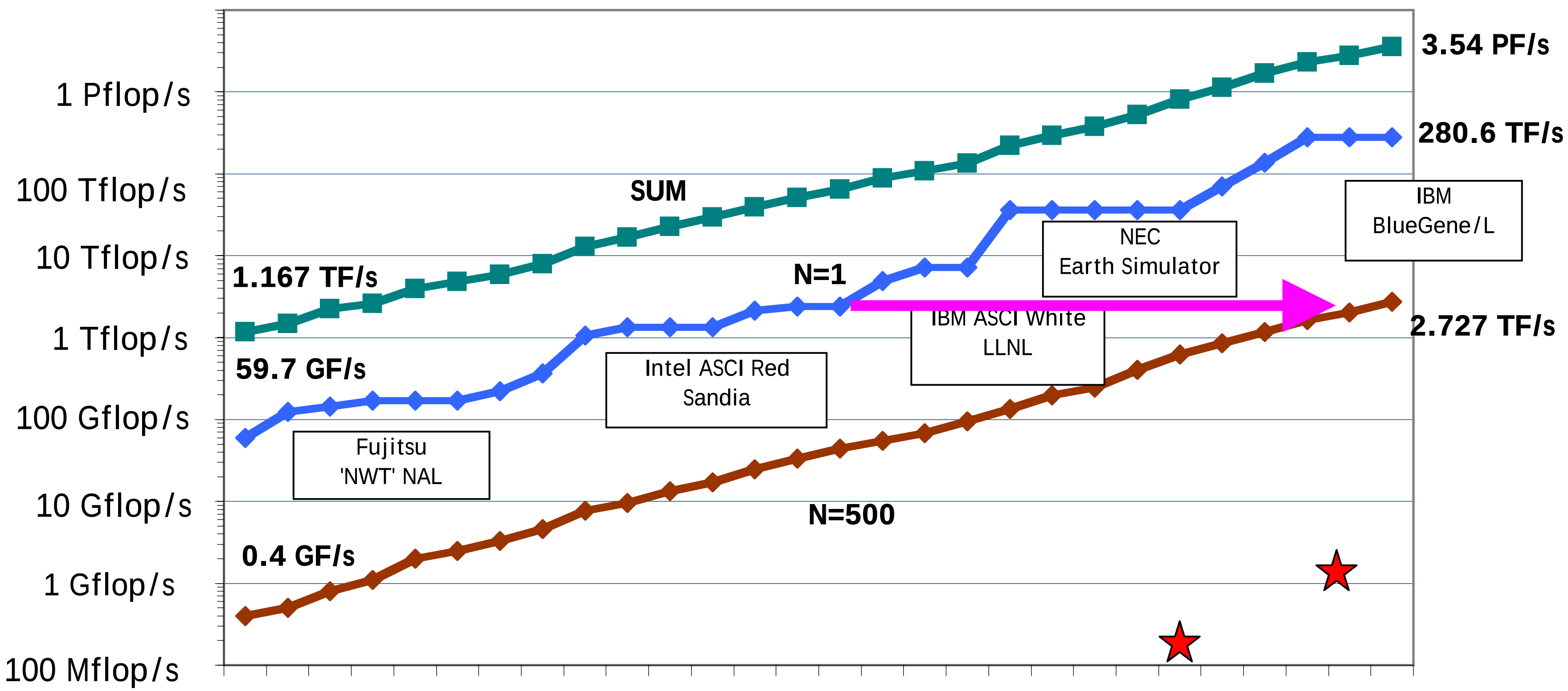


2X transistors/chip every 1.5 years
Called "Moore's Law"

Microprocessors have become smaller, denser, and more powerful.

Gordon Moore (co-founder of Intel) predicted in 1965 that the transistor density of semiconductor chips would double roughly every 18 months.

Performance Development (TOP500)



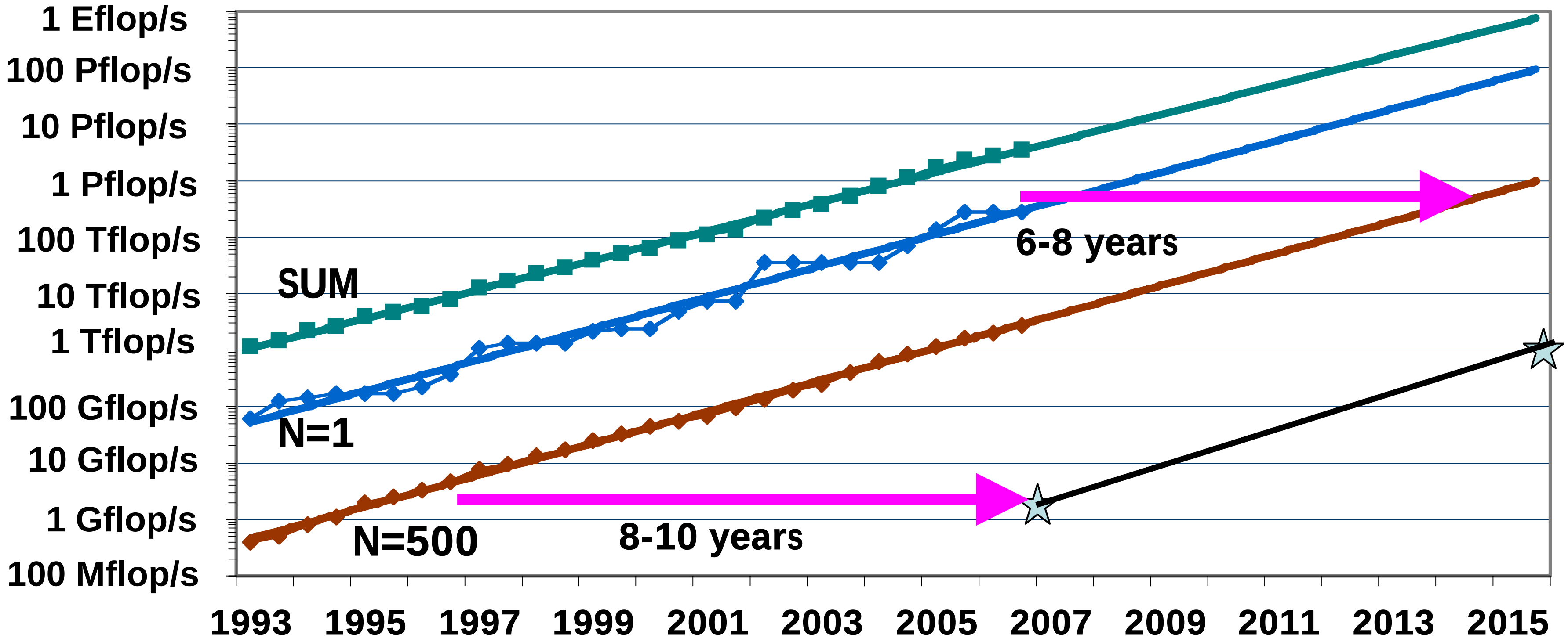
see www.top500.org

Signpost System in 2005

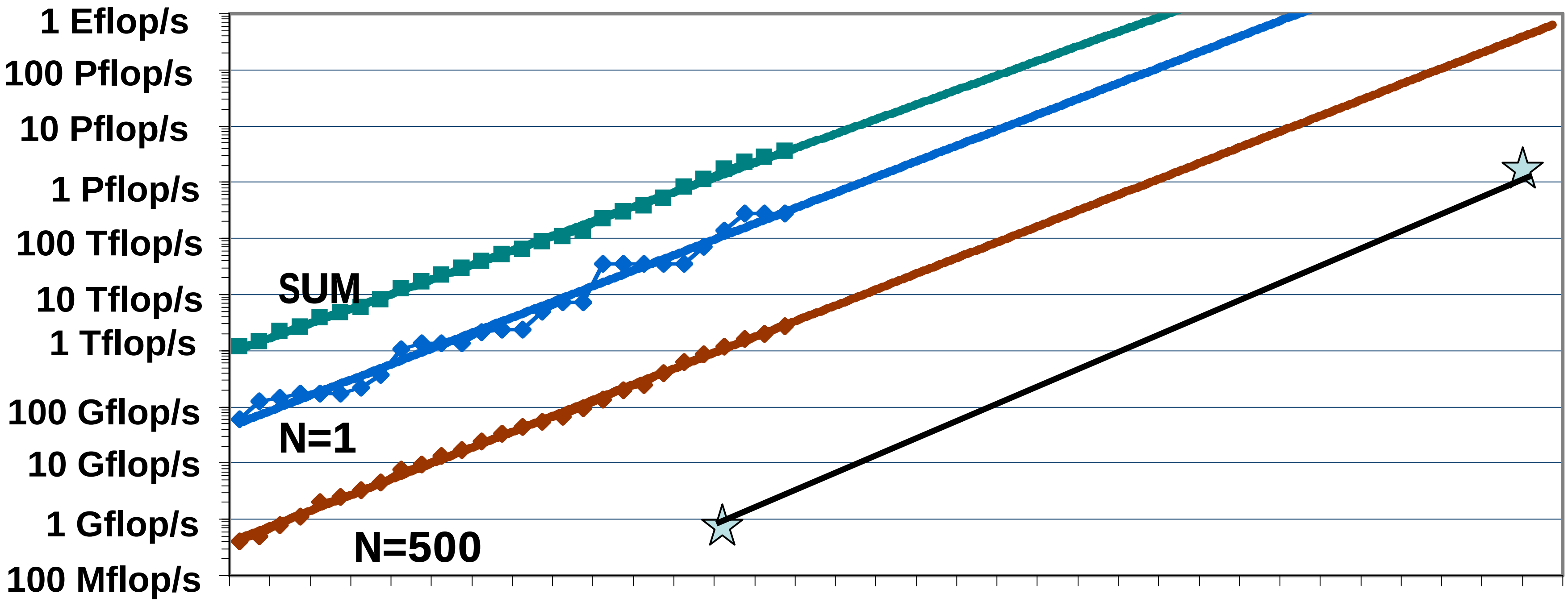
IBM BG/L @ LLNL

- 700 MHz
- 65,536 nodes
- 180 (360) Tflop/s peak
- 32 TB memory
- 135 Tflop/s LINPACK
- 250 m^2 floor space
- 1.8 MW power

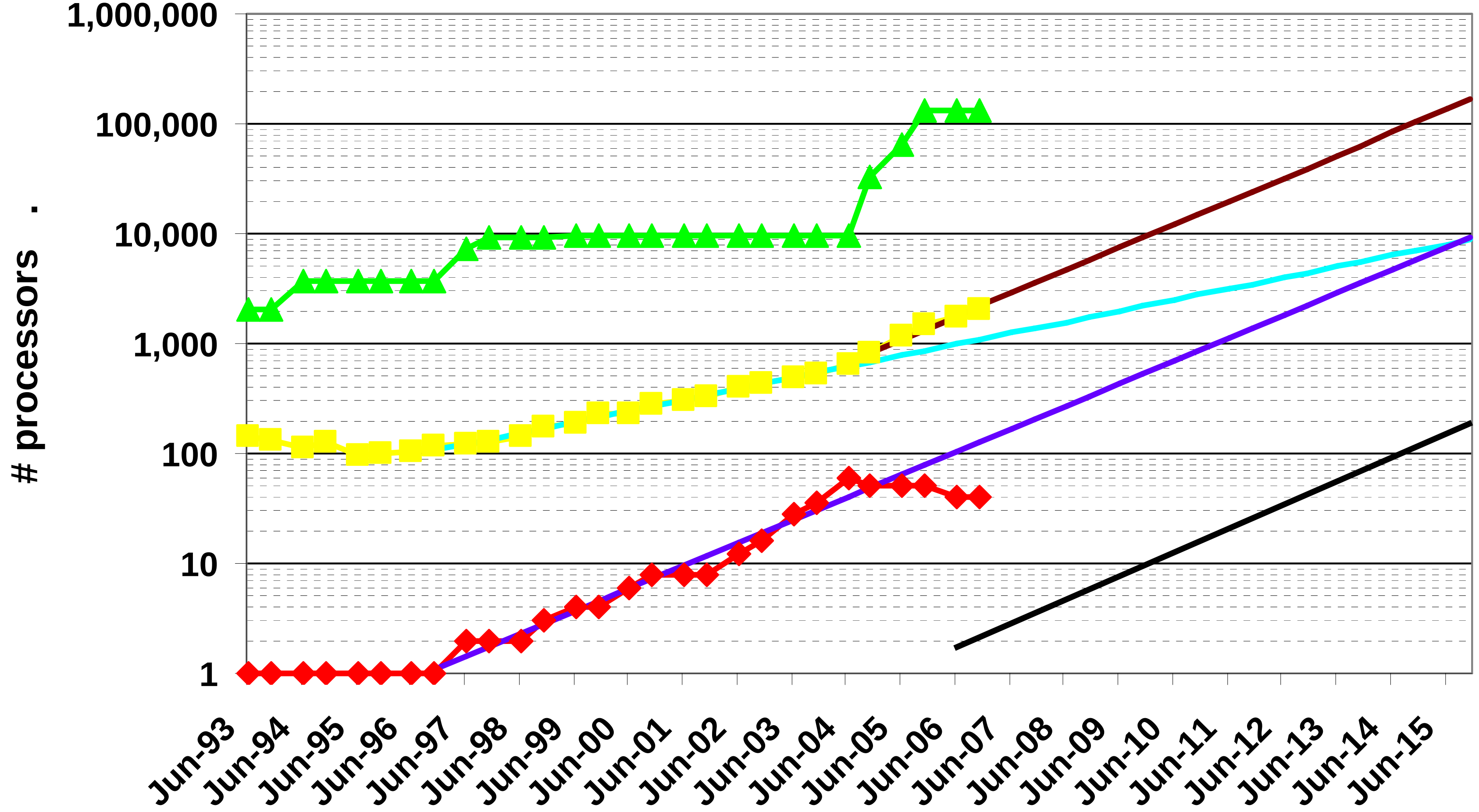
Performance Projection



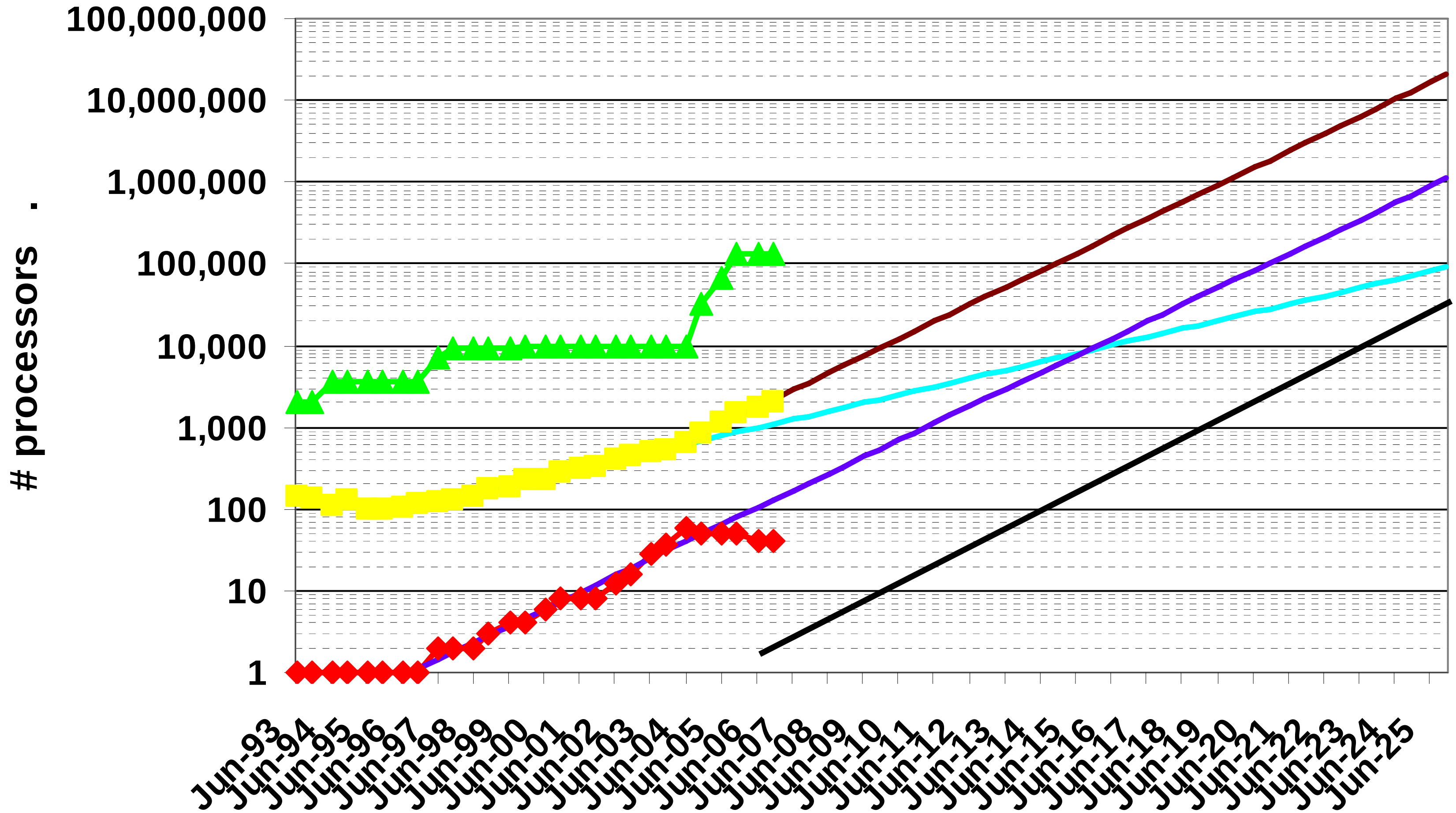
Performance Projection



Concurrency Levels



Concurrency Levels- There is a Massively Parallel System Also in Your Future



NERSC

BERKELEY LAB

Office of Science
U.S. DEPARTMENT OF ENERGY

Traditional Sources of Performance Improvement are Flat-Lining

- New Constraints
  - **15 years of *exponential* clock rate growth has ended**
- But Moore's Law continues!
  - **How do we use all of those transistors to keep performance increasing at historical rates?**
  - **Industry Response: #cores per chip doubles every 18 months *instead* of clock frequency!**
- Is multicore the correct response?

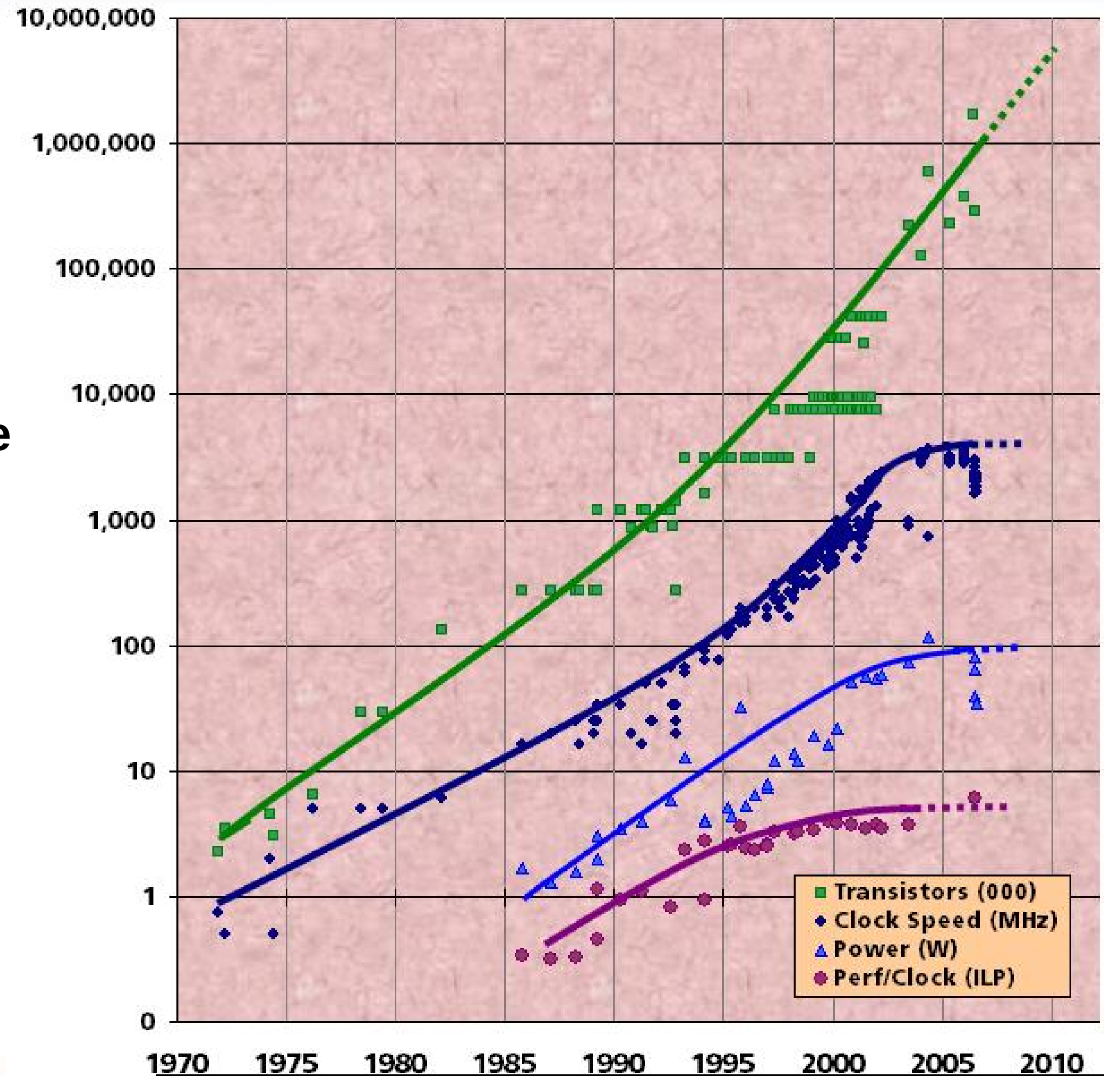
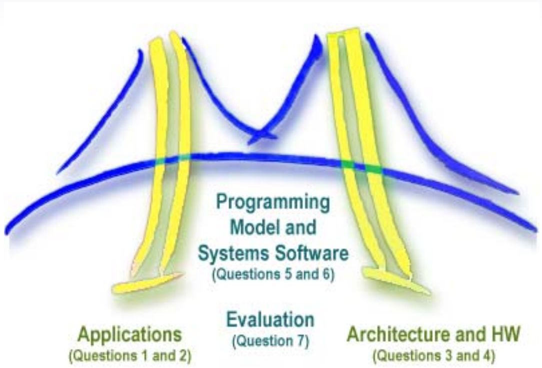


Figure courtesy of Kunle Olukotun, Lance Hammond, Herb Sutter, and Burton Smith

Is Multicore the Correct Response?



- **"The View from Berkeley",** http://view.eecs.berkeley.edu/wiki/Main_Page
- **Kurt Keutzer:** "This shift toward increasing parallelism is not a triumphant stride forward based on breakthroughs in novel software and architectures for parallelism; instead, this plunge into parallelism is actually a retreat from even greater challenges that thwart efficient silicon implementation of traditional uniprocessor architectures."
- **David Patterson:** "Industry has already thrown the hail-mary pass. . . But nobody is running yet."

Supercomputing Today

- Microprocessors have made desktop computing in 2007 what supercomputing was in 1995.
- Massive Parallelism has changed the “high-end” completely.
- Most of today's standard supercomputing architecture are “hybrids”, clusters built out of commodity microprocessors and custom interconnects.
- The microprocessor revolution will continue with little attenuation for at least another 10 years
- The future will be massively parallel, based on multicore

Overview

- **Introducing NERSC and Computing Sciences at Berkeley Lab**
- **Current Trends in Supercomputing (High-End Computing)**
- **What Supercomputers Do**
- **What Supercomputers Can’t Do**

Computational Science- Recent News

"An important development in sciences is occurring at the intersection of computer science and the sciences that has the potential to have a profound impact on science. It is a leap from the application of computing … to the *integration of computer science concepts, tools, and theorems* into the very fabric of science." -*Science* 2020 Report, March 2006

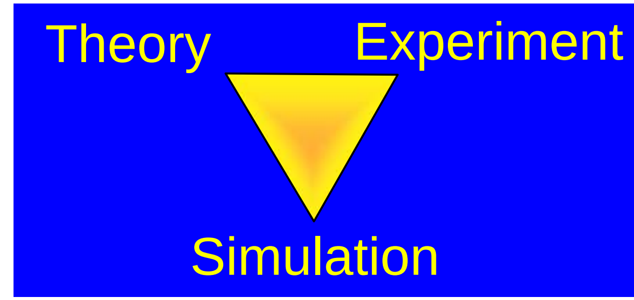
Nature, March 23, 2006

Drivers for Change

- Continued exponential increase in computational power → simulation is becoming third pillar of science, complementing theory and experiment
- Continued exponential increase in experimental data → techniques and technology in data analysis, visualization, analytics, networking, and collaboration tools are becoming essential in all data rich scientific applications

Simulation: The Third Pillar of Science

- **Traditional scientific and engineering paradigm:**
  - **(1) Do theory or paper design**
  - **(2) Perform experiments or build system**



- **Limitations:**
  - **–Too difficult—build large wind tunnels**
  - **–Too expensive—build a throw-away passenger jet**
  - **–Too slow—wait for climate or galactic evolution**
  - **–Too dangerous—weapons, drug design, climate experimentation**

- **Computational science paradigm:**
  - **(3) Use high performance computer systems to simulate and analyze the phenomenon**
    - **Based on known physical laws and efficient numerical methods**
    - **Analyze simulation results with computational tools and methods beyond what is used traditionally for experimental data analysis**

What Supercomputers Do

Introducing Computational Science and Engineering

Three Examples

- simulation replacing experiment that is too difficult
- simulation replacing experiment that is too dangerous
- analyzing massive amounts of data with new tools

Computational Science and Engineering (CSE)

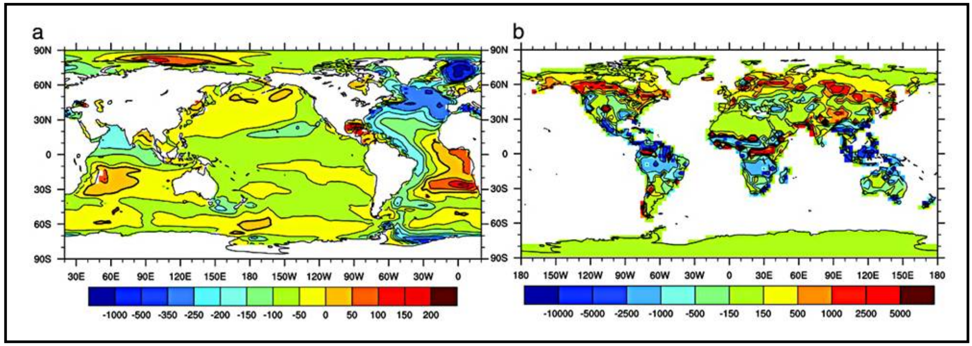
- **CSE is a widely accepted label for an evolving field concerned with the science of and the engineering of systems and methodologies to solve computational problems arising throughout science and engineering**
- **CSE is characterized by**
  - **Multi - disciplinary**
  - **Multi - institutional**
  - **Requiring high-end resources**
  - **Large teams**
  - **Focus on community software**
- **CSE is not “just programming” (and not CS)**
- **Teraflop/s computing is necessary but not sufficient**

Reference: Petzold, L., *et al*., Graduate Education in CSE, *SIAM Rev*., 43(2001), 163-177

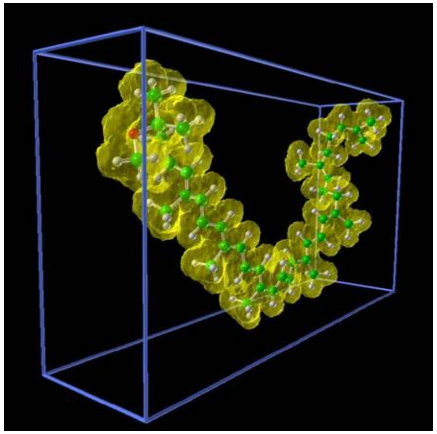
SciDAC - First Federal Program to Implement CSE

- **SciDAC (Scientific Discovery through Advanced Computing) program created in 2001**
  - **About $50M annual funding**
  - **Berkeley (LBNL+UCB) largest recipient of SciDAC funding**

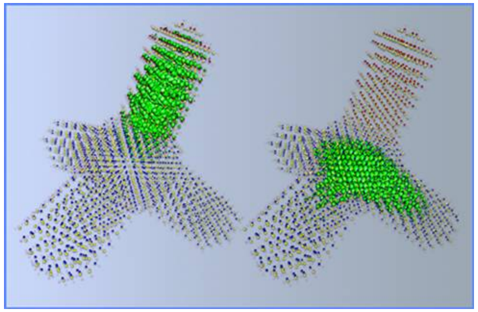
Global Climate



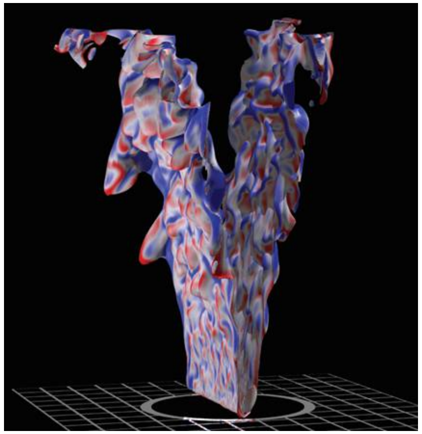
Biology



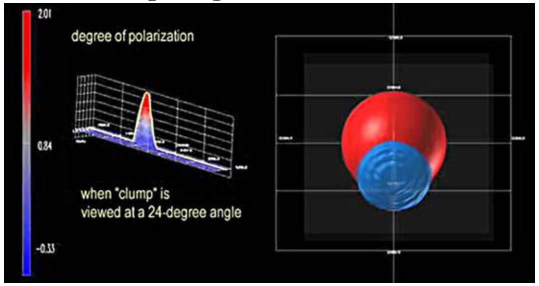
Nanoscience



Combustion



Astrophysics



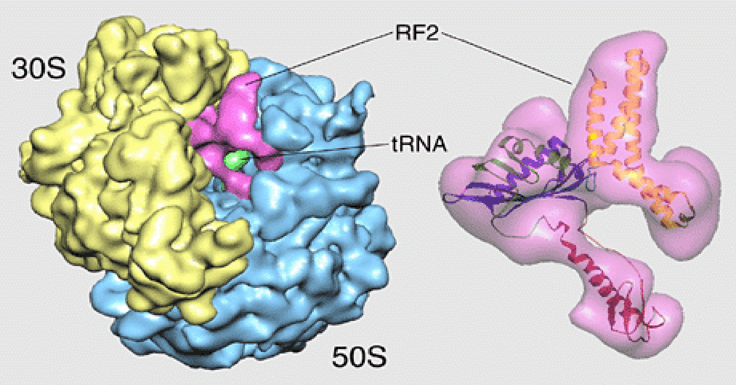
NERSC

Berkeley Lab

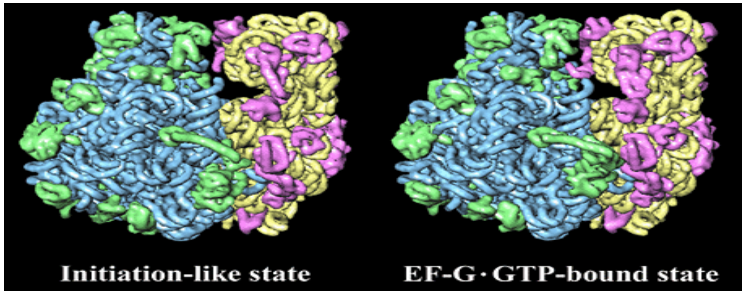
Office of Science
U.S. Department of Energy

Cryo-EM: Significance

- **Protein structure determination is one of the building blocks for molecular biology research**
  - **Provides the mapping of subunits and active sites within a complex**
- **Standard approach is to crystallize protein**
- **However, 30% of all proteins do not crystallize or are difficult to crystallize**



Ribosome bound with release factor RF2 in the presence of a stop codon and a P-site tRNA. (J. Frank)

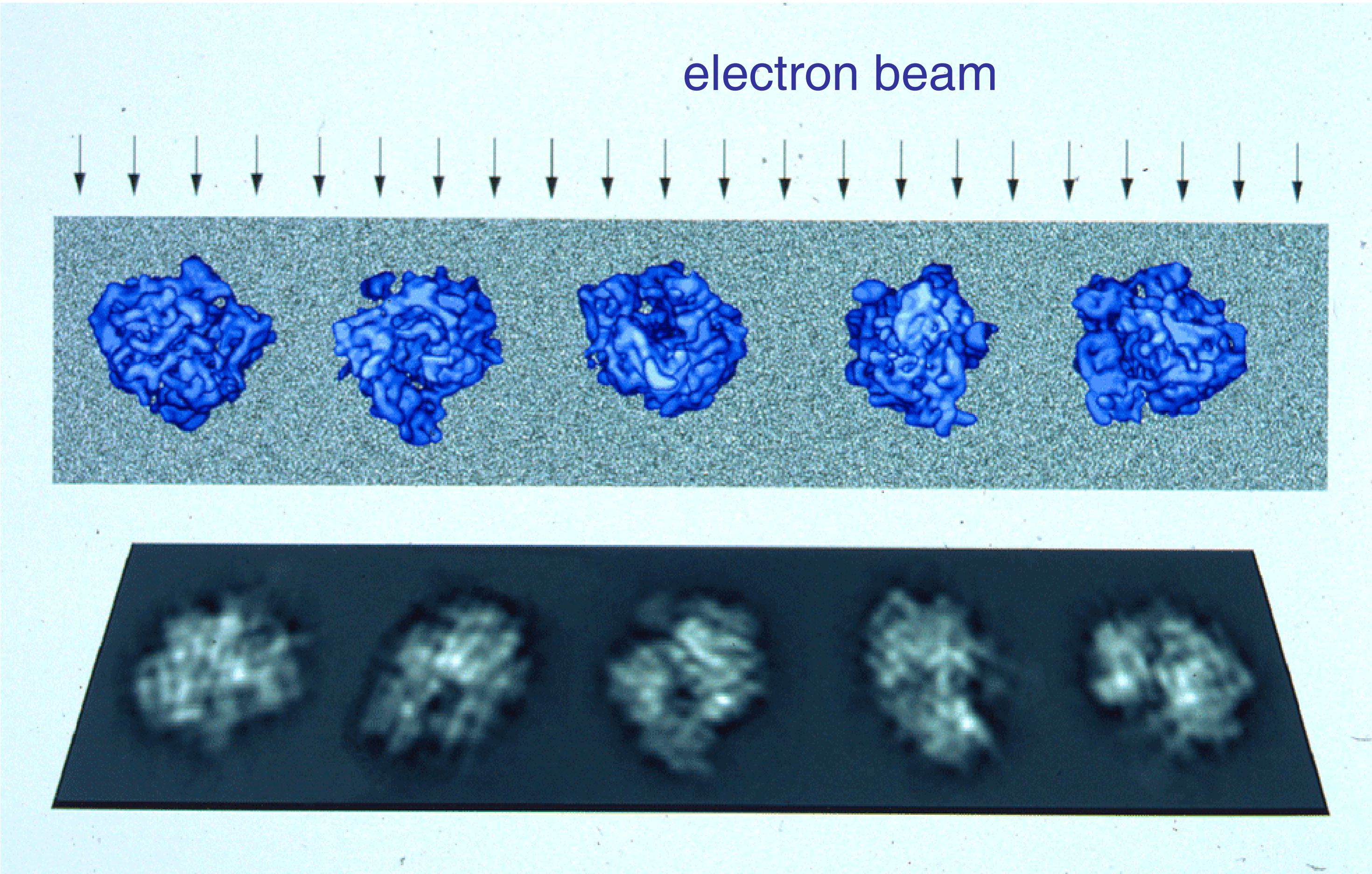


Space-filling atomic models of the *E. coli* ribosome in two conformations related by the ratchet rotation. Blue and green: RNA and proteins of 50S subunit, respectively; yellow and red: RNA and proteins of the 30S subunit. While the RNA undergoes continuous elastic deformations, some of the proteins rotate and change their conformations significantly. (J. Frank)

Cryo EM Team

- Funded by NIH in June 2003
  - Five-year program project
- Collaboration among many institutions
  - LBNL
    - ❑ CRD (Ng, Yang, Malladi)
    - ❑ Physical Biosciences (Glaser)
    - ❑ Life Sciences (Downing, Nogales)
  - U. Texas Medical School (Penczek)
  - Wadsworth Center, NY (Frank)
  - Baylor College of Medicine (Chiu)
  - Scripps Research Institute (Asturias)

The Reconstruction Problem



3D macromolecule

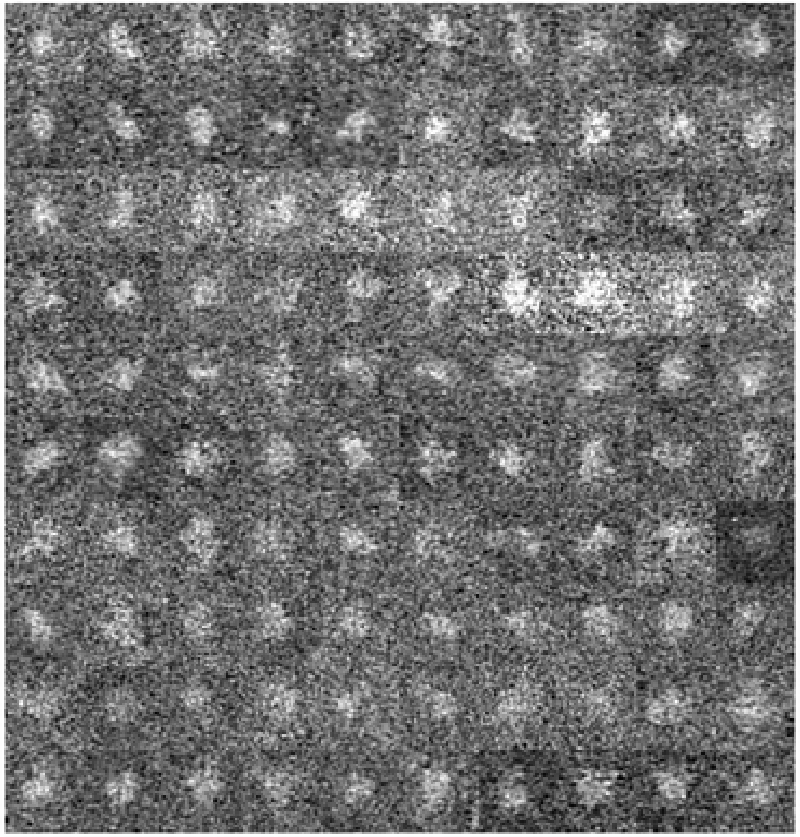
2D projections

(Cryo-EM)

Can we deduce the 3-D structure of the molecule from a set of 2-D projection images with unknown relative orientations?

Challenge

- Nonlinear inverse problem
- Extremely noisy data
- Large volume of data
  - To make the problem well-defined (over-determined)
  - To achieve sufficient signal-to-noise ratio (SNR)
    - Higher SNR yields higher resolution
    - To reach atomic resolution requires 10^6 2-D images



Mathematical Formulation

- Data: $g_i \in \Re^{n^2}, i = 1,2,\ldots m;$
- Unknown parameters:
  - Density: $f \in \Re^{n^3}$
  - Rotations: $(\varphi_i, \theta_i, \phi_i), i = 1,2,\ldots, m$
  - Translations: $(s_{x_i}, s_{y_i}), i = 1,2,\ldots m;$
- Objective

$$\min_{\varphi_i, \theta_i, \phi_i, f, s_{x_i}, s_{y_i}} \sum_{i=1}^{m} \left\| r_i \right\|^2 = \sum_{i=1}^{m} \left\| P\left(\varphi_i, \theta_i, \phi_i, s_{x_i}, s_{y_i}\right) f - g_i \right\|^2$$

“Unified 3-D Structural and Projection Orientation Refinement Using Quasi-Newton Algorithm.” C. Yang, E. Ng and P. A. Penczek. Journal of Structural Biology **149** (2005), pp. 53–64.

Computing the Search Direction

- Objective function $\rho(x) = \dfrac{1}{2}\sum_{i=1}^{m} \left\| r_i \right\|^2$

$$x^T = \left(f \;\; \phi_1 \cdots \phi_m \;\; \theta_1 \cdots \theta_m \;\; \psi_1 \cdots \psi_m \right)$$

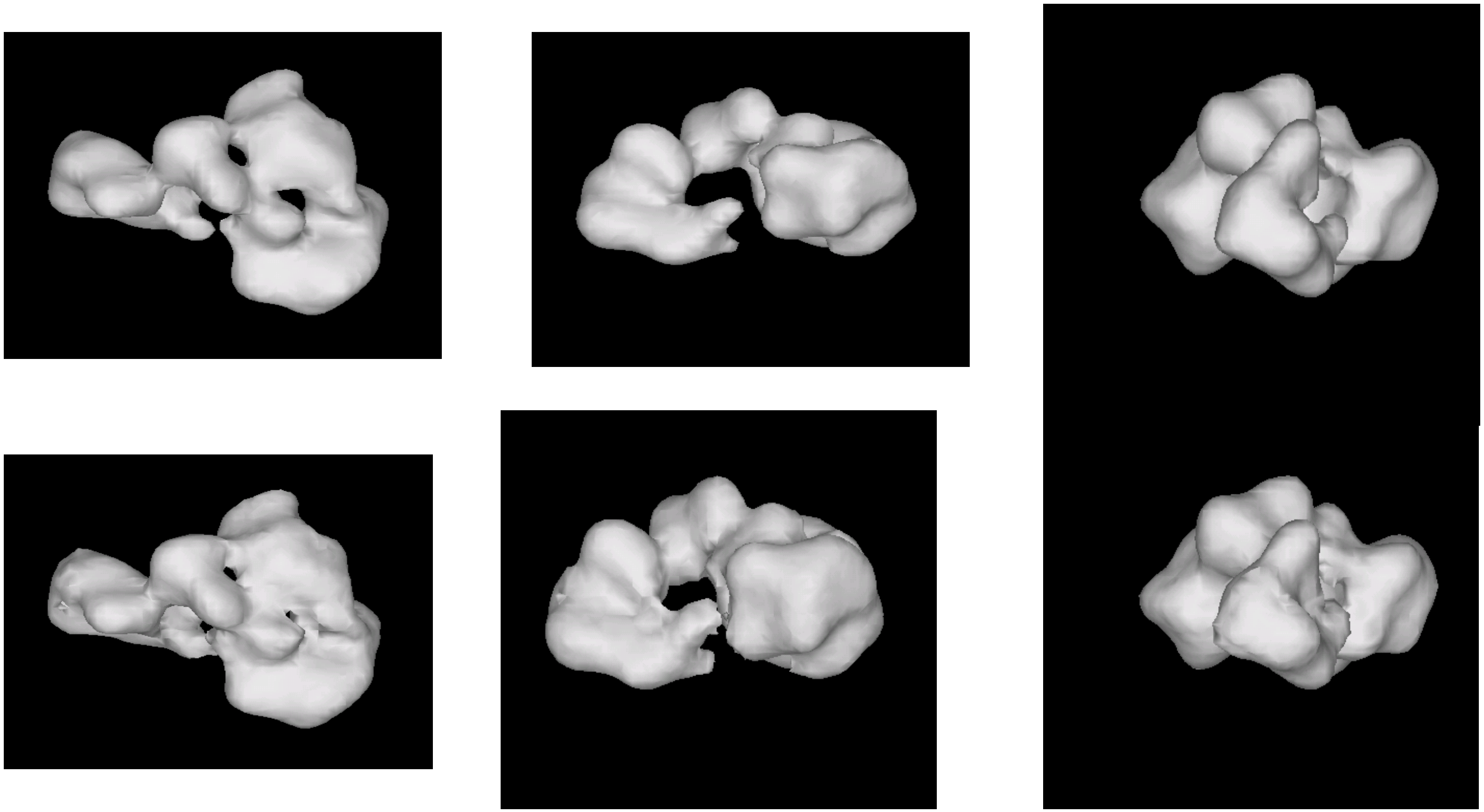
- Gradient $\nabla\rho(x) = J^T r$

$$J = \begin{pmatrix} P_1 & g_1^{\phi} & & & g_1^{\theta} & & & g_1^{\psi} & & \\ P_2 & & g_2^{\phi} & & & g_2^{\theta} & & & g_2^{\psi} & \\ \vdots & & & \ddots & & & \ddots & & & \ddots \\ P_m & & & g_m^{\phi} & & & g_m^{\theta} & & & g_m^{\psi} \end{pmatrix} \qquad r = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{pmatrix}$$

$$g^{\phi_1} = \frac{\partial r_1}{\partial \phi_1} \approx \frac{P(\phi_1 + \Delta\phi) f - P(\phi_1) f}{\Delta\phi}$$

J is mn^2 by n^3+3m
m can be as large as 10^6
n can be as large as 512

Exact vs. Reconstructed Volume

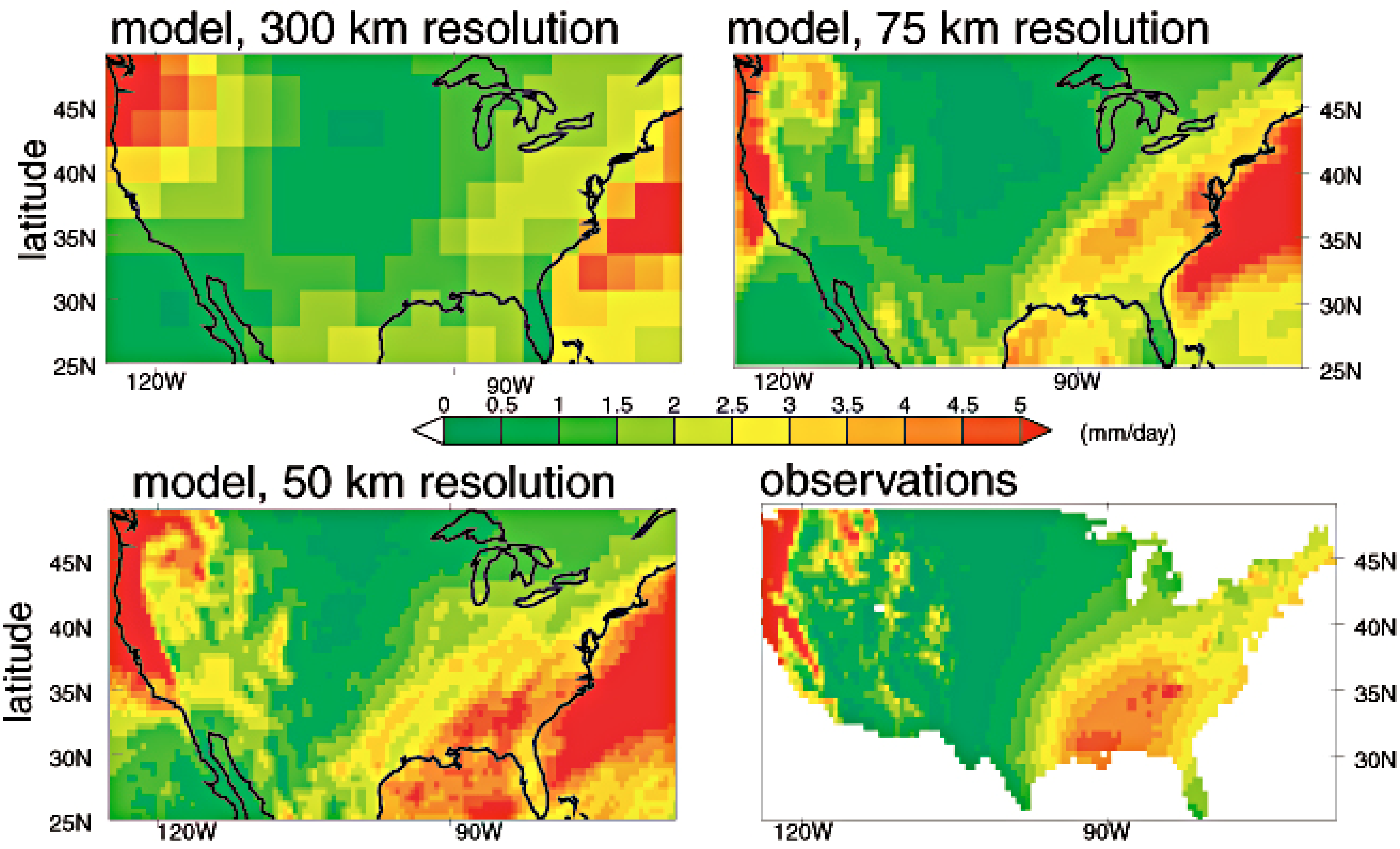


Cryo-EM - Summary

- The computer IS the microscope!
- Image resolution is directly correlated to the available compute power
- Naïve and complete ab initio calculation of a protein structure might require 10**(18) operations

Wintertime Precipitation

As model resolution becomes finer, results converge towards observations



Tropical Cyclones and Hurricanes

Research by: Michael Wehner, Berkeley Lab,
Ben Santer, Phil Duffy, and G. Bala, LLNL

- Hurricanes are extreme events with large impacts on human and natural systems
- Characterized by high vorticity (winds), very low pressure centers, and upper air temperature warm anomalies
- Wind speeds on the Saffir-Simpson Hurricane Scale
  - Category one: 74-95 mph (64-82 kt or 119-153 km/hr)
  - Category two: 96-110 mph (83-95 kt or 154-177 km/hr)
  - Category three: 111-130 mph (96-113 kt or 178-209 km/hr)
  - Category four: 131-155 mph (114-135 kt or 210-249 km/hr)
  - Category five: >155 mph (135 kt or 249 km/hr).

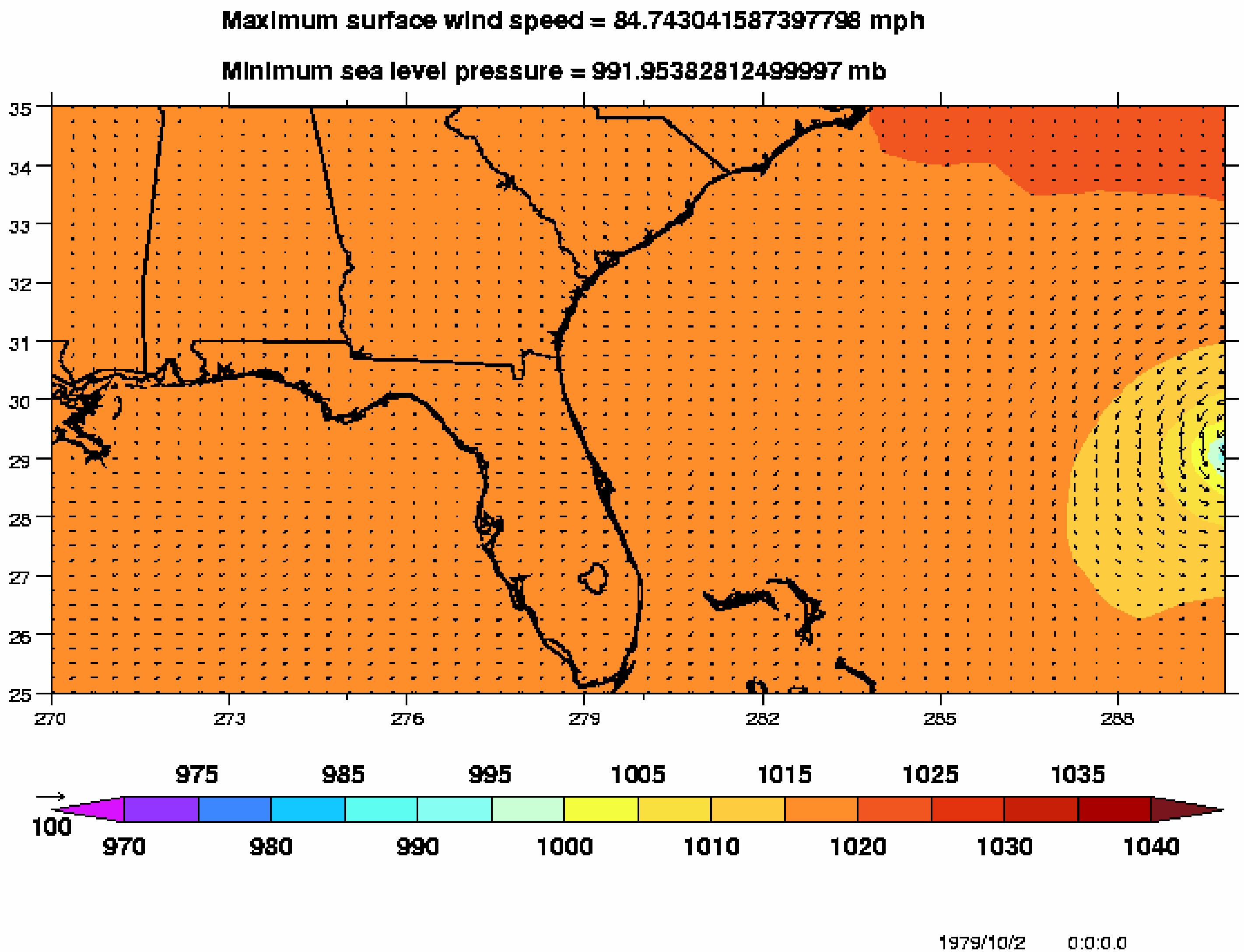
How will the hurricane cycle change as the mean climate changes?

Tropical Cyclones in Climate Models

- Tropical cyclones are not generally seen in integrations of global atmospheric general circulation models at climate model resolutions (T42 ~ 300 km).
- In fact, in CCM3 at T239 (50 km), the lowest pressure attained is 995 mb. No realistic cyclones are simulated.
- However, in high resolution simulations of the finite volume dynamics version of CAM2, strong tropical cyclones are common.

Finite Volume Dynamics CAM

- Run in an ‘AMIP’ Mode
  - Specified sea surface temperature and sea ice extent
  - Integrated from 1979 to 2000
- We are studying four resolutions
  - B: 2^{o}x2.5^{o}
  - C: 1^{o}x1.25^{o}
  - D: 0.5^{o}x0.625^{o}
  - E: 0.25^{o}x0.375^{o}
- Processor Configuration and Cost (IBM SP3)
  - B: 64 processors, 10 wall clock hours / simulated year
  - C: 160 processors, 22 wall clock hours / simulated year
  - D: 640 processors, 33 wall clock hours / simulated year
  - E: 640 processors, 135 wall clock hours / simulated year



Maximum surface wind speed = 84.743041587397798 mph
Minimum sea level pressure = 991.95382812499997 mb
35
34
33
32
31
30
29
28
27
26
25
270
273
276
279
282
285
288
975
985
995
1005
1015
1025
1035
100
970
980
990
1000
1010
1020
1030
1040
1979/10/2 0:0:0.0

NERSC
BERKELEY LAB
Office of Science
U.S. DEPARTMENT OF ENERGY

New Science Question: Hurricane Statistics

What is the effect of different climate scenarios on number and severity of tropical storms?

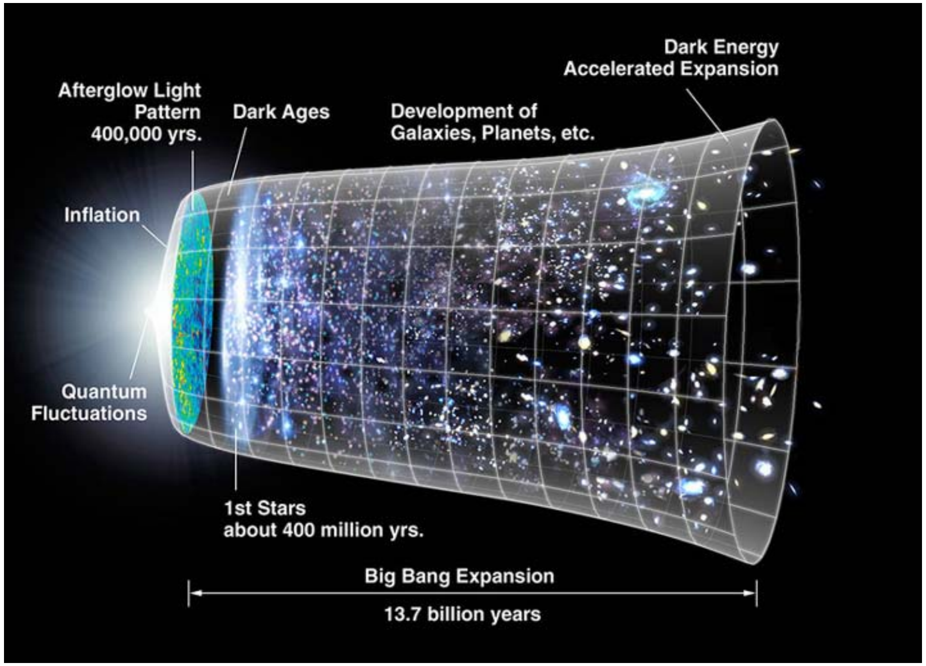
	1979	1980	1981	1982	Obs
Northwest Pacific Basin	>25	~30			40
Atlantic Basin	~6	~12			?

Work in progress—computer power insufficient!

Extreme Weather - Summary

- Computer Simulation permits us to perform experiments that are too dangerous
- We can ask new scientific questions that we could not even think of before
- Current computer power still insufficient to get statistically meaningful results on possible correlation of extreme weather and climate change

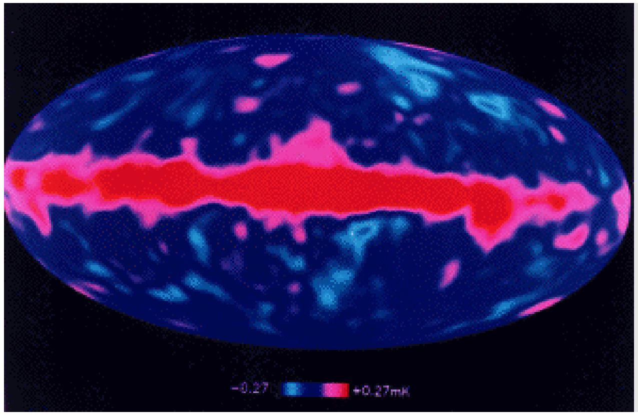
NERSC User George Smoot wins 2006 Nobel Prize in Physics



Cosmic Microwave Background Radiation (CMB): an image of the universe at 400,000 years

Smoot and Mather 1992

COBE Experiment showed anisotropy of CMB



CMB Computing at NERSC

- CMB data analysis presents a significant and growing computational challenge, requiring
  - well-controlled approximate algorithms
  - efficient massively parallel implementations
  - long-term access to the best HPC resources
- DOE/NERSC has become the leading HPC facility in the world for CMB data analysis
  - O(1,000,000) CPU-hours/year
  - O(10) Tb project disk space
  - O(10) experiments & O(100) users (rolling)

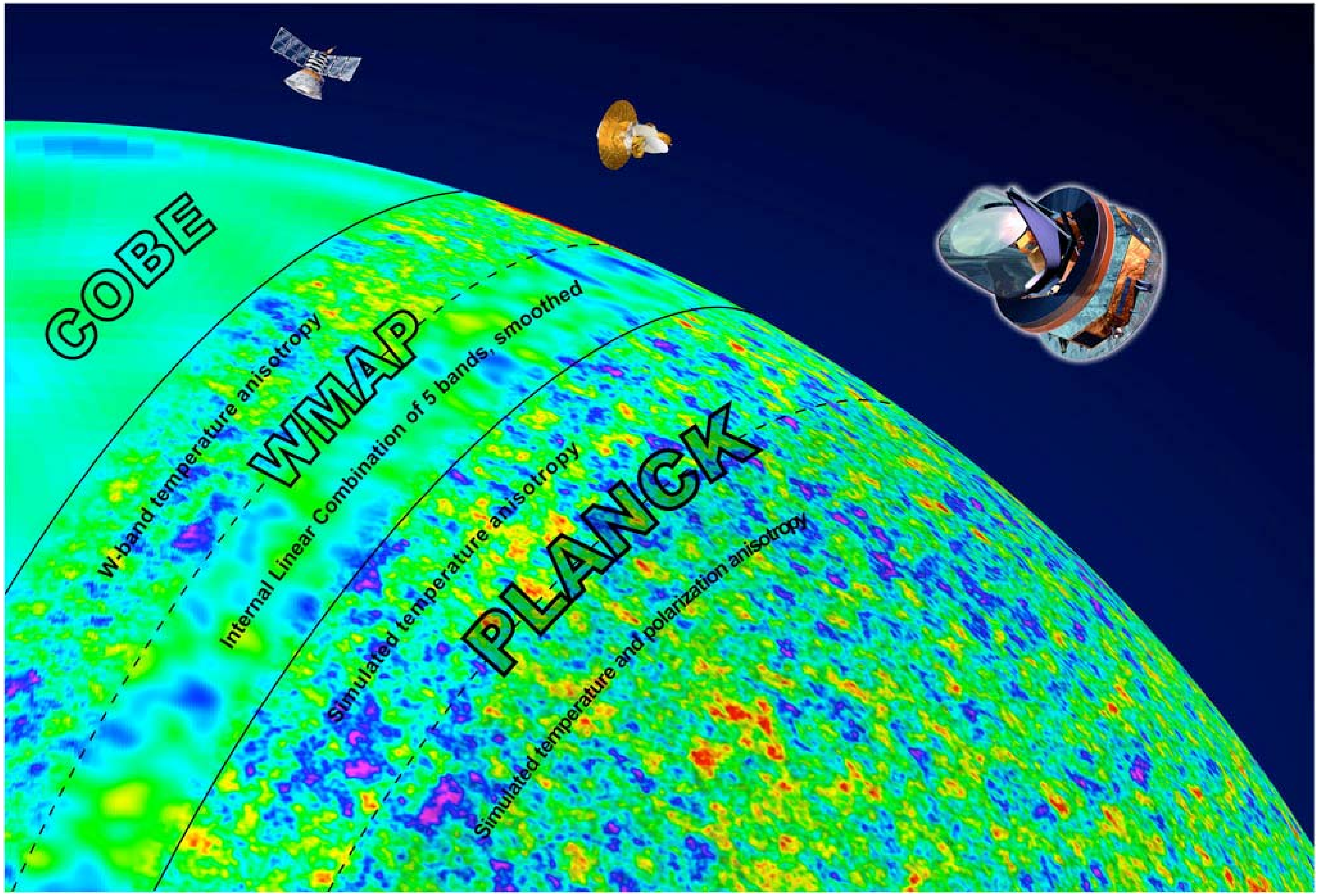
source J. Borrill, LBNL

Evolution Of CMB Data Sets

Experiment	N_t	N_p	N_b	Limiting Data	Notes
COBE (1989)	2×10^9	6×10^3	3×10^1	Time	Satellite, Workstation
BOOMERanG (1998)	3×10^8	5×10^5	3×10^1	Pixel	Balloon, 1st HPC/NERSC
(4yr) WMAP (2001)	7×10^{10}	4×10^7	1×10^3	?	Satellite, Analysis-bound
Planck (2007)	5×10^{11}	6×10^8	6×10^3	Time/ Pixel	Satellite, Major HPC/DA effort
POLARBEAR (2007)	8×10^{12}	6×10^6	1×10^3	Time	Ground, NG-multiplexing
CMBPol (~2020)	10^{14}	10^9	10^4	Time/ Pixel	Satellite, Early planning/design

data compression

Evolution Of CMB Satellite Maps



Algorithms & Flop-Scaling

- Map-making
  - Exact maximum likelihood : $O(N_p^3)$
  - PCG maximum likelihood : $O(N_i N_t \log N_t)$
  - Scan-specific, e.g.. destriping : $O(N_t \log N_t)$
  - Naïve : $O(N_t)$

Speed

Accuracy

- Power Spectrum estimation
  - Iterative maximum likelihood : $O(N_i N_b N_p^3)$
  - Monte Carlo pseudo-spectral :
    - Time domain : $O(N_r N_i N_t \log N_t)$, $O(N_r l_{max}^3)$
    - Pixel domain : $O(N_r N_t)$
    - Simulations
      - exact simulation > approximate analysis !

Speed

Accuracy

NERSC BERKELEY LAB Office of Science U.S. DEPARTMENT OF ENERGY

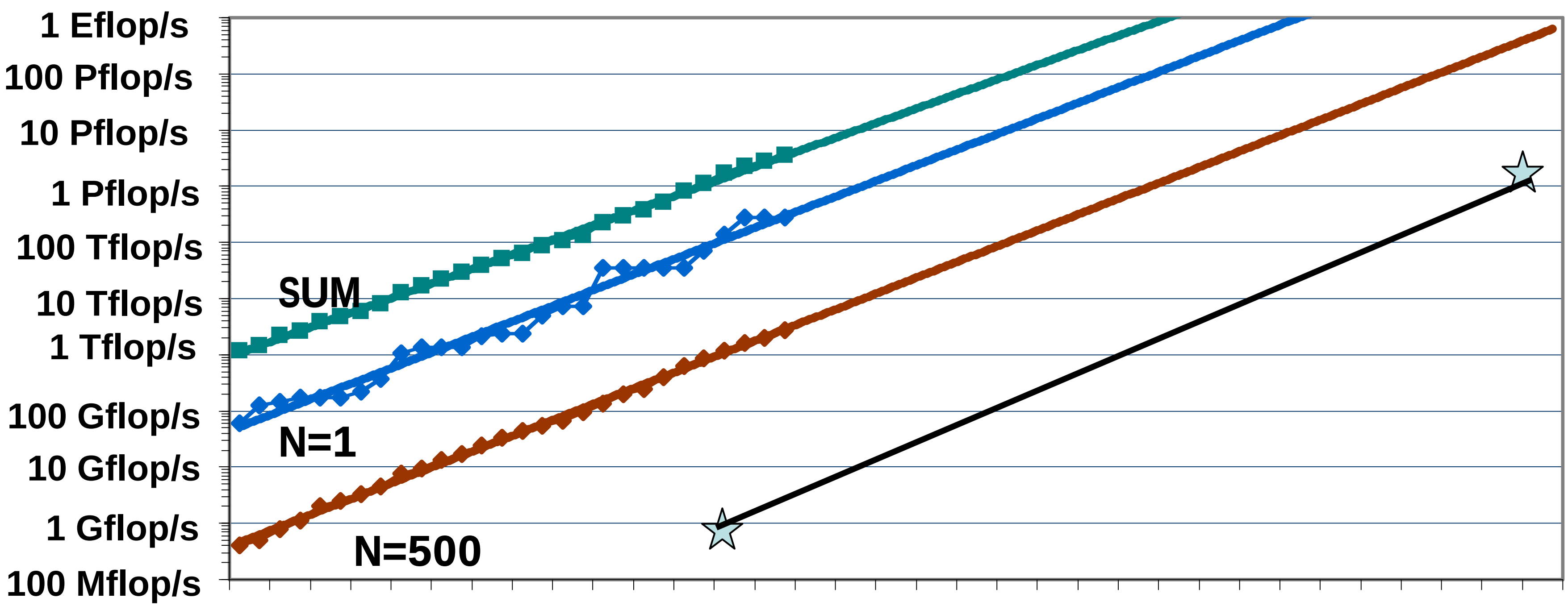
CMB is Characteristic for CSE Projects

- **Petaflop/s and beyond computing requirements**
- **Algorithm and software requirements**
- **Use of new technology, e.g. NGF**
- **Service to a large international community**
- **Exciting science**

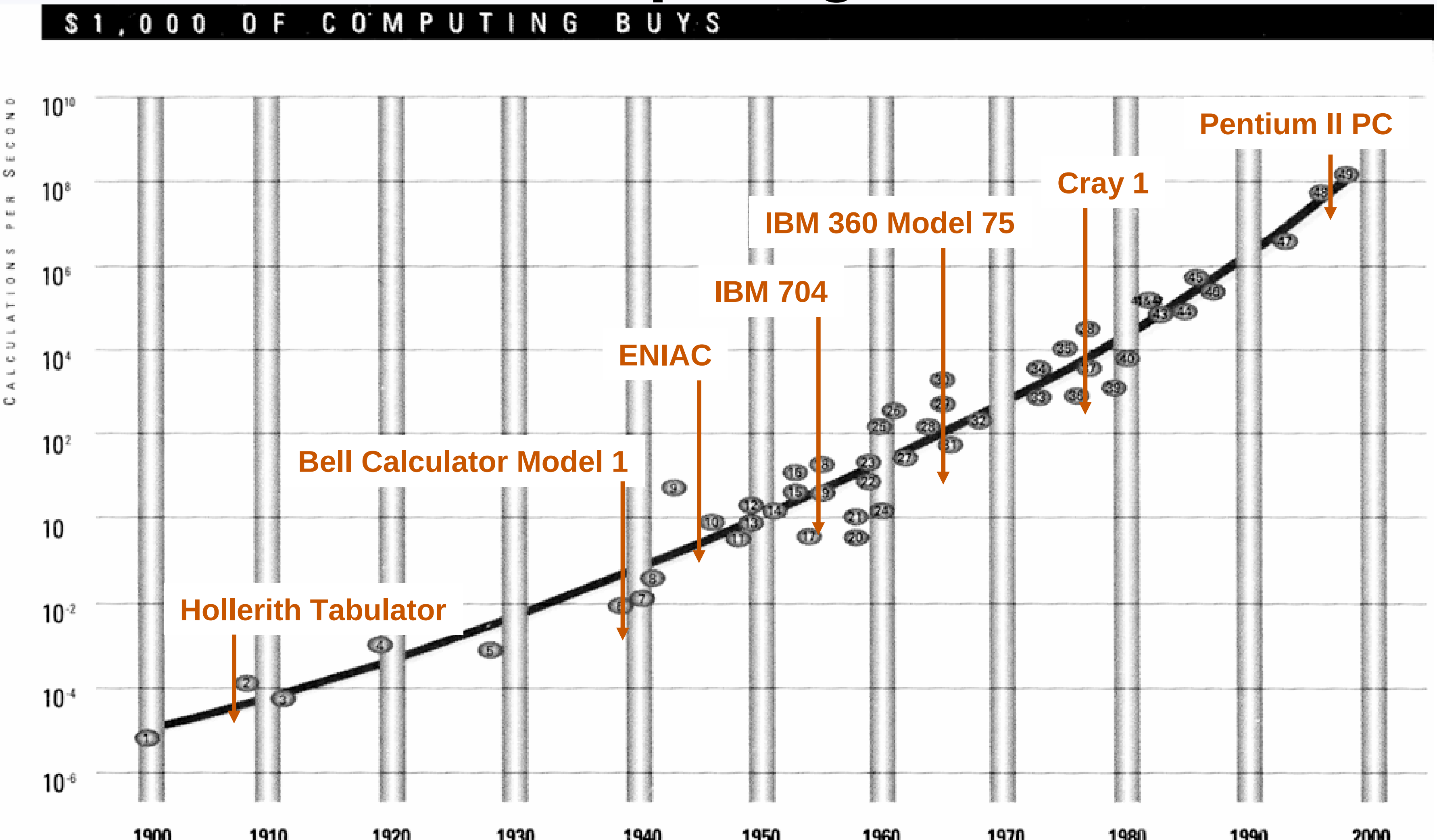
Overview

- Introducing NERSC and Computing Sciences at Berkeley Lab
- Current Trends in Supercomputing (High-End Computing)
- What Supercomputers Do
- What Supercomputers Still Can’t Do

Performance Projection (Top500)



The Exponential Growth of Computing, 1900-1998



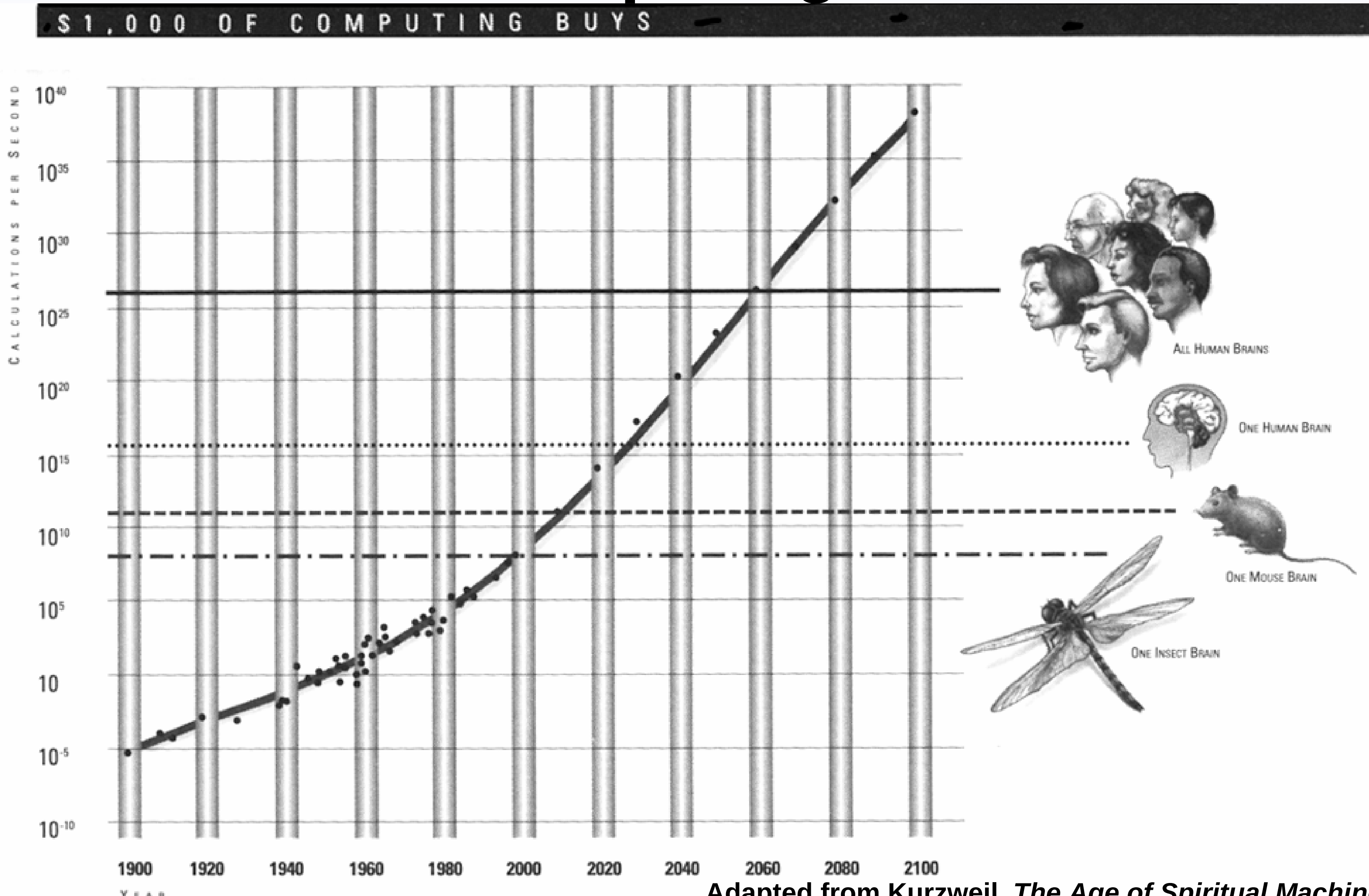
Adapted from Kurzweil, *The Age of Spiritual Machines*

NERSC

BERKELEY LAB

Office of Science
U.S. DEPARTMENT OF ENERGY

The Exponential Growth of Computing, 1900-2100



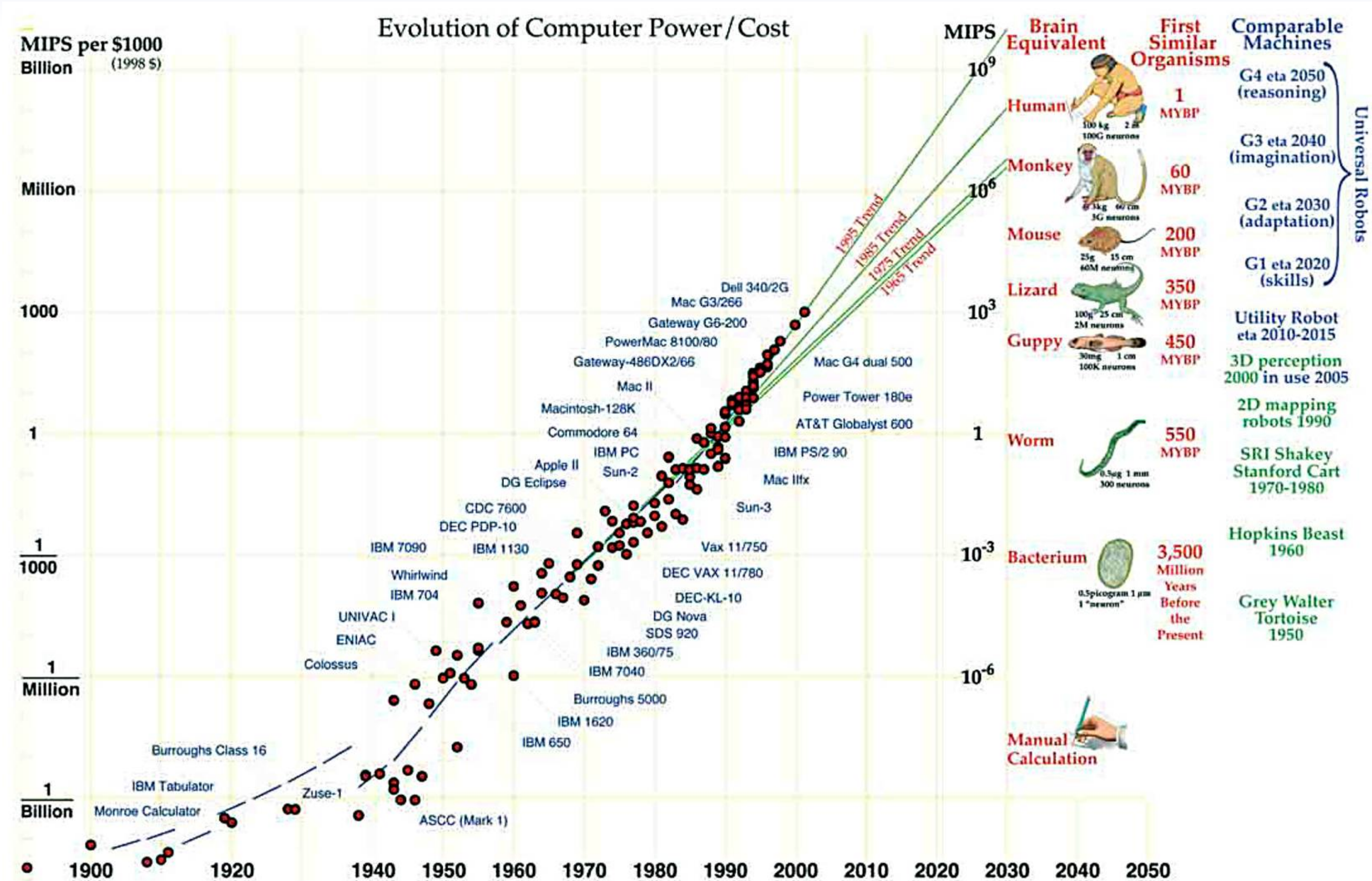
Adapted from Kurzweil, *The Age of Spiritual Machines*

NERSC

BERKELEY LAB

Office of Science
U.S. DEPARTMENT OF ENERGY

Growth of Computing Power and “Mental Power”



Hans Moravec, *CACM* 10, 2003, pp 90-97

NERSC

BERKELEY LAB

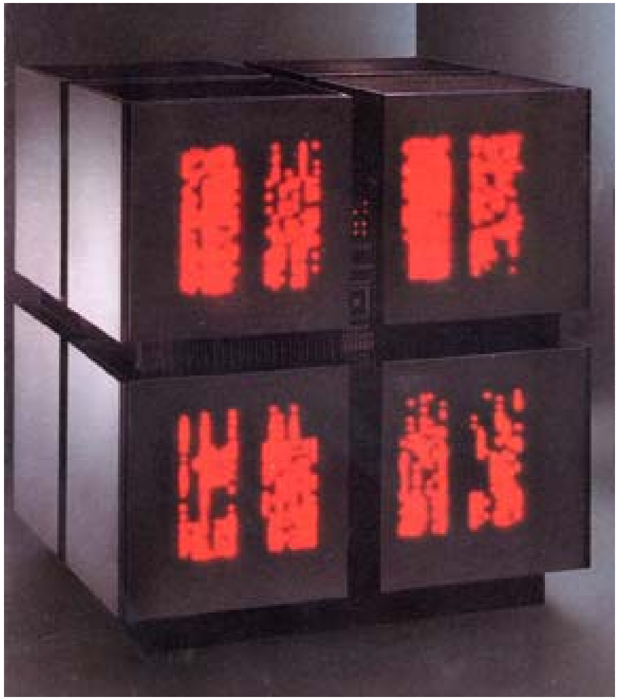
Office of Science
U.S. DEPARTMENT OF ENERGY

Why This Simplistic View is Wrong

- **Unsuitability of Current Architectures**
  - **Teraflop systems are focused on excelling in computing; only one of the six (or eight) dimensions of human intelligence**
- **Fundamental lack of mathematical models for cognitive processes**
  - **That's why we are not using the most powerful computers today for cognitive tasks**
- **Complexity limits**
  - **We don't even know yet how to model turbulence, how then do we model thought?**

History Lesson: 1987

- "Legendary" CM-2 by Thinking Machines
- Architecture evolved into CM-5 (1992) built as MPP for scientific applications
- Early history of AI applications on parallel platforms has been lost

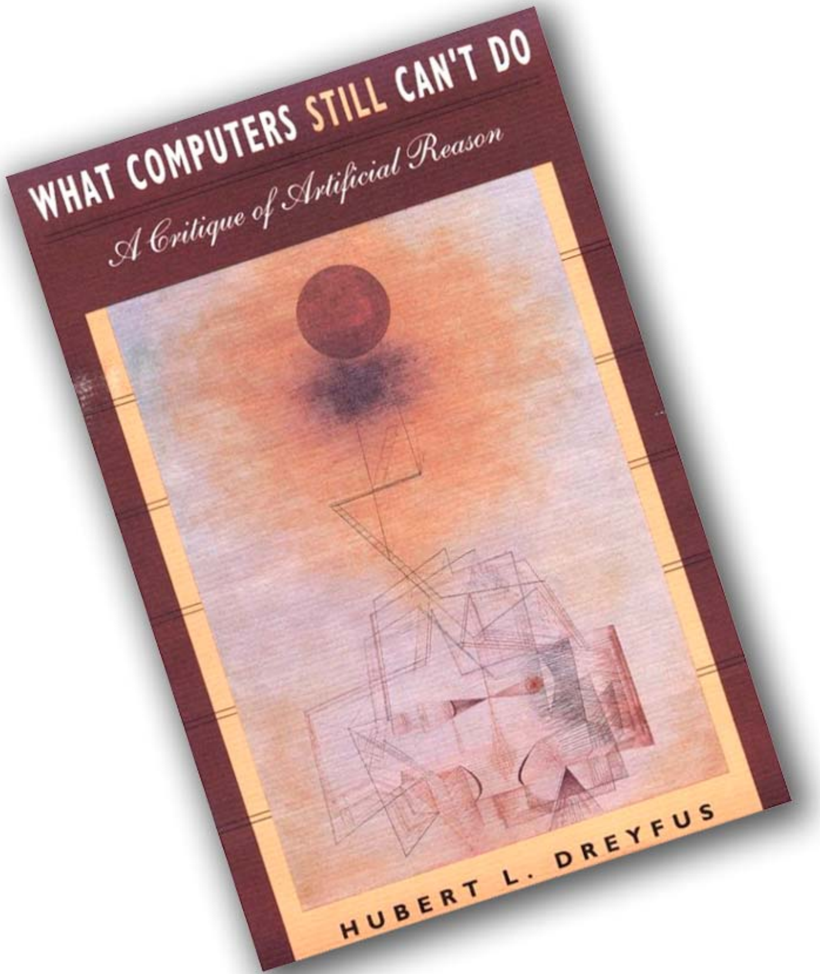


History Lesson: 1997

- IBM Deep Blue beats Gary Kasparov (May 1997)
- one of the biggest success stories of machine intelligence,
- however, the chess computer “Deep Blue”, did not teach us anything about how a chess grandmaster thinks
- no further analysis or further developments

Motivation for the Title of my Talk

"The computer model turns out not to be helpful in explaining what people actually do when they think and perceive."
Hubert Dreyfus, pg.189



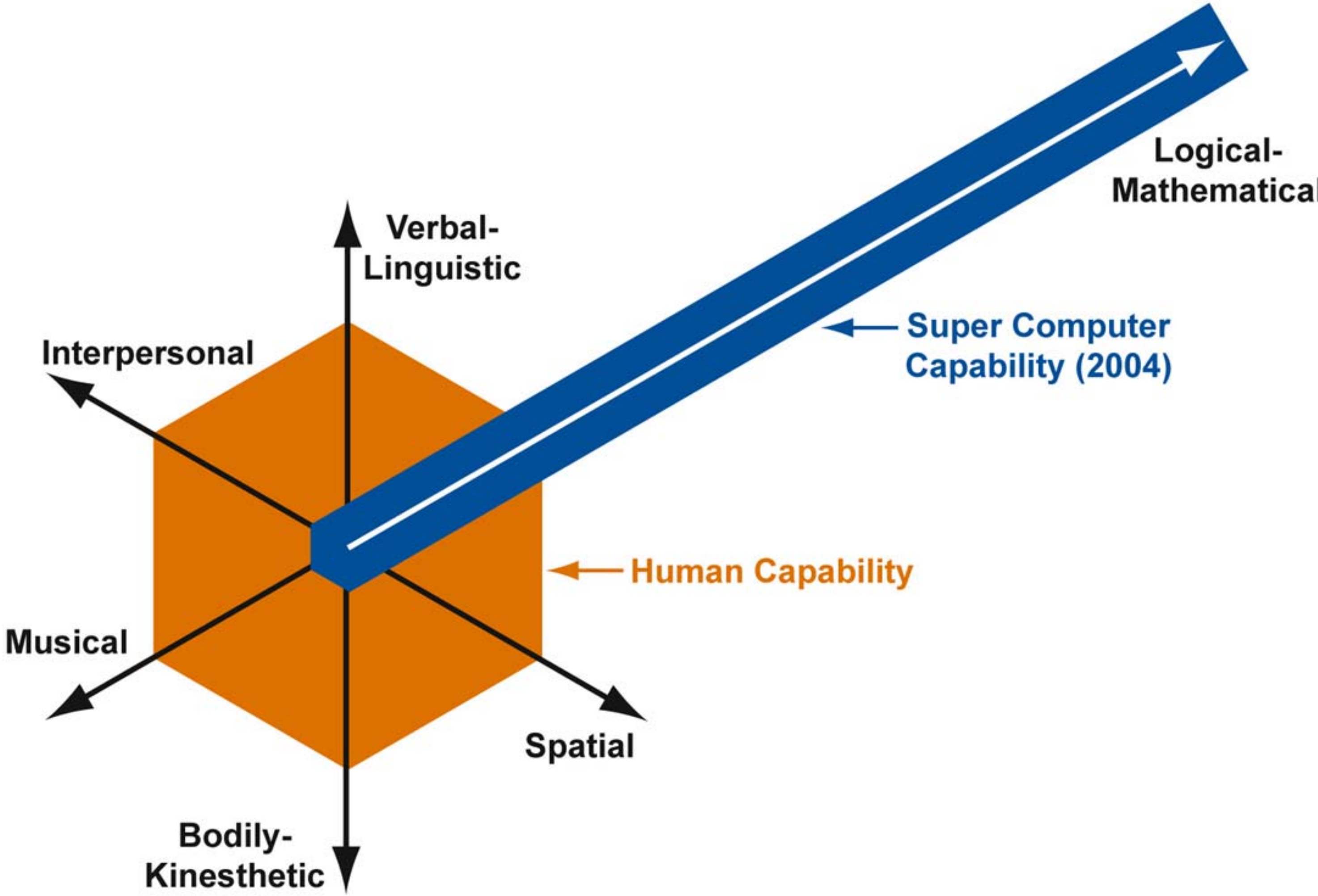
Example: one of the biggest success stories of machine intelligence, the chess computer "Deep Blue", did not teach us anything about how a chess grandmaster thinks.

Six Dimensions of Intelligence

1. **Verbal-Linguistic**
   ability to think in words and to use language to express and appreciate complex concepts
2. **Logical-Mathematical**
   makes it possible to calculate, quantify, consider propositions and hypotheses, and carry out complex mathematical operations
3. **Spatial**
   capacity to think and orientate in physical three-dimensional environment
4. **Bodily-Kinesthetic**
   ability to manipulate objects and fine-tune physical skills
5. **Musical**
   sensitivity to pitch, melody, rhythm, and tone
6. **Interpersonal**
   capacity to understand and interact effectively with others

Howard Gardner. *Frames of Mind: The Theory of Multiple Intelligences*. New York: Basic Books, 1983, 1993.

Current State of Supercomputers



The Research Opportunities for CSE

There are vast areas of science and engineering where CSE has not even begun to make an impact

- **current list of CSE applications is almost the same as fifteen years ago**
- **in many scientific areas there is still an almost complete absence of computational models**
- **even in established areas many researchers do not know how to use leading-edge computational resources**

Research opportunities for computer scientists and applied mathematicians

- **the current set of architectures is capturing only a small cognitive abilities subset of human**
- **our tools for analyzing vast amounts of data are still primitive**

Resources

Argonne Leadership Computing Facility

- Current Blue Gene/L Capabilities
  - BGL: 1024 nodes, 2048 cores, 5.7 TF, 512GB
  - Supports development + INCITE
- Additional 2007 INCITE time at IBM T.J. Watson Research
  - BGW: 20,480 nodes, 40,960 cores, 114 TF, 10 TB
- Coming for early 2008 production
  - 100+ TF Blue Gene system
  - Fast PB file system
  - Many PB tape archive
- Then for early 2009 production
  - Major Blue Gene upgrade
  - Next gen. file system

BGL

BGW

Contact: Paul Davé, dave@alcf.anl.gov

Accounts: http://accounts.alcf.anl.gov

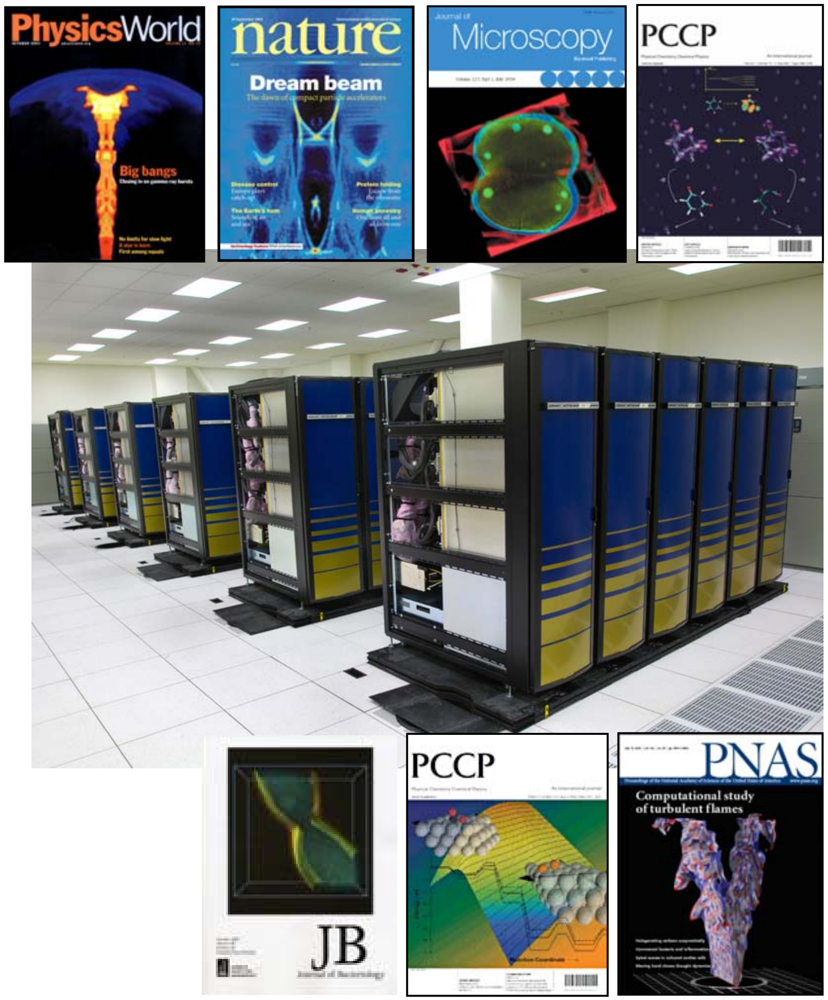
National Center for Computational Sciences
Oak Ridge National Laboratory

- **Leadership Computing Facility provides the most powerful open science systems in the world for government, academia, and industry**
- **Focuses on a small number of important applications that need the full size of these systems to solve time-sensitive problems**
- **Contact: Buddy Bland, blandas@ornl.gov**

Phoenix - Cray X1E: 1,024 Vector Processors

Jaguar – 119 TF Cray XT4: 11,706 Dual-Core Processors, 46 TB memory

NERSC at Berkeley Lab



For quick start-up allocations for CSGF projects contact Francesca Verdier, fverdier@lbl.gov

General accounts URL: http://www.nersc.gov/nusers/accounts/

Cray XT-4 “Franklin”

19,344 compute cores

102 Tflop/sec peak

39 TB memory

350 TB usable disk space

50 PB storage archive

INCITE Program

- INCITE program started in 2004 at NERSC to provide access to large supercomputer to researchers from government, academia, and industry
- Today includes four laboratories from DOE Office of Science: ANL, LBNL, ORNL, PNNL
- Call for 2008 proposals is available today
- DOE will allocate ~150 million CPU hours in 2008

Office of Science U.S. DEPARTMENT OF ENERGY
INCITE
Advancing America's Science and Industrial Competitiveness

See http://hpc.science.doe.gov

2007 INCITE Allocations:
45 projects, 95 million hrs

ANL

LBNL

NERSC

ORNL

Tell us about your computational needs

- There are many resources in the DOE HPC space available to meet the computational needs of your research. The SciDAC Outreach Center can help you find the right resources.
- Take the CSGF HPC Survey
  - http://outreach.scidac.gov/survey/csgf/
- Send us your computing inquiries:
  - Email : help@outreach.scidac.gov
  - Phone : 1-866-470-5547