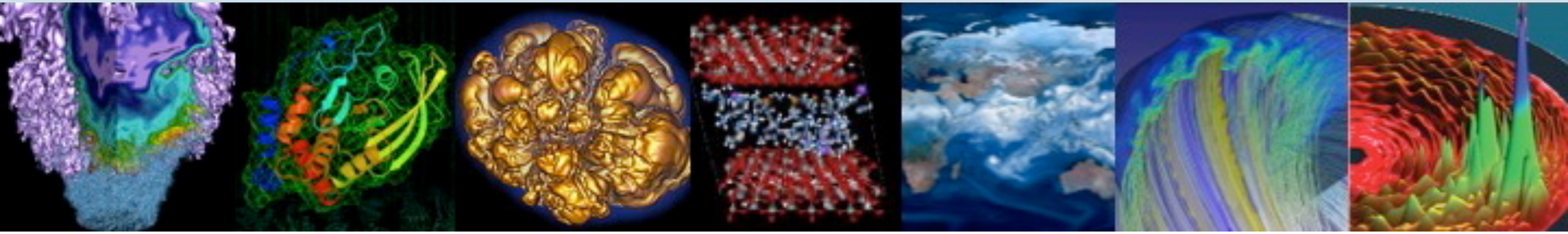


Saving the World with Computing

Kathy Yelick

EECS Professor, U.C. Berkeley

Associate Laboratory Director for Computing Sciences and NERSC Director, Lawrence Berkeley National Laboratory

BERKELEY LAB

Why are you Interested in Computer Science?

I want to:

A. Build computer hardware and software

B. Create new companies and industries

C. Solve important problems facing the world

D. Work on teams with other creative people

E. All of the above

Using Computers for Science and Engineering

Computers are used to understand things that are:

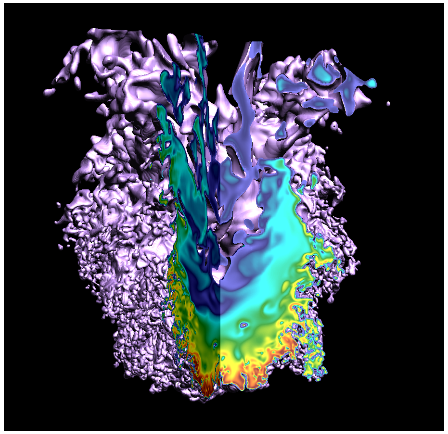
- too big
- too small
- too fast
- too slow
- too expensive or
- too dangerous

for experiments

Understanding the universe

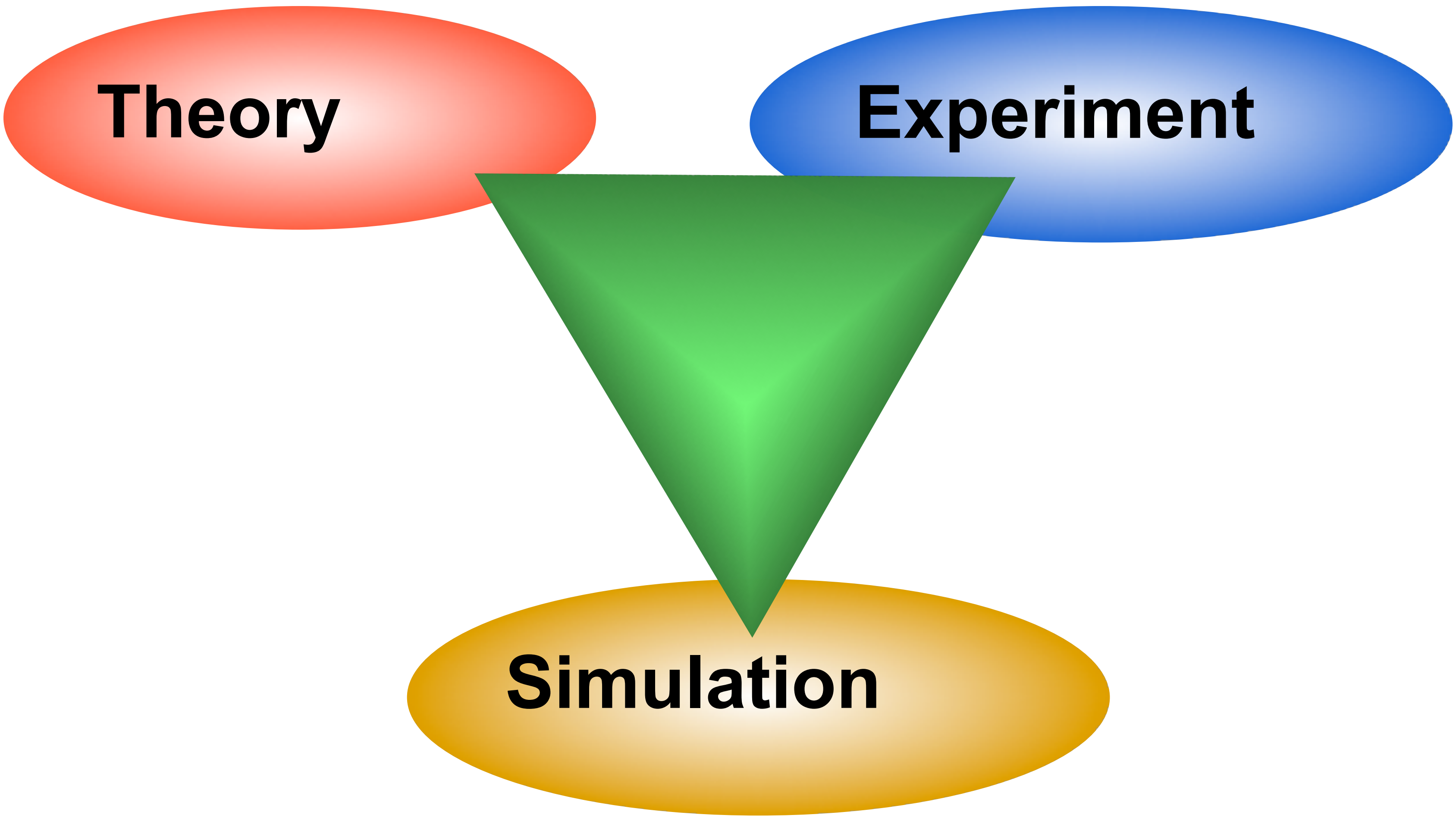
Proteins and diseases like Alzheimer's

Industrial products and processes



Energy-efficient combustion engines

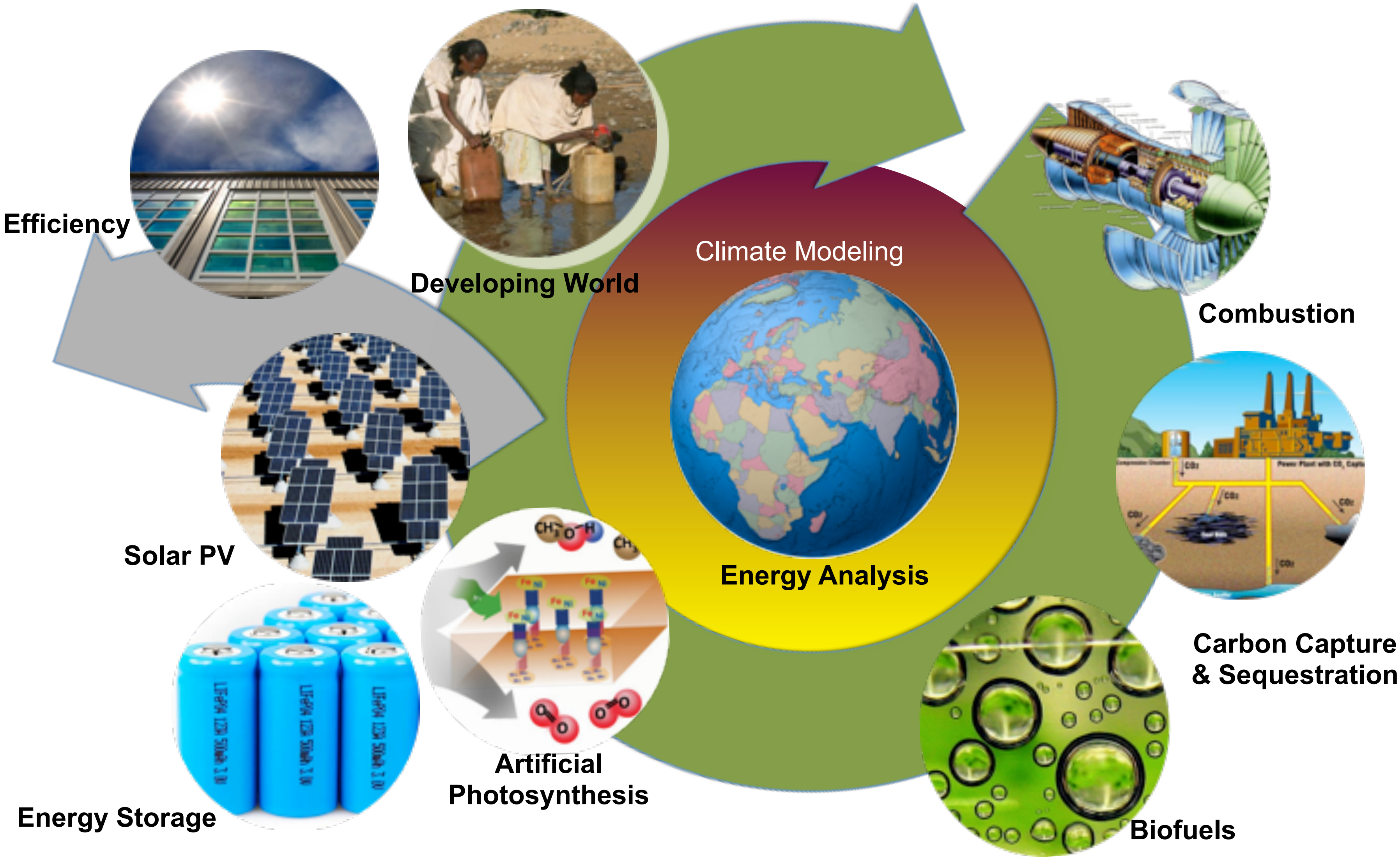
Simulation:
The "Third Pillar" of Science



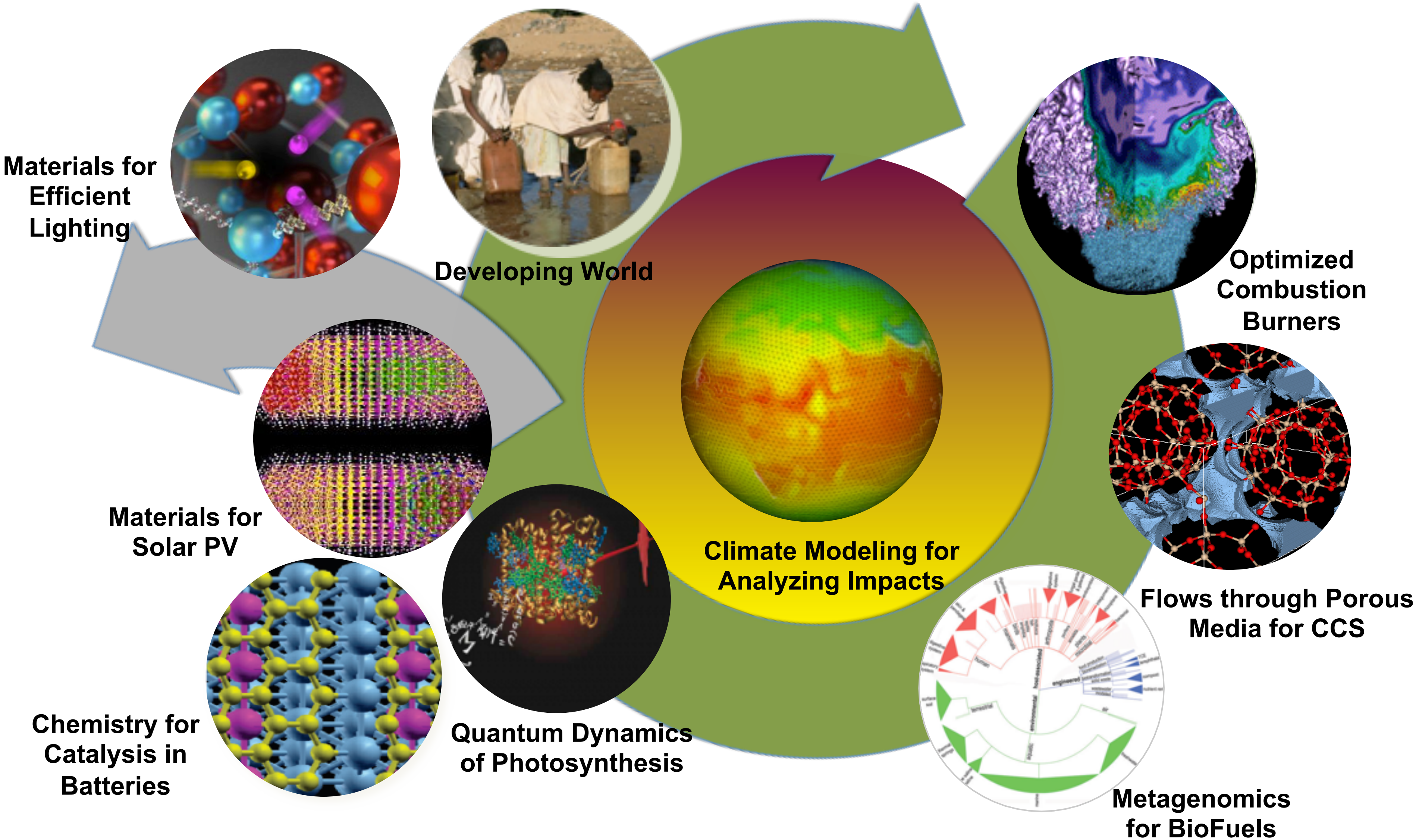
Addressing Challenges using Computing

- Two of the most significant challenges
  - Our changing world: understanding climate change, alternative energy sources, mitigation techniques, etc.
  - Health and medicine: understanding the human body, development of treatments, and disease prevention

Carbon Cycle 2.0 Initiative at Berkeley Lab

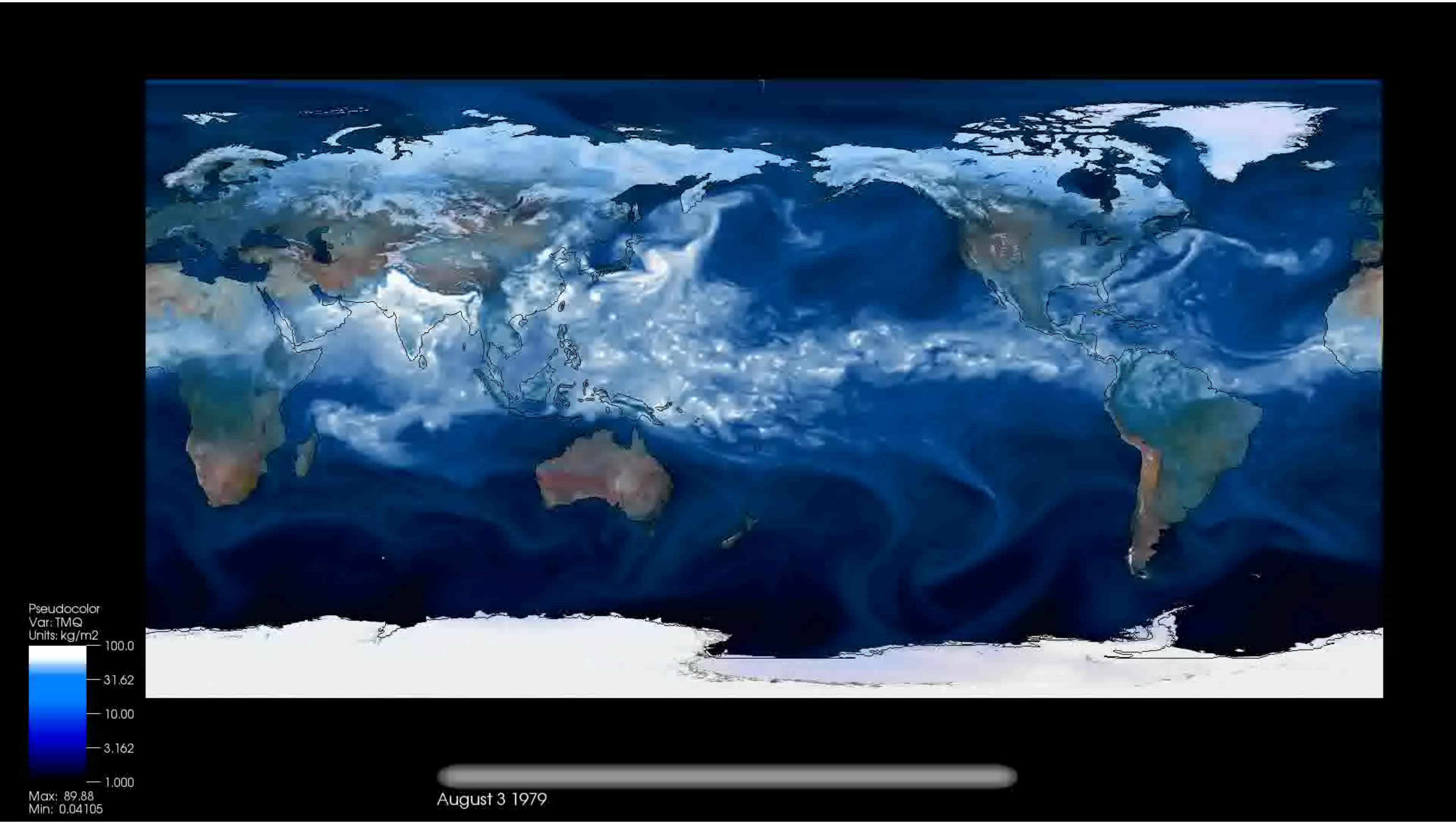


Computing for Carbon Cycle 2.0

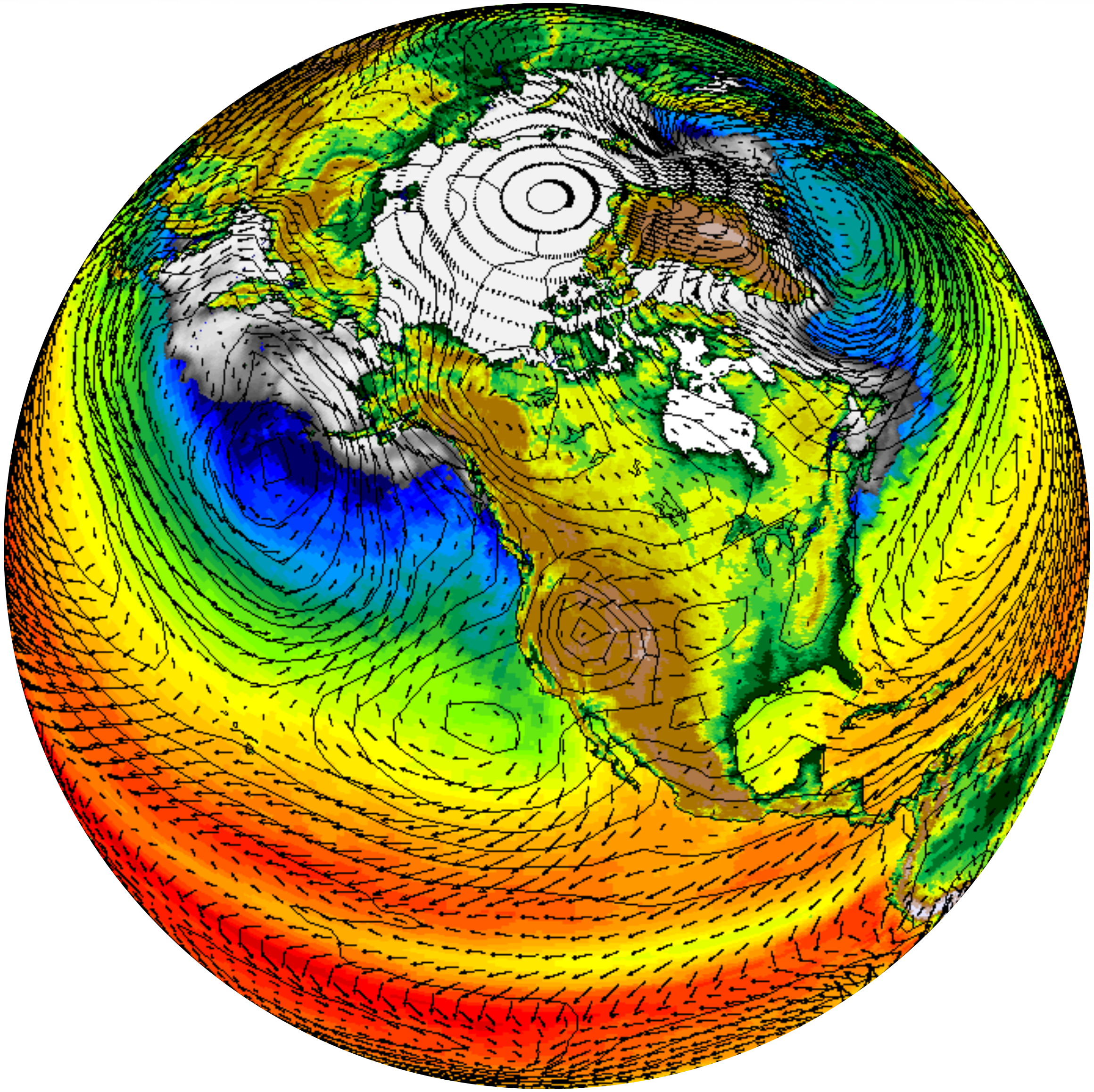


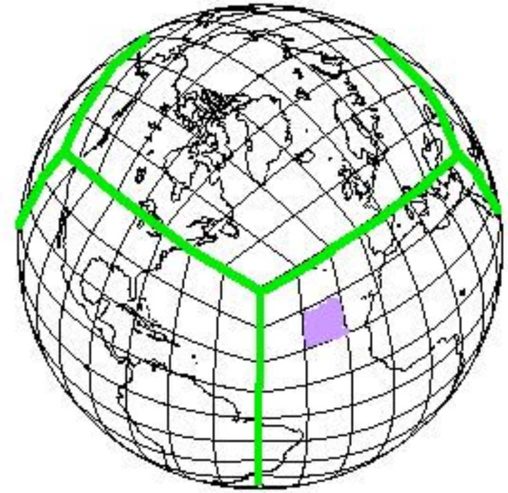
1979 Hurricane Season

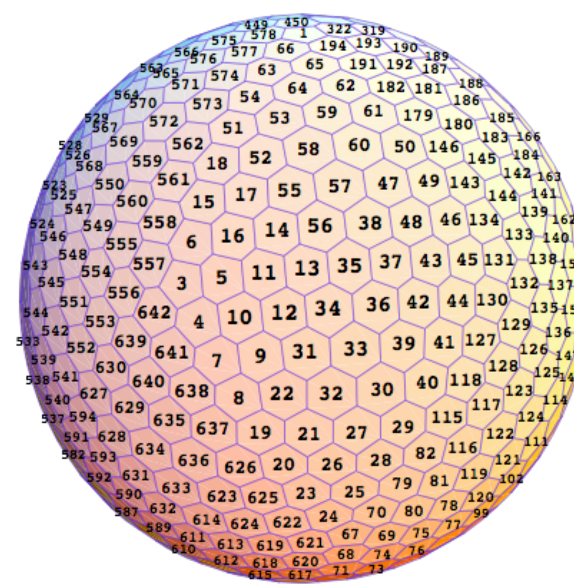
Movie from Michael Wehner and Prabhat at LBNL



Data Structures for Simulations



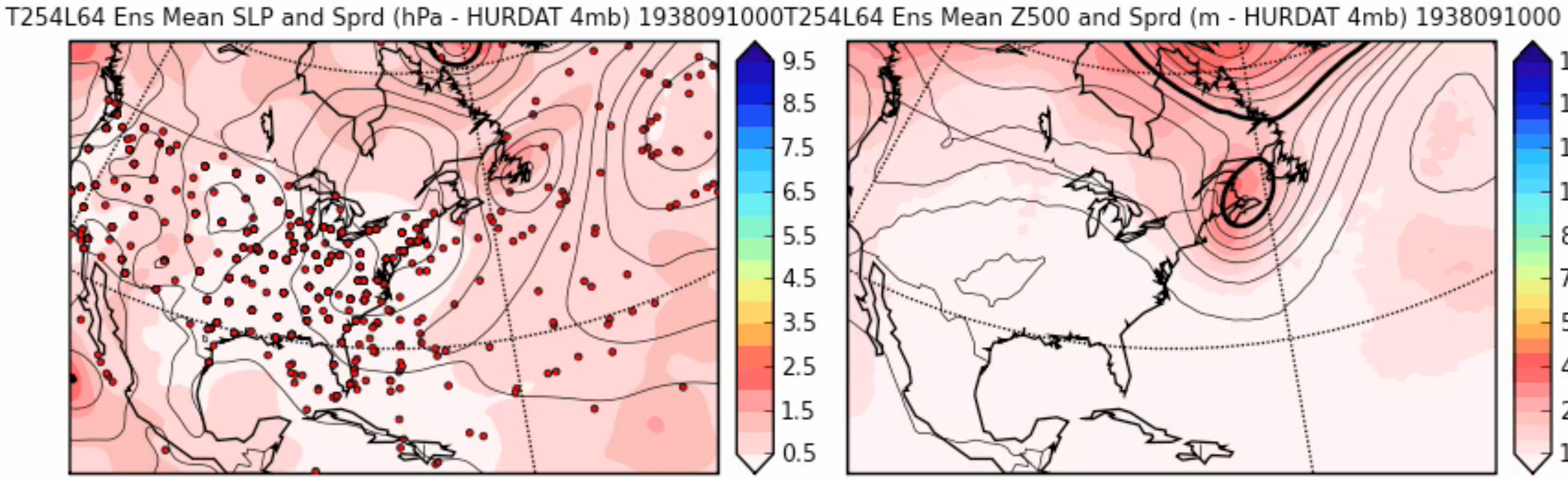




Climate Change Requires Lots of Data

"validate" that the computer models are working as expected

Simulation of 1938 hurricane hitting New York

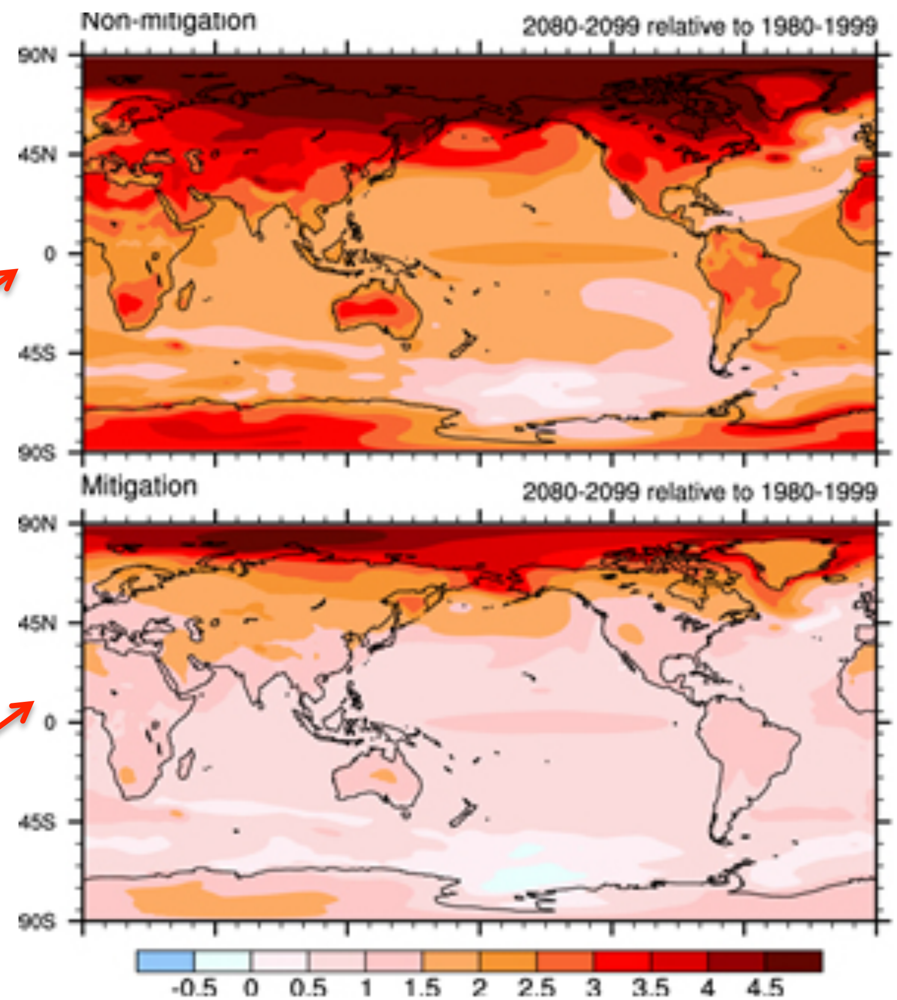


Mitigating Global Climate Change

Can global warming impacts be diminished if greenhouse gases are cut?

Average surface air temperatures rise by >3°C if emissions increase at current rate

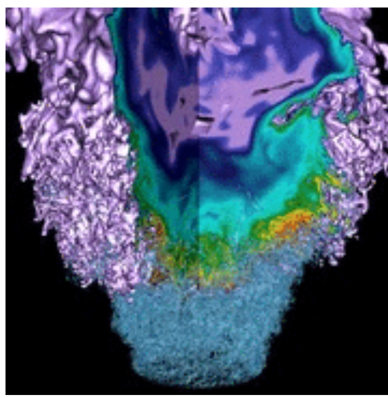
Temperatures rise by <2°C across nearly all populated areas if emissions are cut by 70%

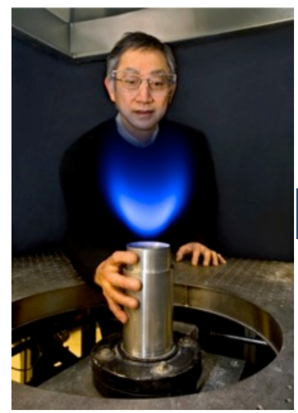


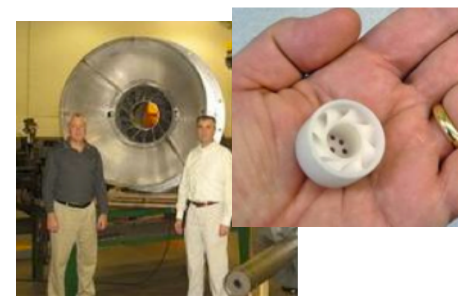
Simulations Aid in the Energy Efficient Devices

- Combustion simulations improve future designs
  - Model fluid flow, burning and chemistry
  - Uses advanced math algorithms
  - Petascale (10^{15} ops/sec) systems today

Simulations reveal features not visible in lab experiments



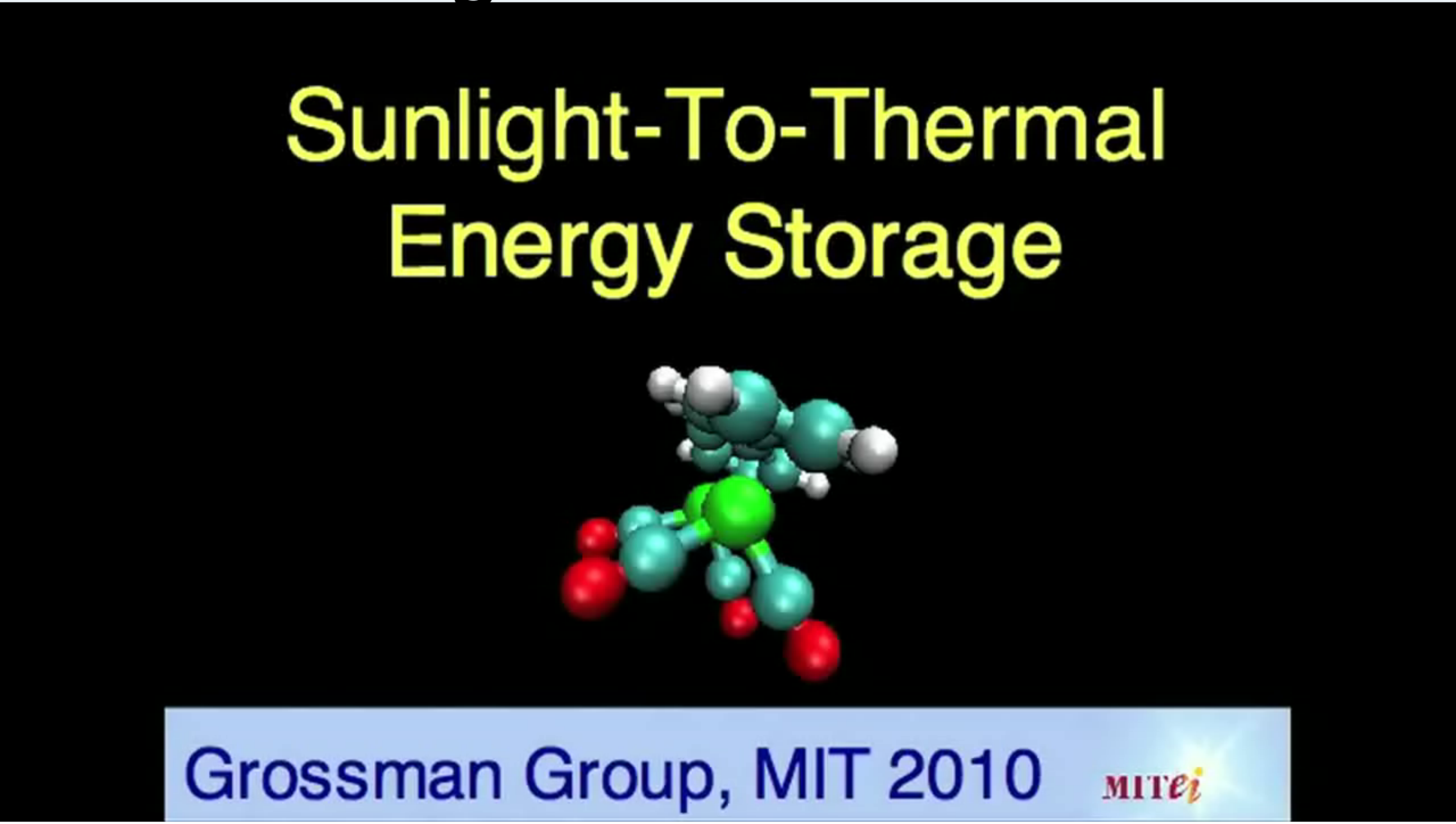




Energy efficient, low emissions technology licensed by industry

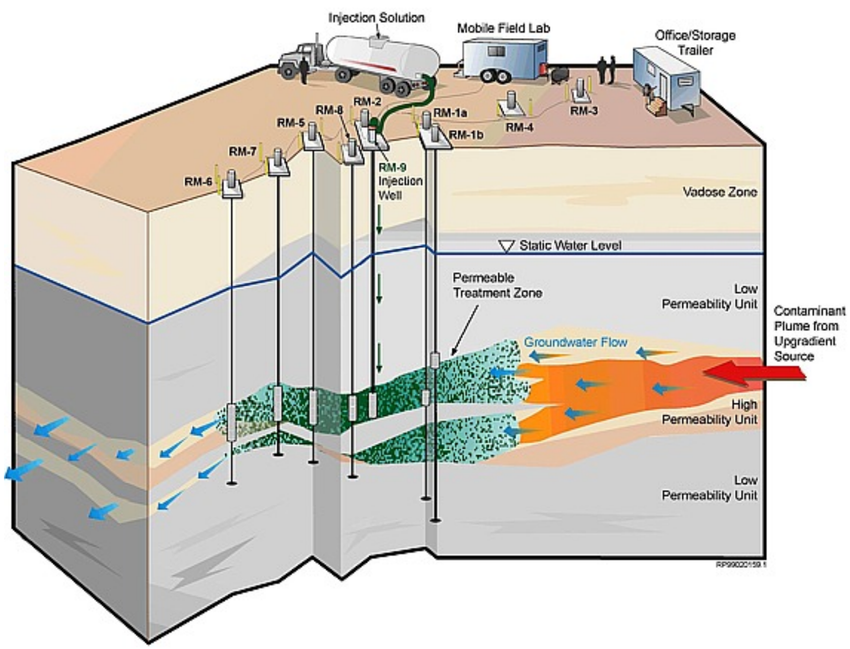
- Need exascale (10^{18} ops/sec) computing to design for alternative fuels, new devices

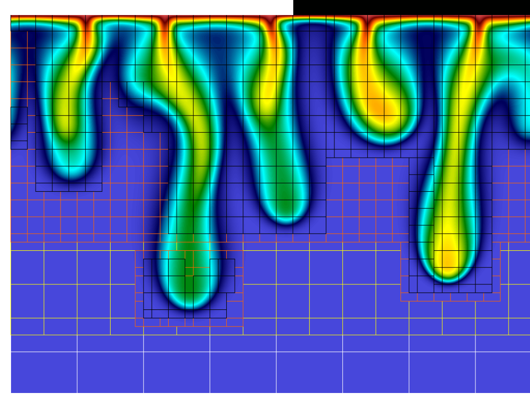
Simulating New Kinds of Batteries



Simulations to Get Rid of CO_2

- Carbon sequestration: "The process of removing carbon from the atmosphere or from flue gasses and depositing it in a reservoir."
- $CaO + CO_2 \rightarrow CaCO_3$



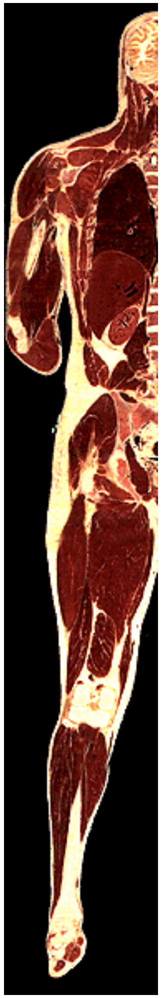


George Pau, LBNL

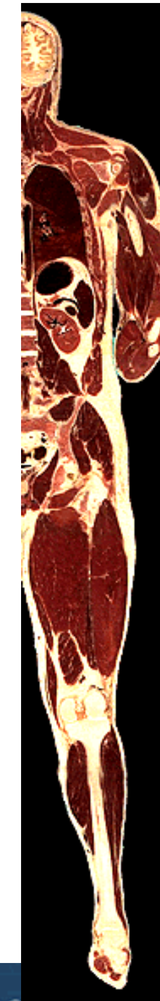
Towards a Digital Human: The 20+ Year Vision

- Imagine a "digital body double"
  - 3D image-based medical record
  - Includes diagnostic, pathologic, and other information
- Used for:
  - Diagnosis
  - Less invasive surgery-by-robot
  - Experimental treatments

Digital Human Today: Imaging



- The Visible Human Project
  - 18,000 digitized sections of the body
    - Male: 1mm sections, released in 1994
    - Female: .33mm sections, released in 1995
  - Goals
    - study of human anatomy
    - testing medical imaging algorithms
  - Current applications:
    - educational, diagnostic, treatment planning, virtual reality, artistic, mathematical and industrial
    - Used by > 1,400 licensees in 42 countries



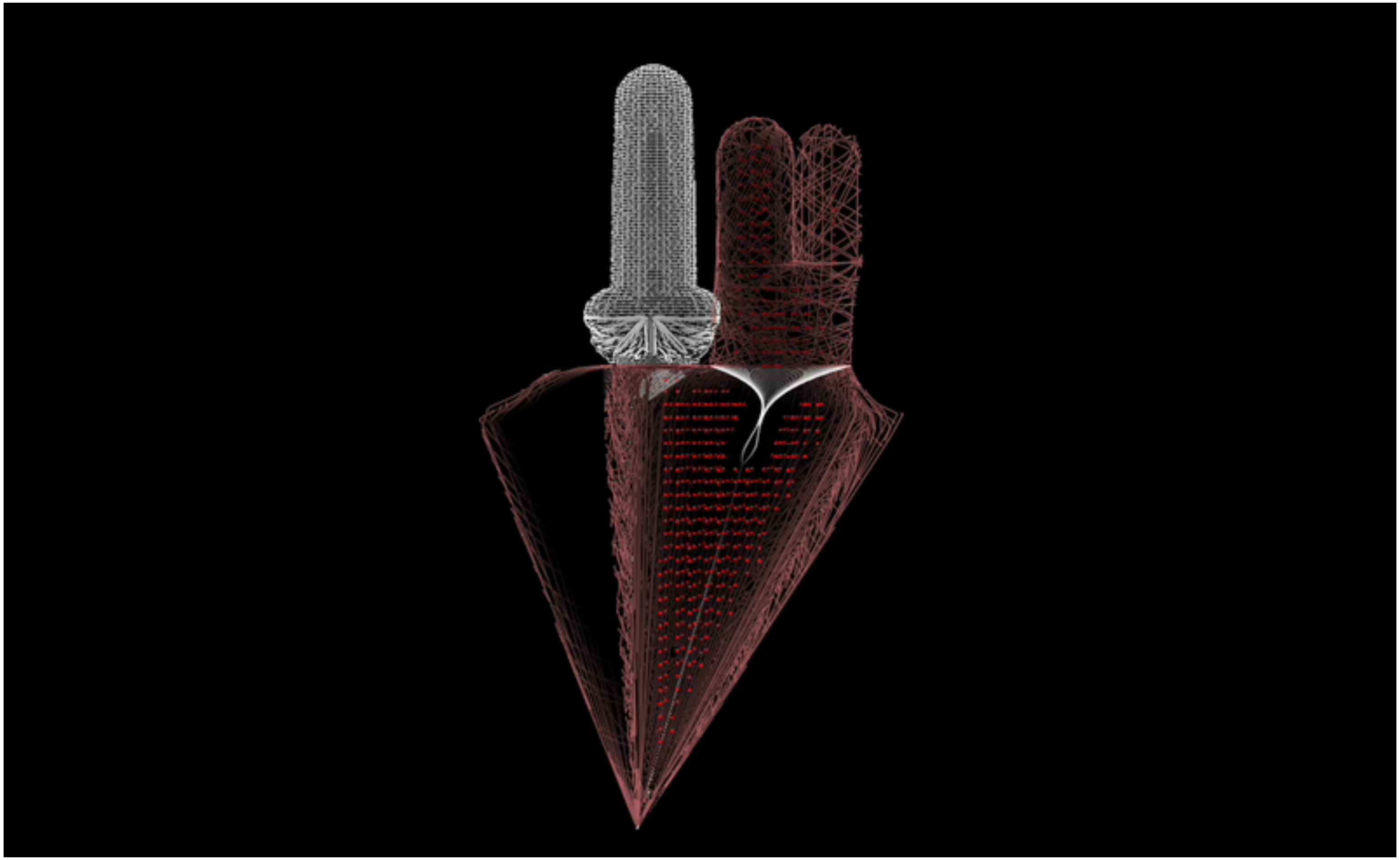
THE VISIBLE HUMAN PROJECT®

Image Source: www.madsci.org

Experimental Data: Visible Human

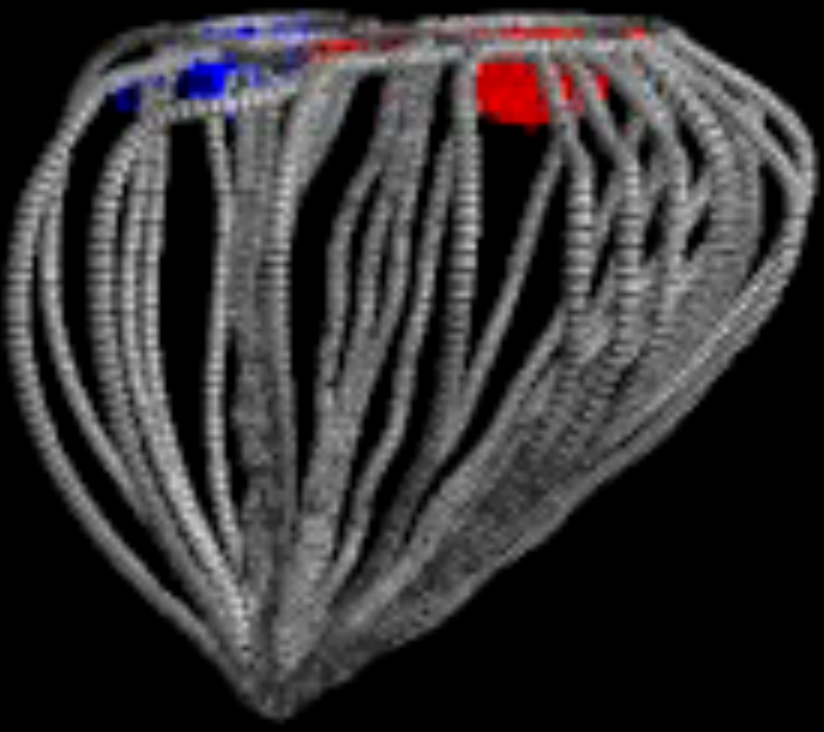
Heart Simulation

Movie from Boyce Griffith's PhD thesis, NYU

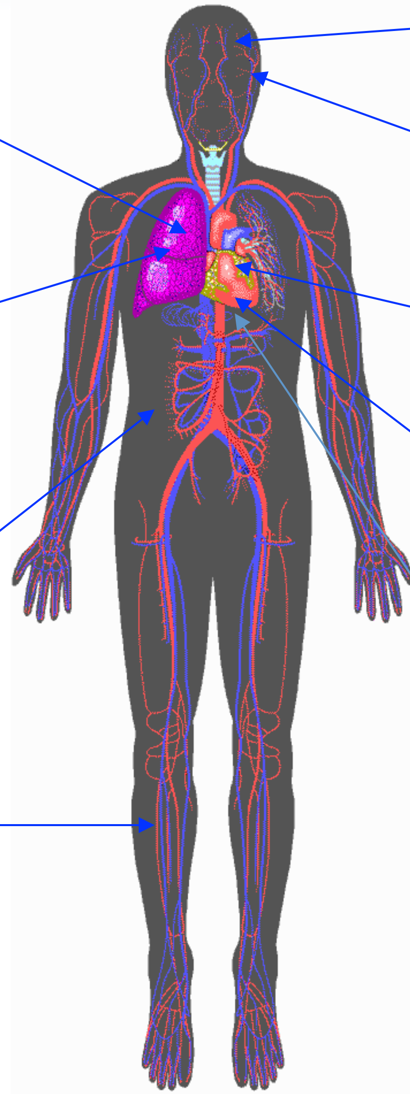


Heart Simulation

Movie from Charles Peskin and Dave McQueen at NYU



Organ Simulation



Brain
UCSD (Ellisman), IBM

Lung transport
Vanderbilt

Cochlea
Caltech, UMichigan

Lung flow
U. Iowa (Lin & Hoffman)

Cardiac flow
NYU, UCB, UCD…

Cardiac cells/muscles
SDSC, Auckland, UW, Utah

Kidney mesh generation
Dartmouth

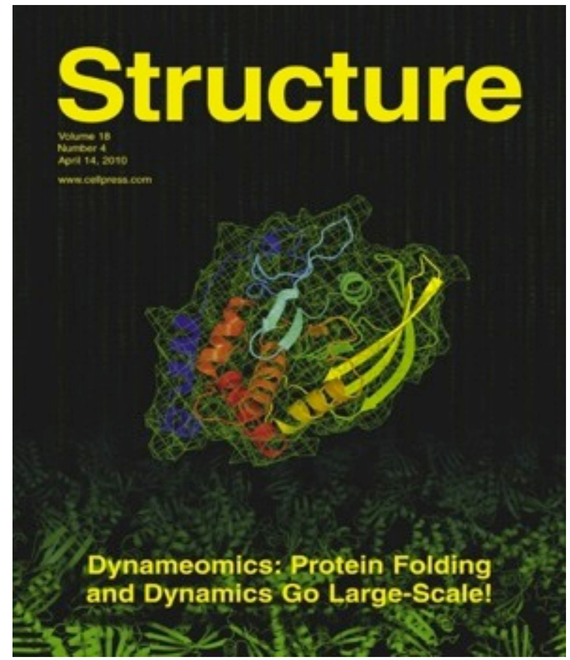
Electrocardiography
Johns Hopkins,…

Skeletal mesh generation

Just a few of the efforts at understanding and simulating parts of the human body

Screening Proteins

- Large number of simulations covering a variety of related proteins,…



<u>Dynameomics Database</u>

Improve understanding of disease and drug design, e.g., 11,000 protein unfolding simulations stored in a public database. [V. Daggett, UW]

Big D and Big C: Computing on Big Data to help Cure Cancer

ESSAY

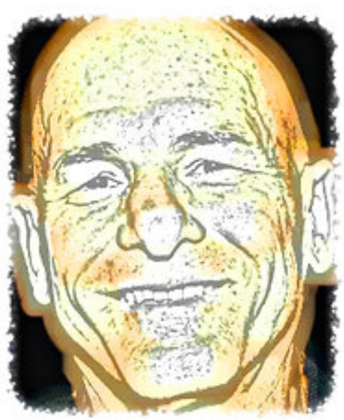
Computer Scientists May Have What It Takes to Help Cure Cancer

By DAVID PATTERSON

Published: December 5, 2011

The war against cancer is increasingly moving into cyberspace. Computer scientists may have the best skills to fight cancer in the next decade — and they should be signing up in droves.

Enlarge This Image



David Patterson

One reason to enlist: Cancer is so pervasive. In his Pulitzer Prize-winning book, "The Emperor of All Maladies," the oncologist Siddhartha Mukherjee writes that cancer is a disease of frightening fractions: One-fourth of deaths in the United States are caused by cancer; one-third of women will face cancer in their lifetimes; and so will half of men.

As he wrote, "The question is not if we will get this immortal disease, but when."

RECOMMEND
TWITTER
LINKEDIN
SIGN IN TO E-MAIL
PRINT
REPRINTS
SHARE

Why Study Computer Science?

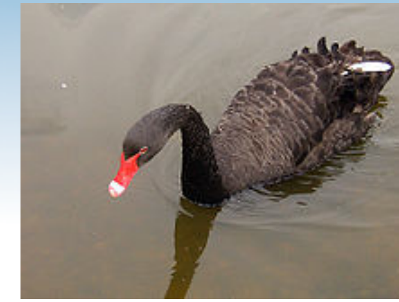
1) Because computers can help solve important problems
2) **Because programming is fun and there are plenty of new problems to solve**

Trends in Computer Science

Which of the following are true?

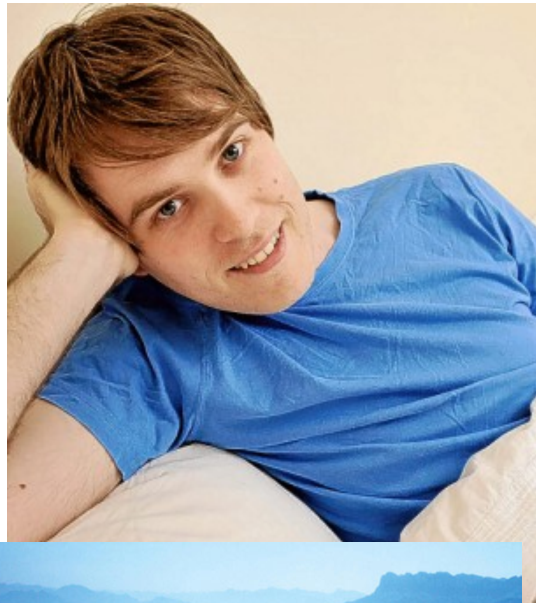
A. Moore's Law says that processor performance doubles every 18 months
B. Moore's Law has ended
C. Current computers are fast enough for most applications
D. None of the above
E. All of the above

Black Swans of Computing



Google™

2012 Computing with 1992 Technology



Technology for Innovation

Which of the following are true?

A. Google developed its own programming language to hide machine failures
B. iPhones are programmed using Java
C. Web search algorithms use only integer arithmetic, not floating point (real) numbers
D. Scientific computing is done mostly using "Vector Supercomputers"
E. All of the above

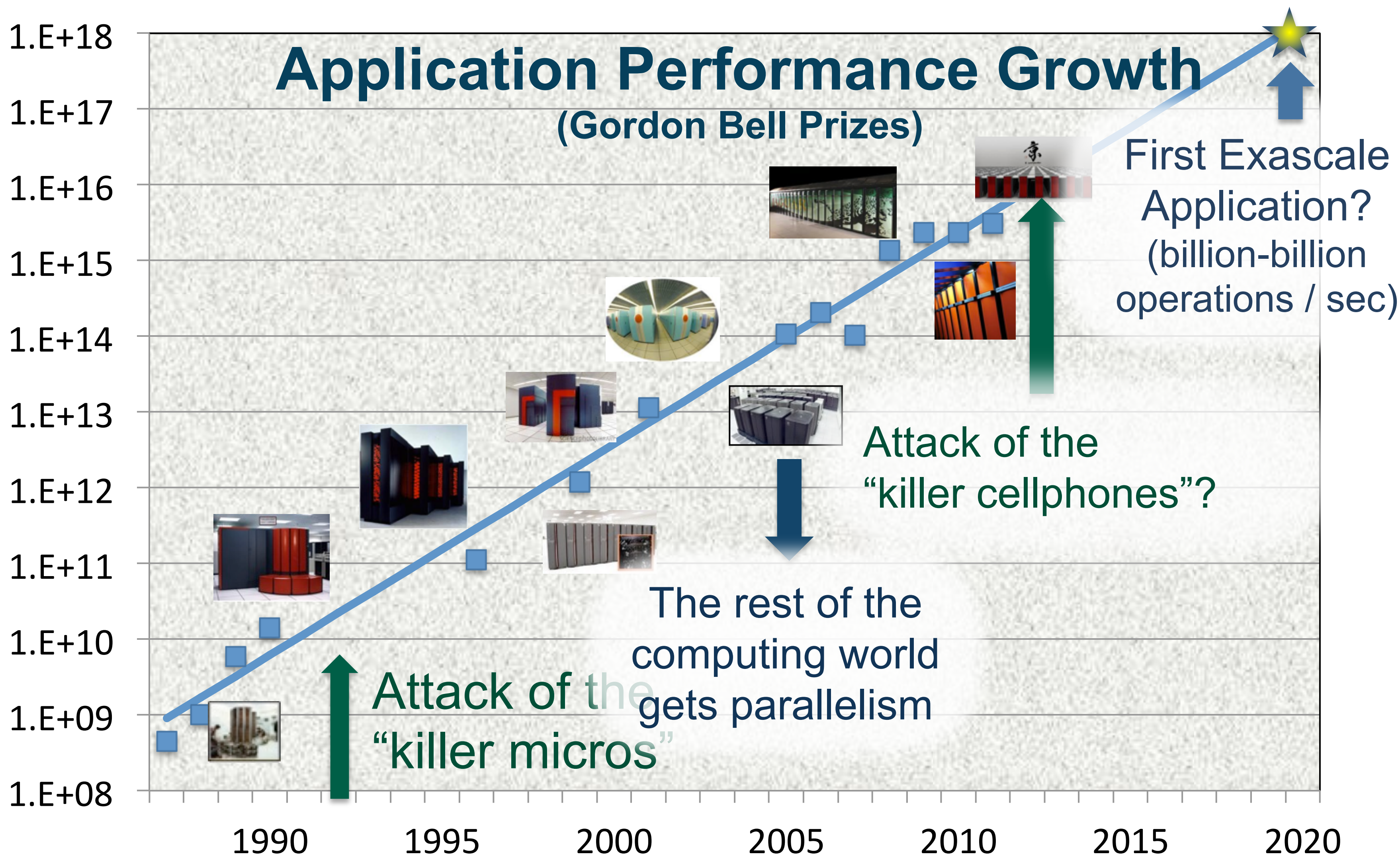
Units of Measure in High Performance Computing (HPC)

- High Performance Computing (HPC) units are:
  - Flops: floating point operations
  - Flops/s: floating point operations per second
  - Bytes: size of data (a double precision floating point number is 8)

- Typical sizes are millions, billions, trillions…

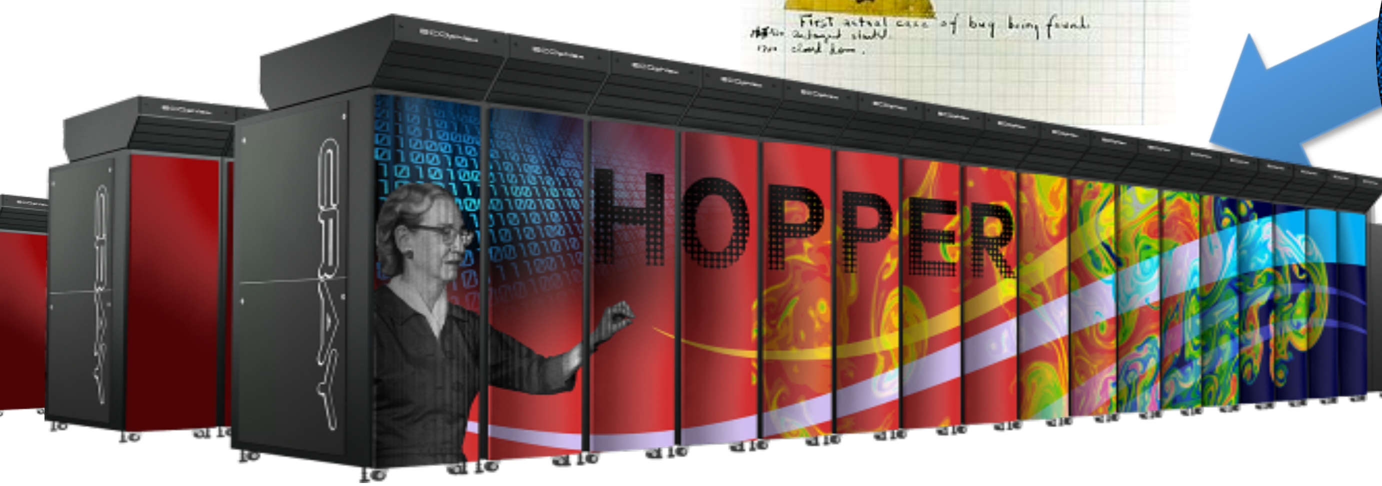
	Kilo	Kflop/s = 10^3 flop/sec	Kbyte = 2^{10} = 1024 ~ 1,000 bytes
	Mega	Mflop/s = 10^6 flop/sec	Mbyte = 2^{20} = 1048576 ~ 10^6 bytes
1988	Giga	Gflop/s = 10^9 flop/sec	Gbyte = 2^{30} ~ 10^9 bytes
1998	Tera	Tflop/s = 10^{12} flop/sec	Tbyte = 2^{40} ~ 10^{12} bytes
2008	Peta	Pflop/s = 10^{15} flop/sec	Pbyte = 2^{50} ~ 10^{15} bytes
	Exa	Eflop/s = 10^{18} flop/sec	Ebyte = 2^{60} ~ 10^{18} bytes
	Zetta	Zflop/s = 10^{21} flop/sec	Zbyte = 2^{70} ~ 10^{21} bytes
	Yotta	Yflop/s = 10^{24} flop/sec	Ybyte = 2^{80} ~ 10^{24} bytes

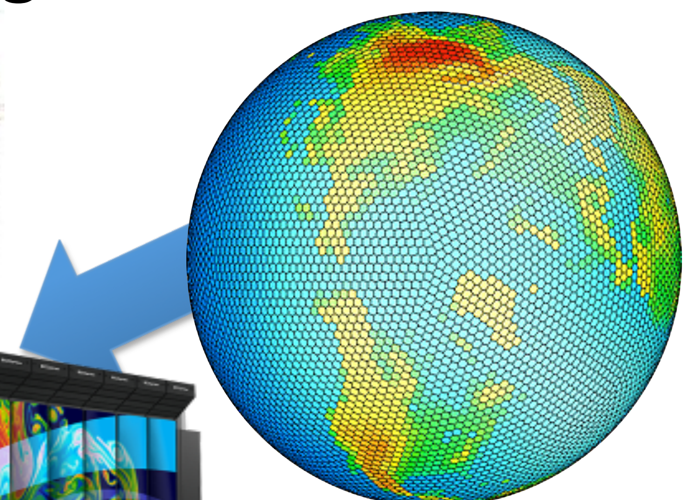
High End Computing Revolutions



The Fastest Computers (for Science) Have Been Parallel for a Long Time

- Fastest Computers in the world: top500.org
- Our Hopper Computer has 150,000 cores and > 1 Petaflop (10^{15} math operations / second)
- Programming and "debugging" are challenging



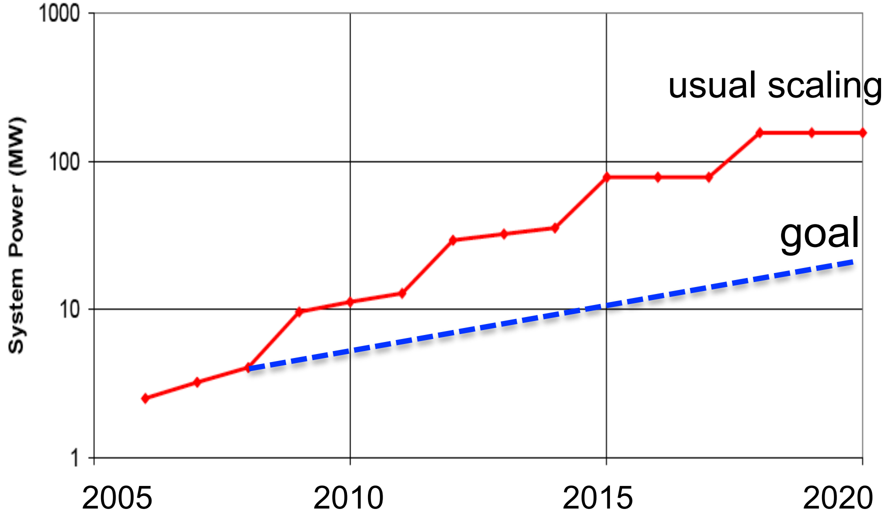


Supercomputing is done by parallel programming

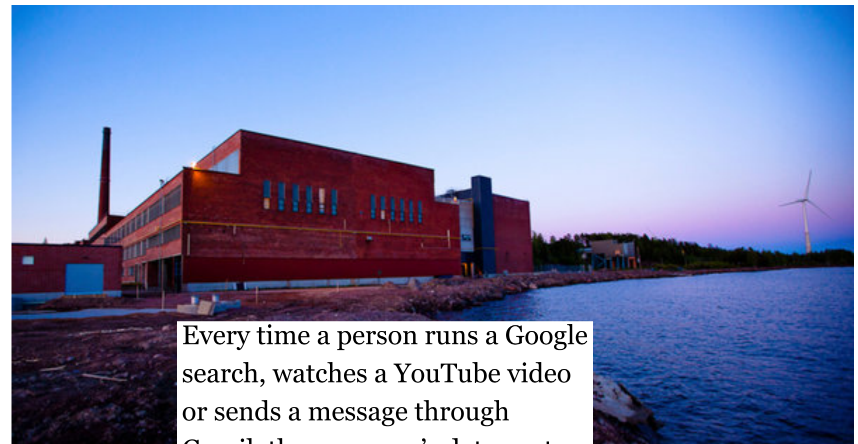
Energy Challenge for Computing

At ~$1M per MW, energy costs are substantial

An exaflop in 2020 would use ~200 MW with "usual" scaling



Google Details, and Defends, Its Use of Electricity



Google

Every time a person runs a Google search, watches a YouTube video or sends a message through Gmail, the company's data centers full of computers use electricity. Those data centers around the world continuously draw almost 260 million watts — about a quarter of the output of a nuclear power plant.

NSA Maxes Out Baltimore Power Grid

August 6th, 2006 : Rich Miller

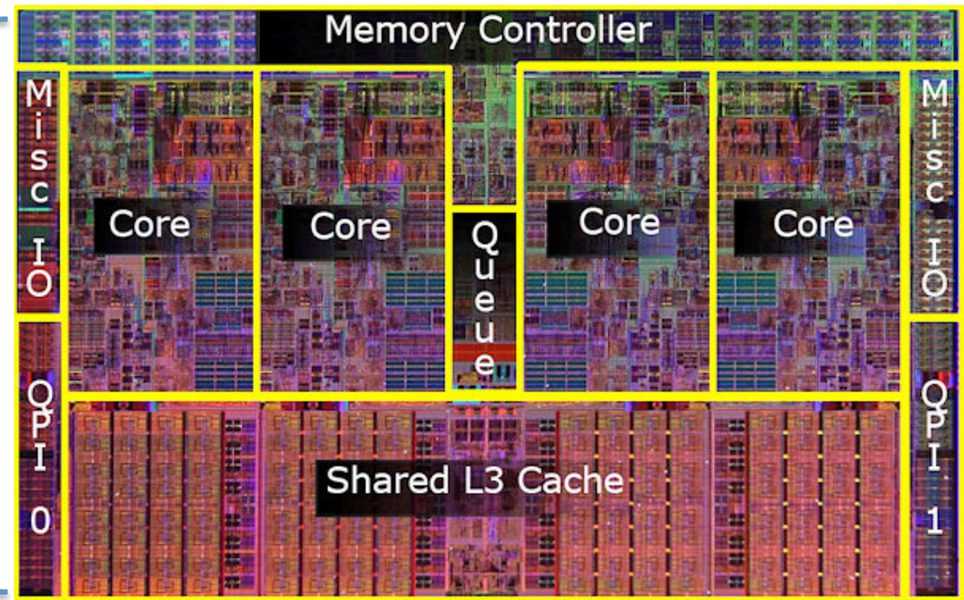
The National Security Agency's technology infrastructure at Fort Meade, Md. has maxed out the electric capacity of the Baltimore area power grid, creating a major challenge for the agency, sources told the Baltimore Sun. An excerpt:

The worldwide data center power in was about 26 gigawatts in 2010 (up from 17 in 2005)

New Processor Designs are Needed to Save Energy

Cell phone processor
(0.1 Watt, 4 Gflop/s)

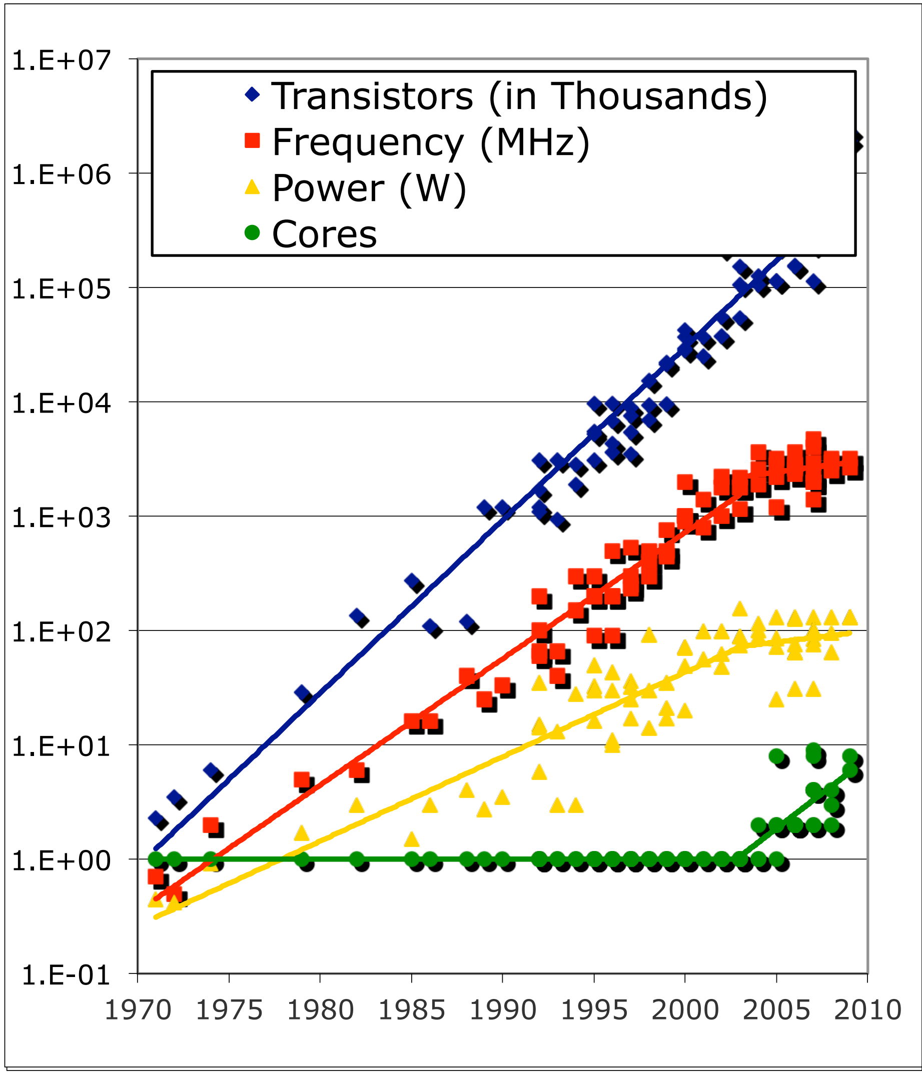
Server processor
(100 Watts, 50 Gflop/s)



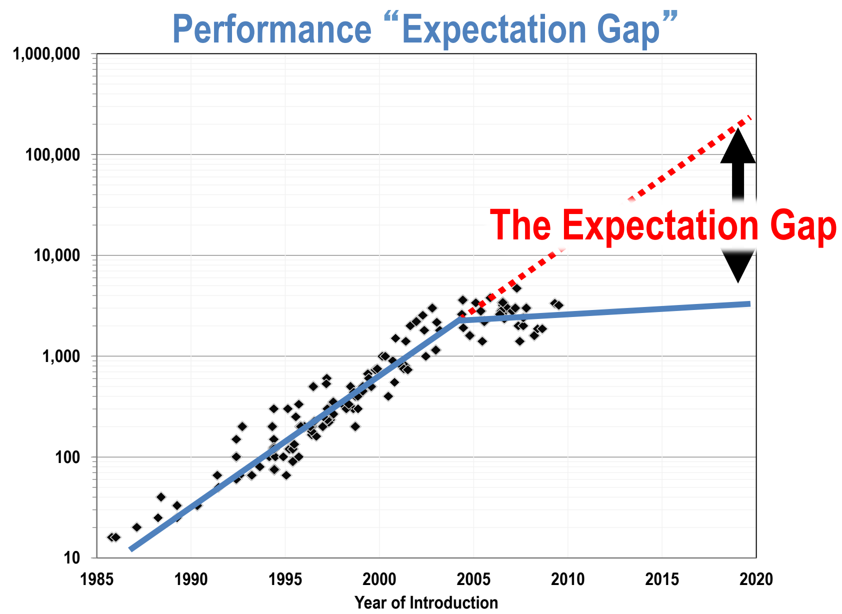
- The server is about 10x faster than the cell phone processor
- But uses 1000x more power → cell phone is 100x more efficient
- Why: Power is proportional to $V^2 f$, and increasing frequency (f) also requires increase voltage V → cube
- Next computers built from graphics, games, cell phones,…

All Computers are Parallel Computers

- Power density limit single processor clock speeds
- Cores per chip is growing
- How to program them?
  - Parallel “loops”
  - Parallel map
  - Parallel divide-and-conquer
  - (Message passing)

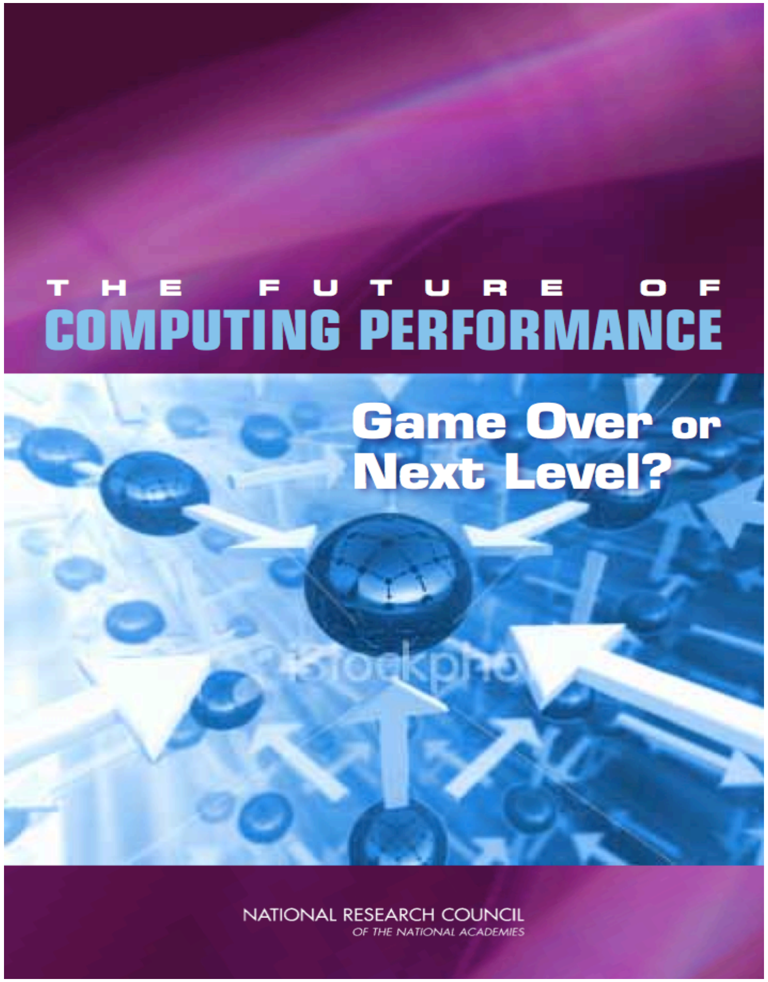


Power Limits Computing Performance Growth



Processor industry was running at "maneuvering speed"

- David Liddle

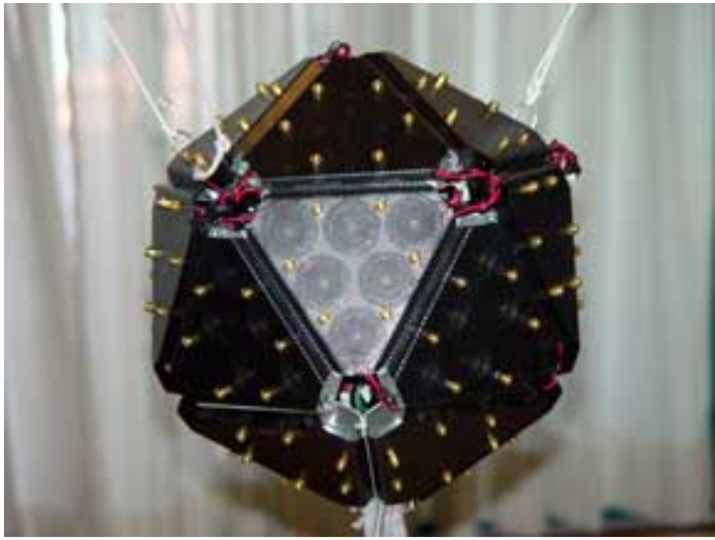


Why Study Computer Science?

1) Because computers can help solve important problems
2) Because computers are fun to program
3) **Because computers make a good career**

Computation in Music
(David Wessel)

- Musicians have an insatiable appetite for computation
  - More channels, instruments, more processing, more interaction!
  - Latency must be low (5 ms)
  - Must be reliable (No clicks)
- Music Enhancer
  - Enhanced sound delivery systems for home sound systems using large microphone and speaker arrays
  - Laptop/Handheld recreate 3D sound over ear buds
- Hearing Augmenter
  - Handheld as accelerator for hearing aid



Berkeley Center for New Music and Audio Technology (CNMAT) created a compact loudspeaker array: 10-inch-diameter icosahedron incorporating 120 tweeters.

Real-Time Deformation and Fracture in a Game Environment

Eric Parker
Pixelux Entertainment

James O'Brien
U.C. Berkeley

Video Edited by Sebastian Burke

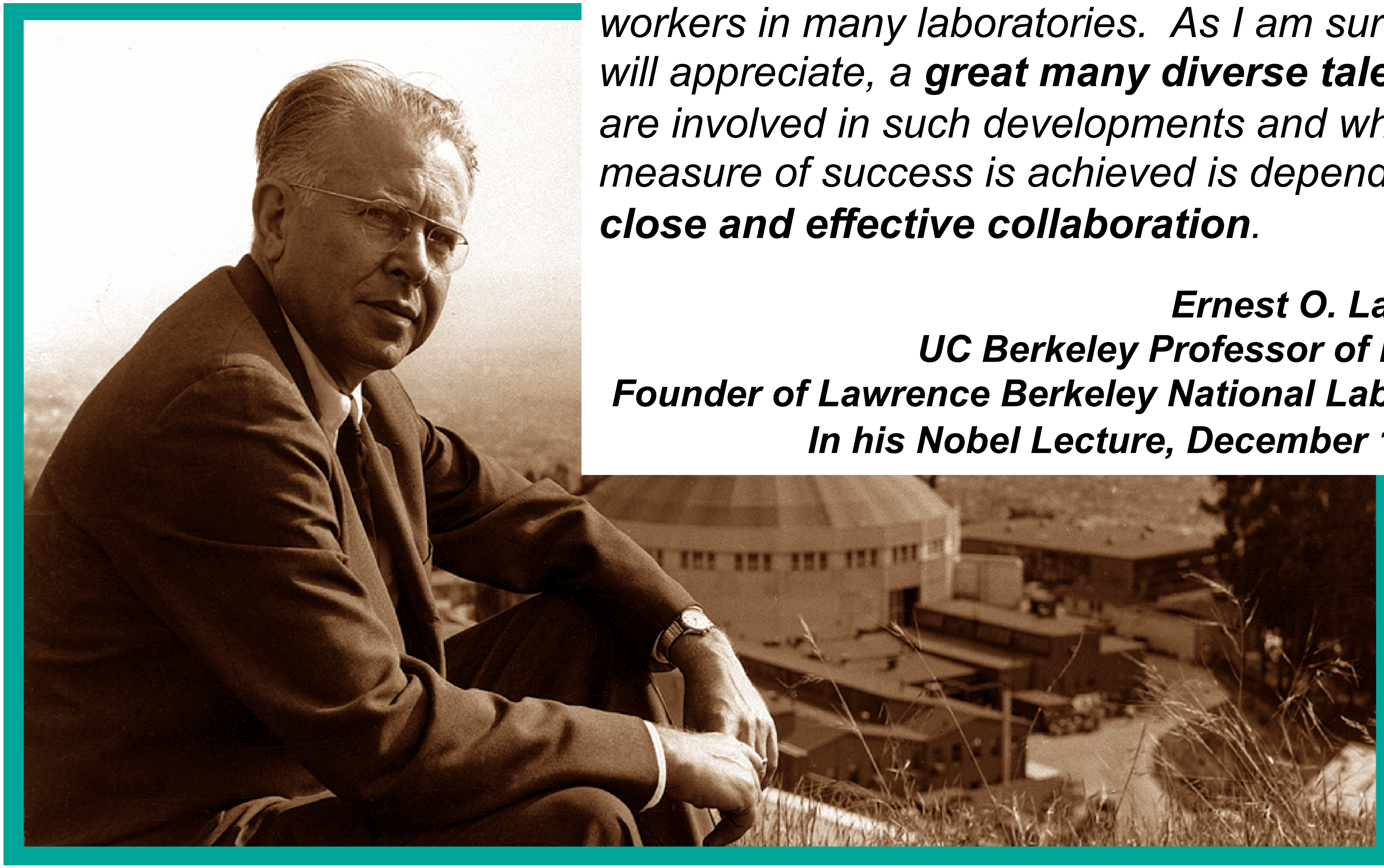
From the proceedings of SCA 2009, New Orleans

Writing Software

Which of the following are true?

A. Most computer software is written by brilliant hackers, working alone
B. Parallel programming is a *solved problem*
C. Speed of programming and speed of programs are the top goals in software
D. Most software is rewritten from scratch every few years
E. None of the above

Computational Science is Necessarily Collaborative



… as from the beginning the work has been a team effort involving many able and devoted co-workers in many laboratories. As I am sure you will appreciate, a ***great many diverse talents*** *are involved in such developments and whatever measure of success is achieved is dependent on* ***close and effective collaboration****.*

Ernest O. Lawrence
UC Berkeley Professor of Physics
Founder of Lawrence Berkeley National Laboratory
In his Nobel Lecture, December 11, 1951

Internships Available: http://csee.lbl.gov/

Why Study Computer Science?

1) Because computers can help solve important problems
2) Because computers are fun to program
3) Because computers make a good career
4) **Because you will get to work with lots of great people**