Computation Across the Curriculum

Charlie Peck, Earlham College

First Annual Tapia LittleFe Scientific Programming Competition

Brought to you by a cast of thousands...

Why Computer Science Matters

Charlie Peck, Mobeen Ludin, Ivan Babic, Sanele Mahlalela
MEAP @ IUPUI
July, 2012

Well, how did we get here?

- Computational power
- Storage volume
- Algorithmic and computational approaches

Digital World

- Simulations based on models
- Analysis of large data sets
- Visualization of large data sets
- Techniques that depend on high performance computing gear to be practical; automating vs enabling
- Third and fourth methods of ~~scientific~~ intellectual inquiry
- 20th vs 21st century ~~science~~ inquiry; ~~atoms and icebergs~~ the occupants of El Sidrón

Computational Methods

- Why use simulations?

  too small (atoms, molecules)

  too large (galaxies, the universe)

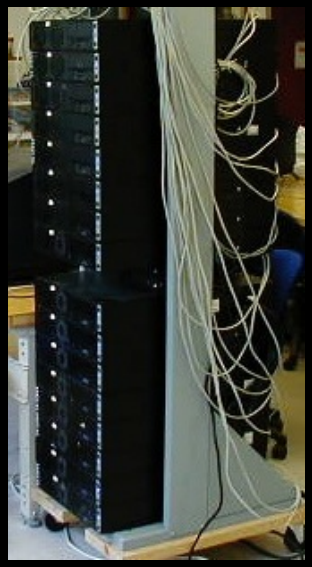
  too fast (photosynthesis, protein folding)

  too slow (geological processes, climate change)

  too complex (blood circulation, weather)

  too dangerous (toxic materials, nuclear stockpile stability)

High Performance Computing Gear



	Bazaar	Cairo	BobSCEd	Al-salam
Year	2000	2003	2006	2010
GFLOPS	18	128	666	~3000
Size	32U	16U	8U	13U

Moore's law, every 18 months the density of transistors in integrated circuits roughly doubles

Moore's Law in Action

1950	32 Bytes	10^0	phone booth
1975	640 KB	10^3	shoe box
2000	256 MB	10^6	pack of gum
2008	2 GB	10^9	credit card
2030	1 TB	10^12	?

Typical amount of RAM in a "desktop" computer

Data > Information > Knowledge

- Complete works of Shakespeare - ~ 5 MB
- Human genome - ~ 1 GB
- Complete works of Beethoven - ~ 20 GB
- Medical imaging - ~ 30 GB per scan
- Library of Congress - ~ 10 TB
- All US academic libraries - ~ 2 PB
- Large Hadron Collider - ~1.5 GB/second

Increasingly Readings > Data ...

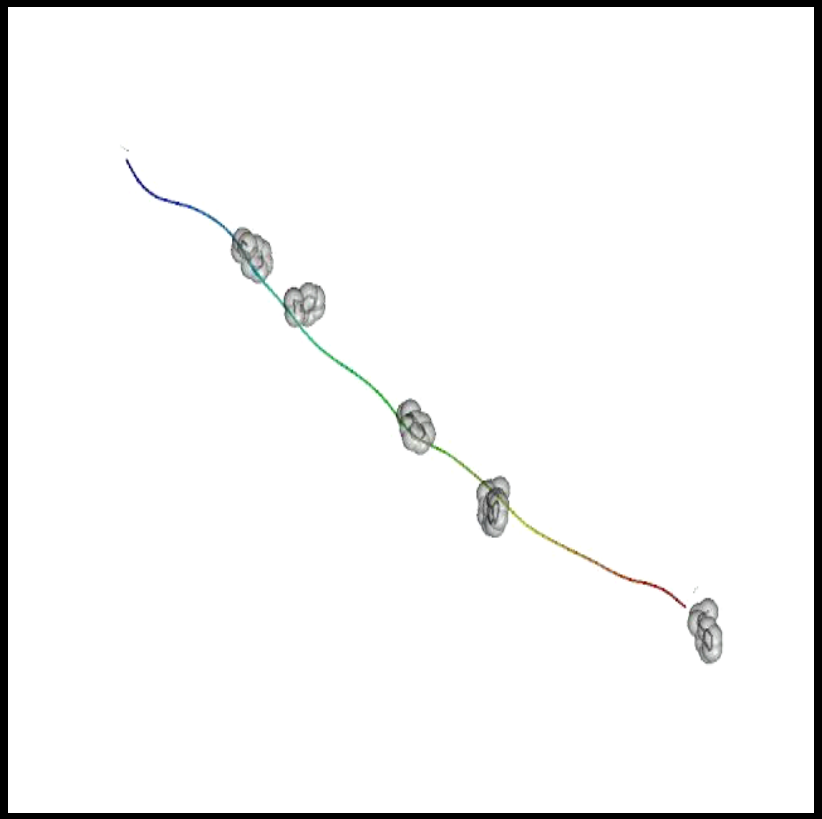
Parallel and Distributed Computing

- Decomposing large problems into smaller ones, solving the smaller problems, and then reducing those answers to find "the answer"
- Domain decomposition
- Functional decomposition
- Shared memory systems
- Message passing systems

Natural Sciences

- Modeling and Simulation
  Protein folding
  Earthquakes
  Phylogenetic reconstruction
  Genome construction
- Data Sets
  Sloan Digital Sky Survey
  Protein Data Bank
  Arctic aerial photographs
  Geographical information systems (GIS)

Protein Folding



Source: Pande Lab, Stanford University

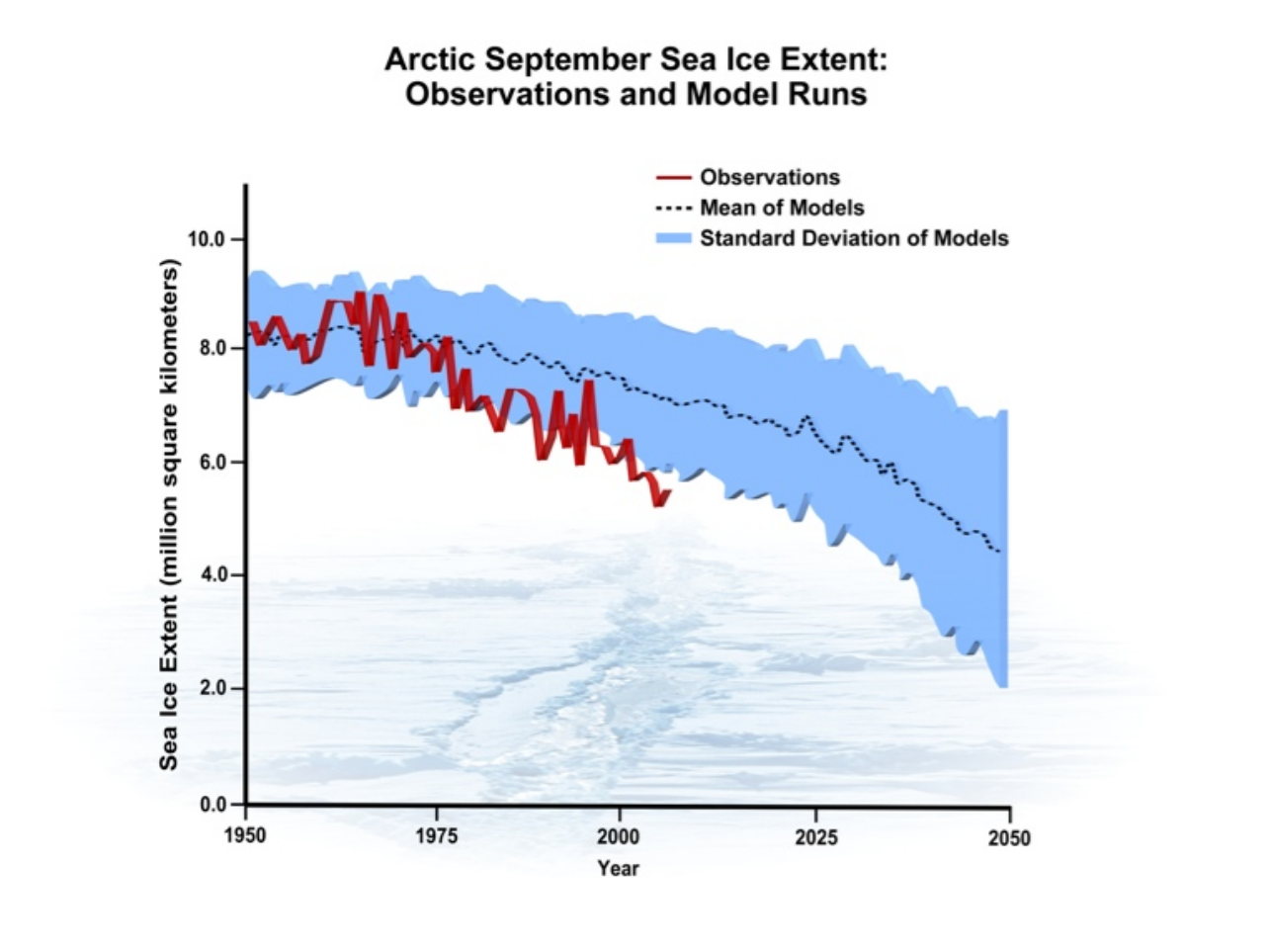
Result of an ensemble molecular dynamics simulation (Gromacs) of the villin headpiece

Earthquakes

Combined view with Camera Match

Source: San Diego Supercomputer Center

Climate Change



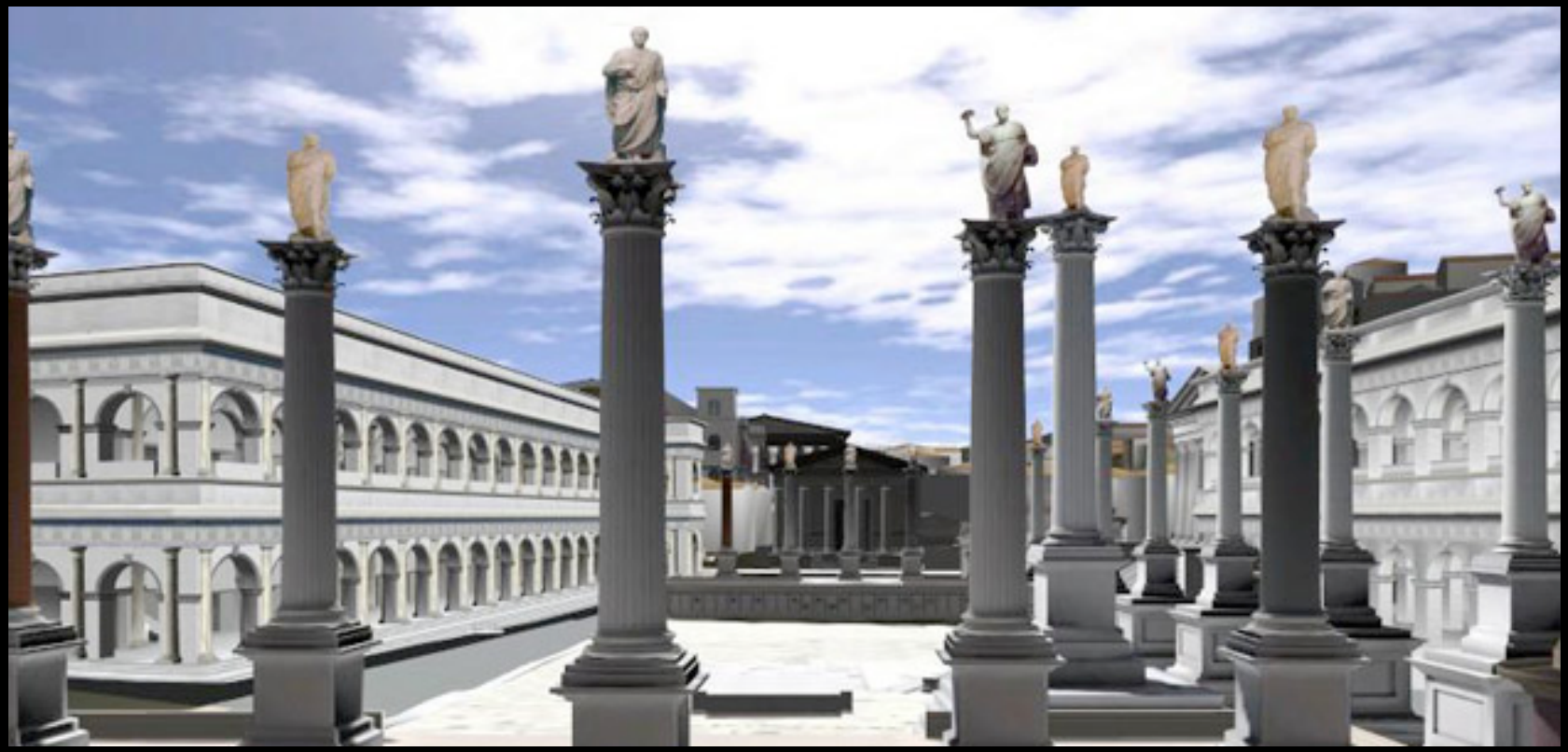
Source: National Center for Atmospheric Research (NCAR) and the University of Colorado's National Snow and Ice Data Center (NSIDC)

Humanities

- Modeling and Simulation
  Game theory
  Topic modeling
  Text analysis

- Data Sets
  Library of Congress
  Project Gutenberg
  Newspaper morgues

Rome Reborn



Source: University of California Irvine

Looking out from the Roman Forum in a complete interactive 3D model of Rome

Arts

- Modeling and Simulation
  Animation and rendering
  Painting provenance
  Digital music

- Data Sets
  Photograph archives
  Scanned paintings, sculptures, buildings
  Digital recordings

It's Everywhere

- Anthropology *
- Archeology
- Classics
- Economics
- Film
- History
- Language Acquisition
- Library Science
- Literature
- Painting *
- Politics
- Psychology
- Sculpture

Is that really a Van Gogh?

Source: Christian Science Monitor

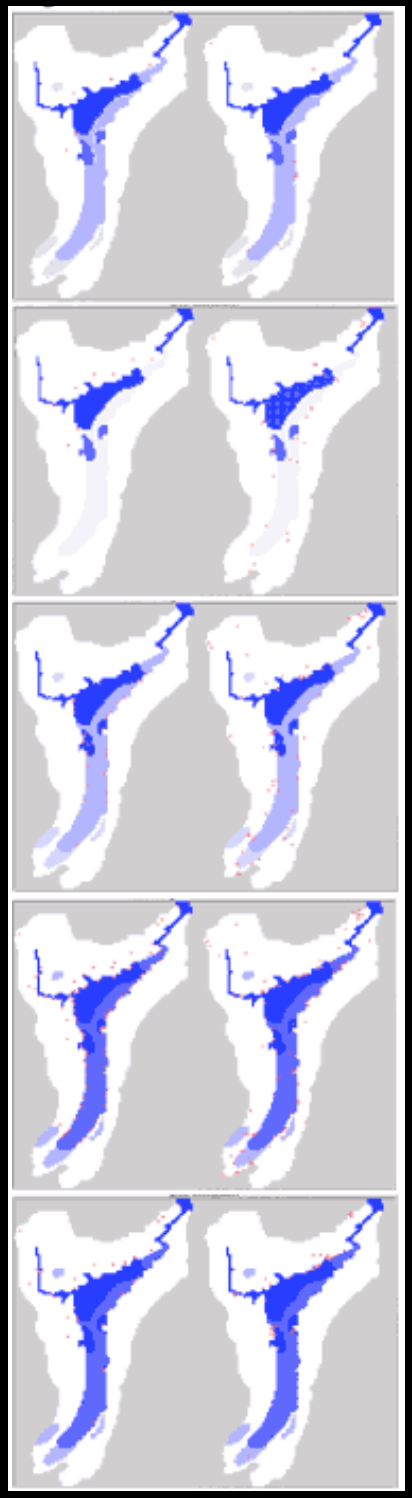
Vase with 15 Sunflowers

Social Sciences

- Modeling and Simulation
  Teacher matching
  Social systems
  Derivatives analysis

- Data Sets
  Census
  Geographical information systems (GIS)
  Voting records
  Transaction records (commercial and civil)

What happened to the Anasazi?



Source: Jonathan Rauch, The Atlantic

Who were those people?

- Forensic analysis
- Archeology
- Genome sequencing
- Anthropology

Technology in the Wild

Science in the Field

Education, Outreach, and Training

- National Computational Science Institute/SC Education Program/ Blue Waters UPEP
- Workshops for undergraduate faculty and students
- 1800 participants at 80 workshops over the past 7 years
- Specialize in first-touch to handoff to a regional provider
- 3D Internet - metaverse and web-based interactive applications
- Curriculum development
- Outreach to high schools

Why Do All This?

- President's Information Technology Advisory Committee: "...computational science is one of the most important technical fields of the 21st century..."
- Rising Above The Gathering Storm: "...vastly improving K-12 and undergraduate science and mathematics education..."
- Bio2010: "...exposure during the early years of their undergraduate careers will help life science students use current computer methods and learn how to exploit emerging computer technologies as they arise..."

Future

- Software
  Making sense of all that stuff we are collecting
  [Readings] > Data > Information > Knowledge
  Visualization, interactive interfaces
- Grid
  Science portals, e.g. TeraGrid, Nanohub, Open Science Grid
  Humanities, Arts, and Social Science portals
- Hardware
  Specialized CPUs, e.g. FPGA and graphics chips
  Cores, cores, and more cores