Powering Cyber-research in the 21st Century

Thom H. Dunning, Jr.
National Center for Supercomputing Applications
University of Illinois at Urbana-Champaign
Outline of Presentation

• NCSA Strategic Directions

• Examples of Current NCSA Projects
  • Cyber-resources
  • Advanced Information Systems (Cyberenvironments)
  • Data-intensive Computing
  • Advanced Visualization

• Institute for Advanced Computing Applications & Technologies
NCSA Strategic Directions

- **Cyber-resources**
  ... enabling scientific discovery at the leading edge
  - Provide leading edge computing and data storage resources and network connectivity as well as user services
  - Assess value of innovative computing technologies for computational science and engineering
  - Develop and deploy a *sustained* petaflops system: *Blue Waters*

- **Cyberenvironments**
  ... harnessing the power of the national cyberinfrastructure
  - Develop integrated, end-to-end software environments to coordinate, automate, and apply high-end resources and capabilities
  - Develop cyberservices and cybertechnologies needed to build cyberenvironments
NCSA Strategic Directions

- Data-intensive Computing
  ... processing massive data sets, extracting knowledge from data
  - Manage, process and analyze large data sets and data streams
  - Develop and apply automated analysis and learning techniques

- Advanced Visualization
  ... providing critical insights, educating the public
  - Develop scientific visualization for scientists and engineers
  - Create popular visualizations for the public

- Cybersecurity (V. Welch, R. Butler)
  ... maintaining the security of our systems and data
  - Prevention
  - Detection and response
NCSA Strategic Directions (cont’d)

• Cybereducation
  ... bringing cyber-enabled science into the classroom
  • Scientific literacy through computational science and technology
  • Virtual School of Computational Science and Engineering
Cyber-resources
... enabling scientific discovery at the leading edge
Increased computing power enables advances in a broad range of science and engineering disciplines:

- Molecular Science
- Weather & Climate Forecasting
- Astronomy
- Earth Science
- Health
Blue Waters Project

NSF’s Strategy for High-end Computing

**Track 1 System**
- UIUC/NCSA (≥1 PF sustained)

**Track 2 Systems**
- PSC (?)
- UT/ORNL (~1PF)
- TACC (500+TF)
- Multiple Systems

**Track 3 Systems**
- Leading University HPC Centers

Science & Engineering Capability (logarithmic scale)

- 2006
- 2007
- 2008
- 2009
- 2010
- 2011
Blue Waters Project

Input from Scientific Community

- **D. Baker, University of Washington**
  Protein structure refinement and determination

- **M. Campanelli, RIT**
  Computational relativity and gravitation

- **D. Ceperley, UIUC**
  Quantum Monte Carlo molecular dynamics

- **J. P. Draayer, LSU**
  Ab initio nuclear structure calculations

- **P. Fussell, Boeing**
  Aircraft design optimization

- **C. C. Goodrich**
  Space weather modeling

- **M. Gordon, T. Windus, Iowa State University**
  Electronic structure of molecules

- **S. Gottlieb, Indiana University**
  Lattice quantum chromodynamics

- **V. Govindaraju**
  Image processing and feature extraction

- **M. L. Klein, University of Pennsylvania**
  Biophysical and materials simulations

- **J. B. Klemp et al., NCAR**
  Weather forecasting/hurricane modeling

- **R. Luetich, University of North Carolina**
  Coastal circulation and storm surge modeling

- **W. K. Liu, Northwestern University**
  Multiscale materials simulations

- **M. Maxey, Brown University**
  Multiphase turbulent flow in channels

- **S. McKee, University of Michigan**
  Analysis of ATLAS data

- **M. L. Norman, UCSD**
  Simulations in astrophysics and cosmology

- **J. P. Ostriker, Princeton University**
  Virtual universe

- **J. P. Schaefer, LSST Corporation**
  Analysis of LSST datasets

- **P. Spentzouris, Fermilab**
  Design of new accelerators

- **W. M. Tang, Princeton University**
  Simulation of fine-scale plasma turbulence

- **A. W. Thomas, D. Richards, Jefferson Lab**
  Lattice QCD for hadronic and nuclear physics

- **J. Tromp, Caltech/Princeton**
  Global and regional seismic wave propagation

- **P. R. Woodward, University of Minnesota**
  Astrophysical fluid dynamics
Blue Waters Project

Attributes of Petascale Computing System

• Maximize Core Performance
  … minimize number of cores needed for a given level of performance as well as lessen the impact of sections of code with limited scalability

• Maximize S&E Application Scalability
  … low latency, high-bandwidth communications fabric

• Solve Memory-intensive Problems
  … large amount of memory
  … low latency, high-bandwidth memory subsystem

• Solve Data-intensive Problems
  … high-bandwidth I/O subsystem
  … large quantity of on-line disk, massive quantity of archival storage

• Provide Reliable Operation
  … maximize system integration
  … mainframe reliability, availability, serviceability (RAS) technologies
## Blue Waters Project

**Blue Waters Computing System**

<table>
<thead>
<tr>
<th>System Attribute</th>
<th>TACC Ranger</th>
<th>NCSA Blue Waters</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vendor</strong></td>
<td>Sun</td>
<td>IBM</td>
</tr>
<tr>
<td><strong>Processor</strong></td>
<td>AMD Barcelona</td>
<td>IBM Power7</td>
</tr>
<tr>
<td><strong>Peak Performance (PF)</strong></td>
<td>0.579</td>
<td>&gt;20</td>
</tr>
<tr>
<td><strong>Sustained Performance (PF)</strong></td>
<td>~0.05</td>
<td>~1</td>
</tr>
<tr>
<td><strong>Number of Cores/Chip</strong></td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td><strong>Number of Processor Cores</strong></td>
<td>62,976</td>
<td>&gt;3</td>
</tr>
<tr>
<td><strong>Amount of Memory (TB)</strong></td>
<td>123</td>
<td>&gt;6</td>
</tr>
<tr>
<td><strong>Amount of Disk Storage (PB)</strong></td>
<td>1.73 (s)</td>
<td>&gt;5</td>
</tr>
<tr>
<td><strong>Amount of Archival Storage (PB)</strong></td>
<td>2.5 (20)</td>
<td>&gt;200</td>
</tr>
<tr>
<td><strong>External Bandwidth (Gbps)</strong></td>
<td>10</td>
<td>&gt;10</td>
</tr>
</tbody>
</table>
Blue Waters Project

Illinois-IBM Collaborative Projects. I

- **Computing Systems Software**
  - Goal: enhance IBM’s HPC software stack
  - Examples
    - Integrated System Management Console
    - Petascale Application Development Environment
    - Computational Libraries
    - Programming models

- **Science and Engineering Applications**
  - Goal: prepare applications to fully utilize Blue Waters’ capabilities
  - Process
    - **Before hardware**: extensive use of simulators and Track 2 systems to optimize processor and communications performance
    - **After hardware**: further optimization for Power7 processor, node, …, full scale system
Blue Waters Project
Illinois-IBM Collaborative Projects. II

• Computing Systems Hardware
  • Goal: enhance the performance of the base Blue Waters system
  • Example
    • Evaluation of accelerators

• Petascale Computing Facility
  • Goal: advance “green” computing
  • Elements
    • On-site cooling towers with ambient water
    • Efficient electrical distribution system
    • PUE < 1.2 (Power Usage Effectiveness)
Blue Waters Project

Petascale Computing Facility

• Modern Data Center
  • 90,000+ ft² total
  • 20,000 ft² machine room

• Energy Efficiency
  • LEED certified (goal: gold)
  • Highly efficient cooling system

www.ncsa.uiuc.edu/BlueWaters
Blue Waters Project
Great Lakes Consortium for Petascale Computation

Goal: Facilitate the widespread and effective use of petascale computing to address frontier research questions in science, technology and engineering at research, educational and industrial organizations across the region and nation.

Charter Members

Argonne National Laboratory
Fermi National Accelerator Laboratory
Illinois Math and Science Academy
Illinois Wesleyan University
Indiana University*
Iowa State University
Illinois Mathematics and Science Academy
Krell Institute, Inc.
Louisiana State University
Michigan State University*
Northwestern University*
Parkland Community College
Pennsylvania State University*
Purdue University*

The Ohio State University*
Shiloh Community Unit School District #1
Shodor Education Foundation, Inc.
SURA – 60 plus universities
University of Chicago*
University of Illinois at Chicago*
University of Illinois at Urbana-Champaign*
University of Iowa*
University of Michigan*
University of Minnesota*
University of North Carolina–Chapel Hill
University of Wisconsin–Madison*
Wayne City High School

* CIC universities*
Advanced Information Systems

... harnessing the power of the national cyberinfrastructure
Advanced Information Systems
National Cyberinfrastructure

Hardware
- Computers
- Data sources
- Data stores
- Networks

Software
- Middleware
- Portals
- Grid-enabled
  - Applications
  - Visualization
  - Data analysis
- ...

CS 591 • 21 October 2009 • University of Illinois at Urbana-Champaign
**Major New Data Sources**

**Computers**
New high-end computers are producing massive amounts of data from ever more detailed computational models.

**Sensors, Surveys and Satellites**
Sensor arrays, aerial surveys and satellite data will revolutionize our understanding of the environment.

**Instruments**
New instruments, *e.g.*, telescopes and detectors, are using advanced digital technologies to support increasingly detailed observations.
Advanced Information Systems
Cyberenvironment for Earthquake Engineering

• Mid-America Earthquake Center
  Consequence-based risk management for seismic events
  • Portal-based collaboration environment
  • Distributed data/metadata sources
  • Multi-disciplinary collaboration
Advanced Information Systems

Cyberenvironments for Environmental Observatories

Experiment
Gather Data
Run Analysis
Publish to the Web for others to explore

Portal
Sensors

Workflow
Modeling Services
Analysis/Visualization

CS 591 • 21 October 2009 • University of Illinois at Urbana-Champaign
Data-intensive Computing

... processing massive data sets, extracting knowledge from data
Large Synoptic Survey Telescope (LSST)

- **New Telescope**
  - Located in Chile (El Peñon) with first light in 2013
  - 8.4-m Mirror with 3 Gigapixel camera
  - Image available sky every 3 days

- **Science Missions**
  - Nature of dark energy and accelerating universe
  - Comprehensive census of solar system objects, create galactic map
  - Explore transients and variable objects

- **Data Sets**
  - 15 terabytes per night (raw)
  - 100 petabytes in 10 years

Data-transport, processing and analysis pipelines
Advanced Visualization

... providing critical insights, educating the public
Visualization of an F3 Tornado within a Supercell Thunderstorm Simulation

Computation and Visualizations
National Center for Supercomputing Applications
University of Illinois at Urbana–Champaign
IACAT Vision and Mission

• Vision
  • To make Illinois a world leader in the innovative use of advanced computing and information technologies to address the most challenging research problems in a broad range of disciplines

• Mission
  • Combine research in UIUC’s academic units with advanced technology development and deployment at NCSA to:
    • Create and apply innovative computing *applications*
    • Create and apply innovative computing *technologies*
  • Transfer advances to the broader research community through NCSA’s cyberinfrastructure programs
IACAT Research Themes and Projects

• **Center for Extreme-scale Computation**
  • Petascale Computing Software Tools and Applications (L. Kale, CS; D. Johnson, MSE)
  • Next Generation Computing Technologies (W-m. Hwu, ECE)
  • Multiscale Simulation in Science and Engineering (R. Braatz, CBE)

• **Advanced Information Systems**
  • Cyberenvironment for Research in Environmental Sustainability (P. Kumar, CEE; B. Minsker, CEE; D. Wuebbles, AS)

• **Computing and Creativity**
  • Cultural Informatics (M. Ross, KCPA; G. Garnett, Music; D. Cox, A&D/NCSA)
Questions?