Cyberinfrastructure: The Future and Its Challenges

Oklahoma Supercomputing Symposium 2003
September 25, 2003

Peter A. Freeman

Assistant Director of NSF for
Computer & Information Science & Engineering

In summary then, the opportunity is here to create cyberinfrastructure that enables more ubiquitous, comprehensive knowledge environments that become functionally complete for specific research communities in terms of people, data, information, tools, and instruments and that include unprecedented capacity for computational, storage, and communication... They can serve individuals, teams and organizations in ways that revolutionize *what they can do, how they do it,* and *who participate*.

- The Atkins Report

Overview

- Context
- A ten-year vision
- Challenges
- Q/A

Context

Is There a Definition of Cyberinfrastructure (CI)?

- Not really - means different things to different groups - but there are commonalities
- Literally, infrastructure composed of “cyber” elements
- Includes High-End Computing (HEC, or supercomputing), grid computing, distributed computing, etc. etc.

Is There a Definition of Cyberinfrastructure (CI)?

- **Working definition**: *an integrated system of interconnected computation/communication/information elements that supports a range of applications*

- **Note:** There is an ***extant*** **CI** today. What we are really talking about is an ***emergent*** **CI**.

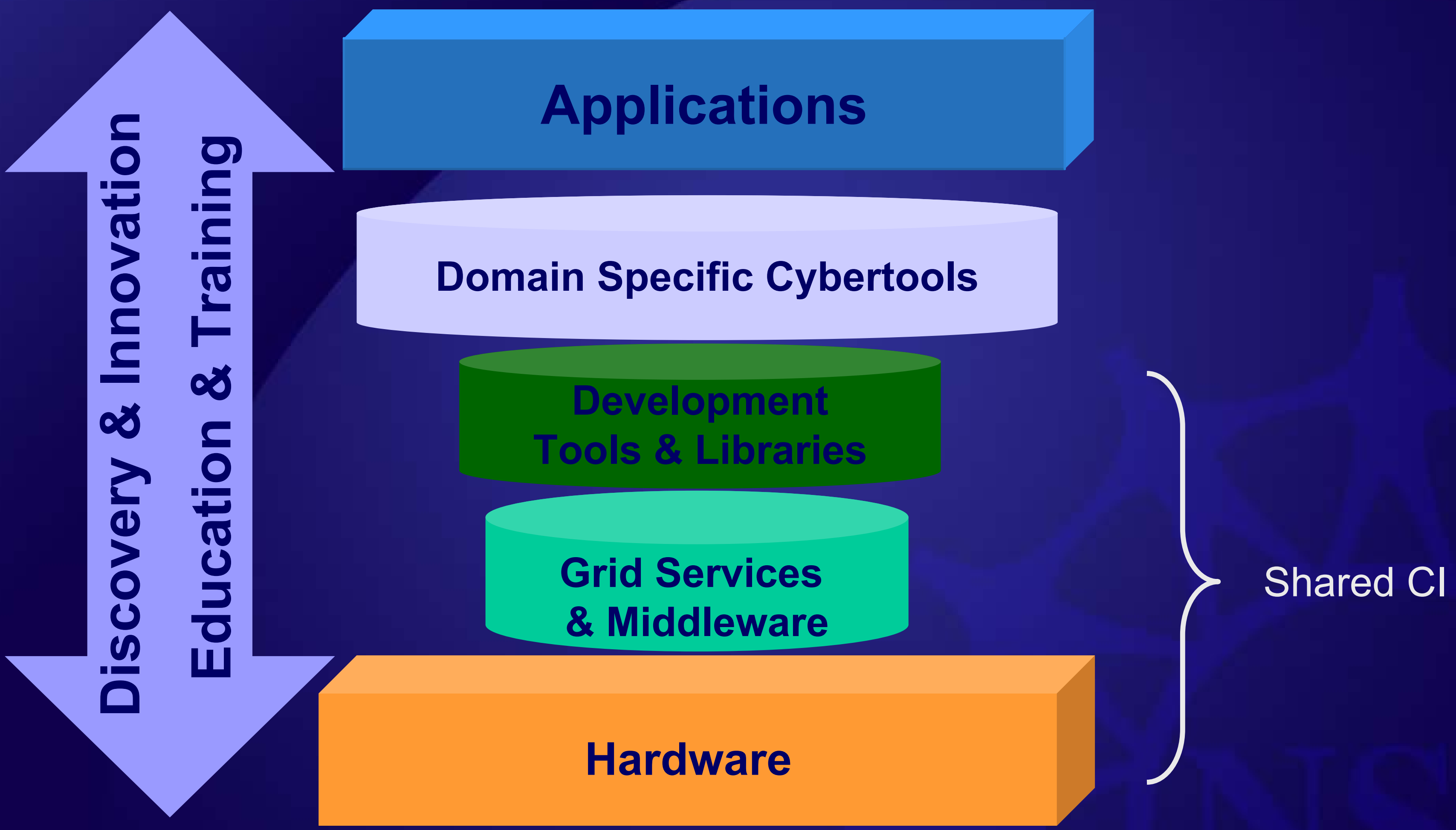
Cyberinfrastructure is the ***means;*** "e-Science" is the ***result***

Cyberinfrastructure consists of …

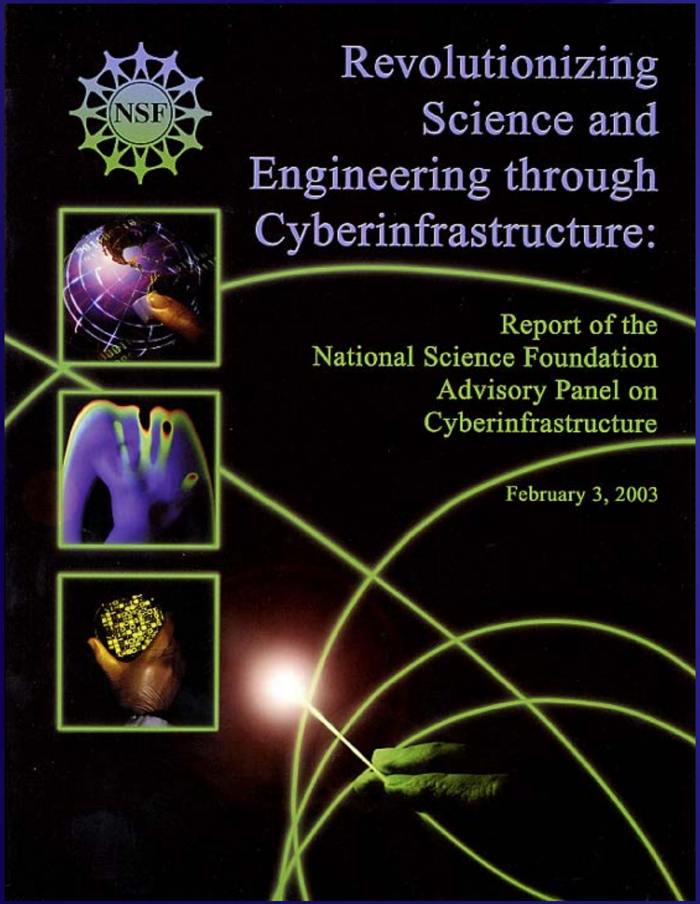
- Computational engines (supercomputers, clusters, workstations, small processors, …)
- Mass storage (disk drives, tapes, …)
- Networking (including wireless, distributed, ubiquitous)
- Digital libraries/data bases
- Sensors/effectors
- Software (operating systems, middleware, domain specific tools/platforms for building applications)
- Services (education, training, consulting, user assistance)

All working together in an integrated fashion.

Integrated Cyberinfrastructure…



The Atkins Report



Daniel E. Atkins, Chair
University of Michigan
Kelvin K. Droegemeier
University of Oklahoma
Stuart I. Feldman
IBM
Hector Garcia-Molina
Stanford University
Michael L. Klein
University of Pennsylvania
David G. Messerschmitt
University of California at Berkeley
Paul Messina
California Institute of Technology
Jeremiah P. Ostriker
Princeton University
Margaret H. Wright
New York University

http://www.cise.nsf.gov/evnt/reports/toc.htm

“[Science is] a series of peaceful interludes punctuated by intellectually violent revolutions . . .[in which] . . . one conceptual world view is replaced by another.”

--Thomas Kuhn

From *The Structure of Scientific Revolutions*

Evolution of the Scientific Enterprise

- Pre-science (< 1000 CE)
- Observational (< 1600 CE)
- Empirical (> 1600 CE)
- Theoretical (>1650 CE)
- Computational (> 1950 CE)
- Informational (> 2000 CE)

A Ten-year Vision for Cyberinfrastructure

In Ten Years, a CI That Is…

- rich in resources, comprehensive in functionality, and ubiquitous;
- easily usable by all scientists and engineers, from students to emertii;
- accessible anywhere, anytime needed by authenticated users;
- interoperable, extendable, flexible, tailorable, and robust;
- funded by multiple agencies, states, campuses, and organizations;
- supported and utilized by educational programs at all levels.

Some Characteristics of a Future Cyberinfrastructure

- Built on broadly accessible, highly capable network: 100’s of terabits backbones down to intermittent, wireless connectivity at very low speeds;
- Contains significant and varied computing resources: 100’s of petaflops at high end, with capacity to support most scientific work;
- Contains significant storage capacity: exabyte collections common; high-degree of DB confederation possible;
- Allows wide range of sensors/effectors to be connected: sensor nets of millions of elements attached;
- Contains a broad variety of intelligent visualization, search, database, programming and other services that are fitted to specific disciplines

Challenges

Technical Challenges

- Computer Science and Engineering broadly
- How to build the components?
- Networks, processors, storage devices, sensors, software
- How to shape the technical architecture?
- Pervasive, many cyberinfrastructures, constantly evolving/changing capabilities
- How to customize CI to particular S&E domains

Operational Challenges

- Data standards
- General interoperability
- Resource allocation
- Security and privacy
- Training
- Continuous evolution

Funding/Ownership Challenges

- Cooperation among agencies
- Cooperation between federal and state/private levels
- Role of campuses
- Interaction with private industry
- $$$$$!

Educational Challenges

- How to make sure that future generations of scientists and engineers can fully utilize CI
  - New paradigms, methods, objectives
- How to retrain current scientists and engineers
- How to make sure that new ideas for extending CI continue to come from those that are using it

CI in Transition

Principles

- Build on what we've learned to date
- Provide new funding opportunities for extant and emerging providers and users
- Encourage partnerships between CI users and computing specialists
- Promote flexibility, interoperability and competition for best ideas

CI in Transition

Funding Strategies

- Maintain essential CI resources and services while providing new funding opportunities for current and future CI providers and users
- Explore new governance models, emphasizing partnerships among computing and domain specialists both domestic and foreign
- Advance the state-of-the-art in cyberinfrastructure capability, including the development of promising new architectures, tools and applications
- Create a portfolio of education, outreach, training and community development activities to enrich, support and expand the impact of cyberinfrastructure on research and education

Summary

- Cyberinfrastructure is already engendering a *revolution* in S&E
- The ubiquity, interconnectedness, and power of CI resources in the future will radically change S&E in the next 10 years
- Education for CI and use of CI in education are the two greatest challenges

The NSF Cyberinfrastructure Objective

- To lead the country in providing an integrated, high-end system of computing, data facilities, connectivity, software, services, and instruments that ...

- enables all scientists and engineers to work in new ways on advanced research problems that would not otherwise be solvable.

Enabling the nation's future through discovery, learning and innovation

Contact Information

Dr. Peter A. Freeman

NSF Assistant Director for CISE

Phone: 703-292-8900

Email: pfreeman@nsf.gov

Visit the NSF Web site at:

www.nsf.gov