Work in Progress:
Status Report on the
Office of Cyberinfrastructure

Daniel E. Atkins
Director,
National Science Foundation
Office of Cyberinfrastructure
datkins@nsf.gov
Nomenclature

- ACLS-Mellon Study: Cyberinfrastructure for Humanities
  http://www.acls.org/cyberinfrastructure/
- HASTAC Project
  http://www.hastac.org/
- e-research
- e-science
- e-infrastructure
- Cyber science

e = electronic | enhanced | enabled
Dualities

**Cyberinfrastructure**

*CI is both an object and means for R&D*

**Collaboration**

*Multi-stakeholder collaboration required to create, provision, and apply CI; CI supports collaborations across time and distance (geographic, disciplinary, institutional)*

**Learning | Education**

*Learning and workforce development initiatives required to create and use CI; CI enables/enhances learning/education*
Vision and Activities Based on Broad and Diverse Community Engagement

- Advances in components of CI-systems for S&E R&E
- Complex, multi-scale, multidisciplinary S&E research challenges
- 30+ disciplinary workshops on CI vision & impact
- NSF internal working groups
- CI Council, Directorate/Office CI Activities, OCI, ACCI

Vision Framework:
- High Performance Computing
- Data, Data Analysis & Visualization
- Virtual Organizations
- Learning & Workforce Development

- All directorates and offices support cyberinfrastructure.
- Science-driven partnerships between creation, provisioning and use of CI
- Supports integrated research and education and broadened access and participation.
Some Science Drivers

- Inherent complexity and multi-scale nature of today's frontier science challenges.
- Requirement for multi-disciplinary, multi-investigator, multi-institutional approach (often international).
- High data intensity from simulations, digital instruments, sensor nets, observatories.
- Increased value of data and demand for data curation & preservation of access.
- Exploiting infrastructure sharing to achieve better stewardship of research funding.
- Strategic need for engaging more students in high quality, authentic science and engineering education.
## NSF CI FY07 Budget Request
Total of $600M in CI Funding with $182M in OCI

### Cyberinfrastructure Funding

(Dollars in Millions)

<table>
<thead>
<tr>
<th>Category</th>
<th>FY 2005 Actuals</th>
<th>FY 2006 Plan</th>
<th>FY 2007 Request</th>
<th>Change over FY 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biological Sciences</td>
<td>$77.00</td>
<td>$84.00</td>
<td>$90.50</td>
<td>$6.50, 7.7%</td>
</tr>
<tr>
<td>Computer and Information Science and Engineering</td>
<td>45.32</td>
<td>63.00</td>
<td>68.00</td>
<td>5.00, 7.9%</td>
</tr>
<tr>
<td>Engineering</td>
<td>52.00</td>
<td>52.00</td>
<td>54.00</td>
<td>2.00, 3.8%</td>
</tr>
<tr>
<td>Geosciences</td>
<td>71.35</td>
<td>71.35</td>
<td>75.00</td>
<td>3.65, 5.1%</td>
</tr>
<tr>
<td>Mathematical and Physical Sciences</td>
<td>56.52</td>
<td>59.30</td>
<td>63.56</td>
<td>4.26, 7.2%</td>
</tr>
<tr>
<td>Social, Behavioral and Economic Sciences</td>
<td>20.39</td>
<td>20.54</td>
<td>20.54</td>
<td>-</td>
</tr>
<tr>
<td>Office of Cyberinfrastructure</td>
<td>123.28</td>
<td>127.12</td>
<td>182.42</td>
<td>55.30, 43.5%</td>
</tr>
<tr>
<td>Office of International Science and Engineering</td>
<td>0.22</td>
<td>1.00</td>
<td>1.05</td>
<td>0.05, 5.0%</td>
</tr>
<tr>
<td>Office of Polar Programs</td>
<td>25.38</td>
<td>26.24</td>
<td>26.24</td>
<td>-</td>
</tr>
<tr>
<td>Subtotal, Research and Related Activities</td>
<td>471.47</td>
<td>504.55</td>
<td>581.31</td>
<td>76.76, 15.2%</td>
</tr>
<tr>
<td>Education and Human Resources</td>
<td>20.27</td>
<td>15.02</td>
<td>15.52</td>
<td>0.50, 3.3%</td>
</tr>
<tr>
<td>Total, Cyberinfrastructure Funding</td>
<td>$491.74</td>
<td>$519.57</td>
<td>$596.83</td>
<td>$77.26, 14.9%</td>
</tr>
</tbody>
</table>

Totals may not add due to rounding.
Achieving the NSF CI (e-science) Vision requires synergy between 3 types of activities

**Transformative Application** - to enhance discovery & learning

**Provisioning** - Creation, deployment and operation of advanced CI

**R&D** to enhance technical and social effectiveness of future CI environments

*Borromean Ring*: The three rings taken together are inseparable, but remove any one ring and the other two fall apart. See [www.liv.ac.uk/~spmr02/rings/](http://www.liv.ac.uk/~spmr02/rings/)
Mission of OCI

The mission of the OCI is to enhance significantly the ability of the NSF community to create, provision, and use the comprehensive cyberinfrastructure essential to 21st century advances in science and engineering. This goal is implicit in many areas of the new NSF Strategic Plan and is being pursued within the context of the evolving Cyberinfrastructure Vision for 21st Century Discovery.

OCI will serve the Foundation and the NSF community in its mission through three types of synergistic activity:

1. provisioning of cyberinfrastructure together with mechanisms for flexible, secure, coordinated sharing among collections of individuals, institutions, and resources;
2. partnerships with others in science/engineering-driven, transformative use of CI in research and education; and
3. partnerships with others in the transfer of the fruits of relevant R&D into the next generation of CI.

OCI is a cross-cutting enterprise that builds mutually beneficial partnerships with all parts of the NSF, with other Federal agencies, and with the large and growing CI/e-science initiatives in other countries. OCI is also the lead in supporting the Advisory Committee for CI (ACCI) for the Foundation.

Draft
New NSF Strategic Plan

Includes many imperatives for innovation in provisioning and transformative application of cyberinfrastructure to discovery and learning.

Available at http://www.nsf.gov/pubs/2006/nsf0648/nsf0648.jsp
Emerging CI Coordination Structure

- Advisory Committee on Cyberinfrastructure
- Other Federal Research & Mission Agency Programs
- National Science & Technology Council (OSTP)
- Committee on Technology
- Subcommittee on Networking & IT R&D
- Inter Agency Data Group
- International e-science, cyber science programs
- CI COUNCIL
  Director
  Dep. Dir
  BIO
  CISE
  EHR
  ENG
  GEO
  MPS
  OISE
  OCI
  OPP
  SBE
- HPC Coordinating Group
- Data Coordinating Group
- Other Coordinating Groups TBD
- ACCI Task Force Groups

Liaison with each Directorate/Office

Additional linkage within NSF through joint appointments and every OCI Program Officer having a liaison role with another Directorate or Office.
Proposed ACCI Task Force Groups

- Digital Data
- Integration of the NSF CI Portfolio & Alignment with S&E Needs
- CI and Competitiveness
- CI-enhanced Learning, Discovery, and Broadened Participation
Several Active Solicitations Posted

Seeking more program officers.

High Performance Computing

Increasingly important tool for understanding

1-50 TeraFLOPS
significant number of systems

100+ TeraFLOPS
multiple systems

1-10 Peta FLOPS
at least one system

NSF Focus FY 2006-10

Track 1: One solicitation funded over 4 years: $200M acquisition + additional O&M cost.

Track 2: Four solicitations over 4 years: $30M/yr acquisition + additional O&M cost. First track 1 approved 8-07

Satellite tobacco mosaic virus, P. Freddolino et al.

Aldehyde dehydrogenase, T. Wymore and S. Brown

I. Shipsey

The Environment

S.-Y. Kim, M. Lodge, C. Taber.

NSF Focus FY 2006-10

100+ TeraFLOPS

leading number of systems

1-50 TeraFLOPS

at least one system

1-10 Peta FLOPS

multiple systems

Life

Matter

Society

John Q Public

Increasingly important tool for understanding
• Challenges: **increased scale, heterogeneity, and re-use value** of digital scientific information and data. Inadequate digital preservation strategy of long-lived data.

• Taking initial steps to **catalyze the development** of a federated, global system of science and engineering data collections that is open, extensible, evolvable, (and appropriately curated and long-lived.)

• Complemented by a **new generation of tools** and services to facilitate data mining, integration, analysis, visualization essential to transforming data into knowledge.

• **NSF Leadership for OSTP/Interagency Working Group on Digital Data**
Instances of Virtual Organizations (VOs)

- **People**

**Interfaces for interaction, workflow, visualization and collaboration for individuals & distributed teams**

**Mechanisms for flexible secure, coordinated resource/services sharing among dynamic collections of individuals, institutions, and resources (the Grid or service layer problem)**

**Distributed, heterogeneous services for:**

- Computation
- Data, information management
- Sensing, observation, activation in the world

Alternate Names for Instances of VOs:
- Co-laboratory
- Collaboratory
- Grid (community)
- Network
- Portal
- Gateway
- Hub
- Virtual Research Environment (VRE)
- Cyberinfrastructure Collaborative
- Other?
Virtual Organizations: Goals

- To catalyze the development, implementation and evolution of a functionally-complete national cyberinfrastructure that integrates both physical and cyberinfrastructure assets and services.

- To promote and support the establishment of world-class VOs that are secure, efficient, reliable, accessible, usable, pervasive, persistent and interoperable, and that are able to exploit the full range of research and education tools available at any given time.

- To support the development of common cyberinfrastructure resources, services, and tools that enable the effective, efficient creation and operation of end-to-end cyberinfrastructure systems for and across all science and engineering fields, nationally and internationally.
• Distributed virtual organizations are **based upon CI** that provides flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions, and resources.

• **Resources and services** include HPC, data/information management, sensor-nets/observatories, linked through global networking and middleware, and accessed by people through web portals and workflow environments.

• Increasing numbers of **virtual organizations are required** by S&E research and education communities. Referred to by many names, e.g. *collaboratory, co-laboratory, grid, gateway, portal, hub, ....*

• **Challenges** being address include tools for more rapid building and ease of use, interoperability/middleware, high performance, end-to-end networking, and dynamic reconfiguration, social issues, assessment of impact, and economic and technical sustainability.
To be a genuinely competitive knowledge economy, Europe must be better

- in producing knowledge through research
- in diffusing it through education
- in applying it through innovation
e-Infrastructures in FP7 - strategy - Virtual Organizations (VO)

Bringing the best brains together
Sharing the best scientific resources

Weather Forecast VO
Biomedics VO
Astrophysics VO

Scientific Data
Sharing Scientific Resources
Communication Network

Producing the best science
VO-substrate: International R&E Networking
Learning supported by CI. (cyber-enabled learning).

Workforce development to create and use CI for S&E research and education.

Broadened participation: Exploit the new opportunities that cyberinfrastructure brings for ... people who, because of physical capabilities, location, or history, have been excluded from the frontiers of scientific and engineering research and education.

Explore CI support for integrated research and education.
OCI Program Officer Presentations

- **HPC Program** - Steve Meacham
- **CI TEAM** - Miriam Heller
- **Implementing the Strategic Vision for Digital Data** - Chris Greer
- **Middleware/Software** - Kevin Thompson
- **Towards Virtual Organization Initiatives** - Abhi Deshmukh
Other OCI Personnel

- Terry Langendoen, Program Officer
- Diana Rhoten, Consultant
- Abani Patra, Consultant
- Judy Hayden, Budget & Operations
- José Muñoz, Deputy Director
- Irene Lombardo, Staff Associate
- Priscilla Bezdek, Prog. & Tech. Specialist
- Deborah White-Wilkins, Sec. to Dir.
- Mary Daley, Sec. to Dep. Dir.
- Courtney Zajdel, STEP Student
Questions & Discussion
ACCI Task Force Groups

- **Portfolio**
  - Jim Bottum, Chair
  - Mark Ellisman
  - J. Tinsely Oden
  - Priscilla Nelson
  - Anne Trefethen
  - Jeron Tromp

- **Digital Data**
  - Henry Brady, Chair
  - Sara Graves
  - Robert Robbins
  - James Kinter
  - Terry Gaasterland
  - Gordon Shepard

- **CI and Competitiveness**
  - James Duderstadt, Chair
  - Stuart Feldman
  - Brian Bershad
  - Neerja Raman
  - Ann Gates
  - Brian Behlendorf

- **CI-Learning, Discovery, and Broadened Participation**
  - Steve Castillo, Chair
  - Diana Oblinger
  - Adebisi Oladipupo
  - David Oxtoby
  - N. Radhakrishnan
  - John Gage
  - Wesley Harris (liaison)