Maximizing Returns on National Investments Through Cyberinfrastructure

> Daniel E. Atkins U. S. National Science Foundation <u>datkins@nsf.gov</u> <u>www.nsf.gov/oci</u>

Workshop on Advanced Collaborative Environments and CERN Large Hadron Collider Users

> December 11-13, 2006 International Conference Centre Geneva, Switzerland





The University of Michigan Upper Atmospheric Research Collaboratory (UARC)

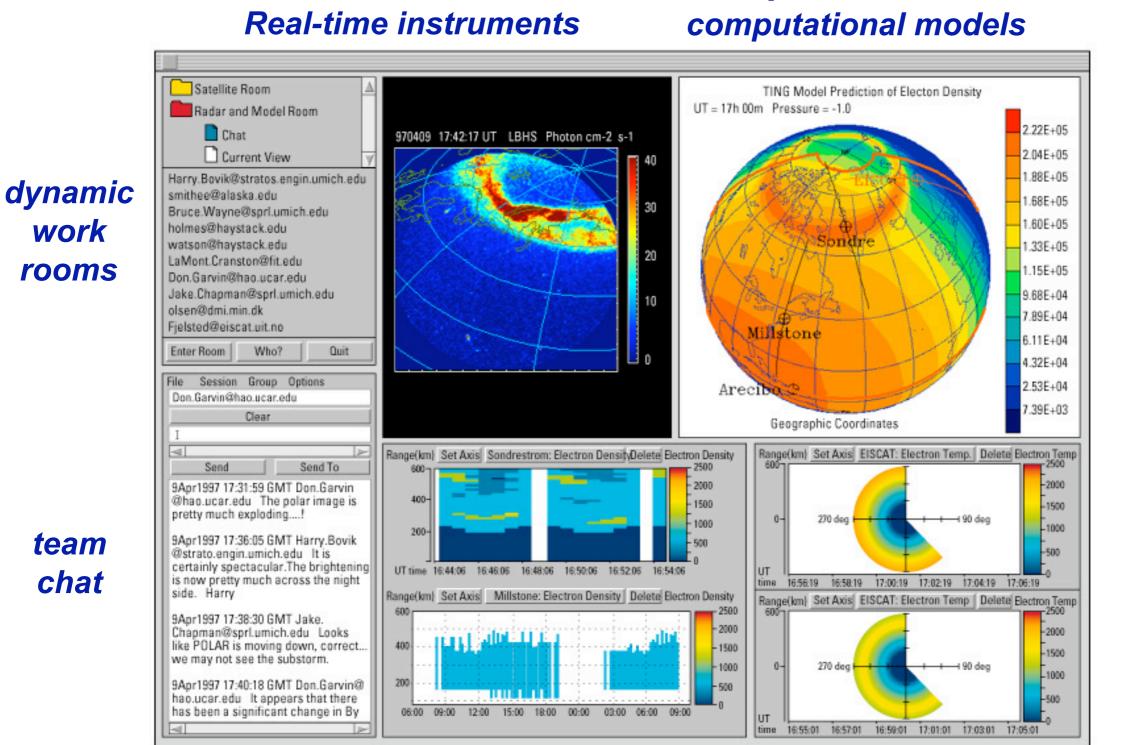


The Initial Facility at Sondrestrom, Greenland





UARC Interface (circa 1988)



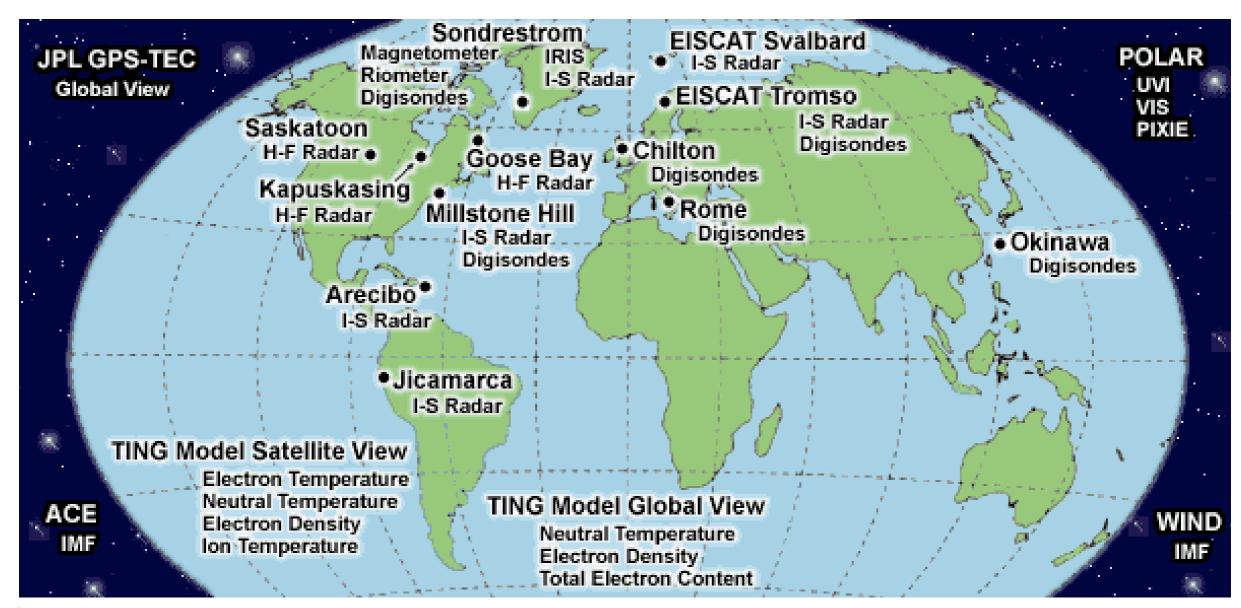
Session replay

Office of Cyberinfrastructure

Archival data

Journals

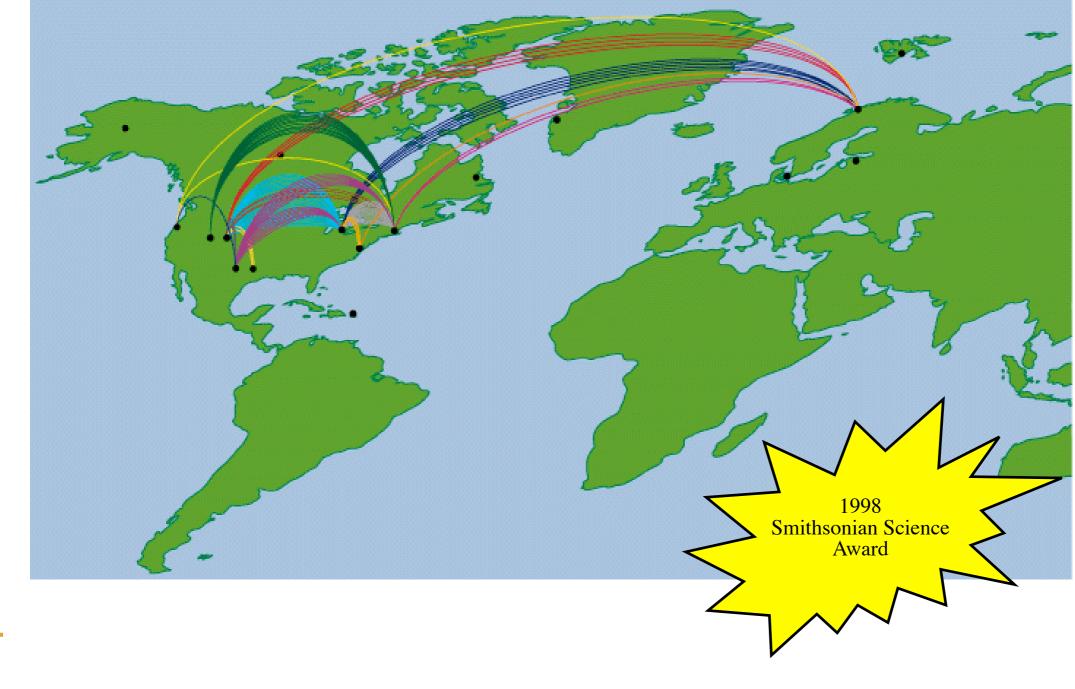
Evolved into a Network of Instruments (one global instrument)





UARC Patterns of Communication

Pattern of Communication, UARC Campaign, April 9, 1997



Office of Cyberinfrastructure

Vignettes from UARC/SPARC

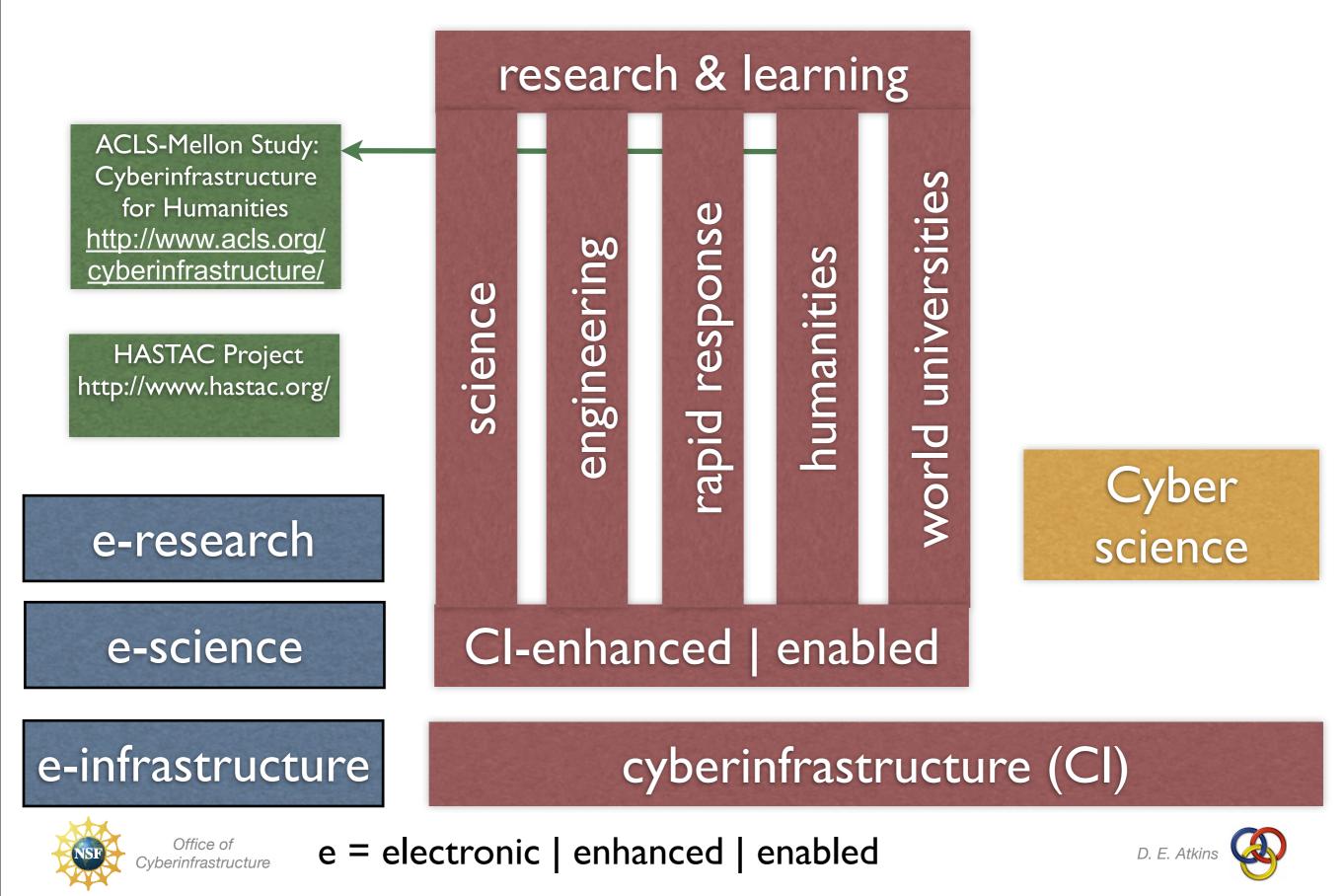
- Shared, tele-instruments, multiple eyes & complementary expertise.
- Rapid response, opportunistic campaigns.
- Isolated instruments became a global instrument chain.
- Cross-mentoring/training.
- New & earlier opportunities/exposure for grad students.
- Enhanced participation. Legitimate peripheral participation.



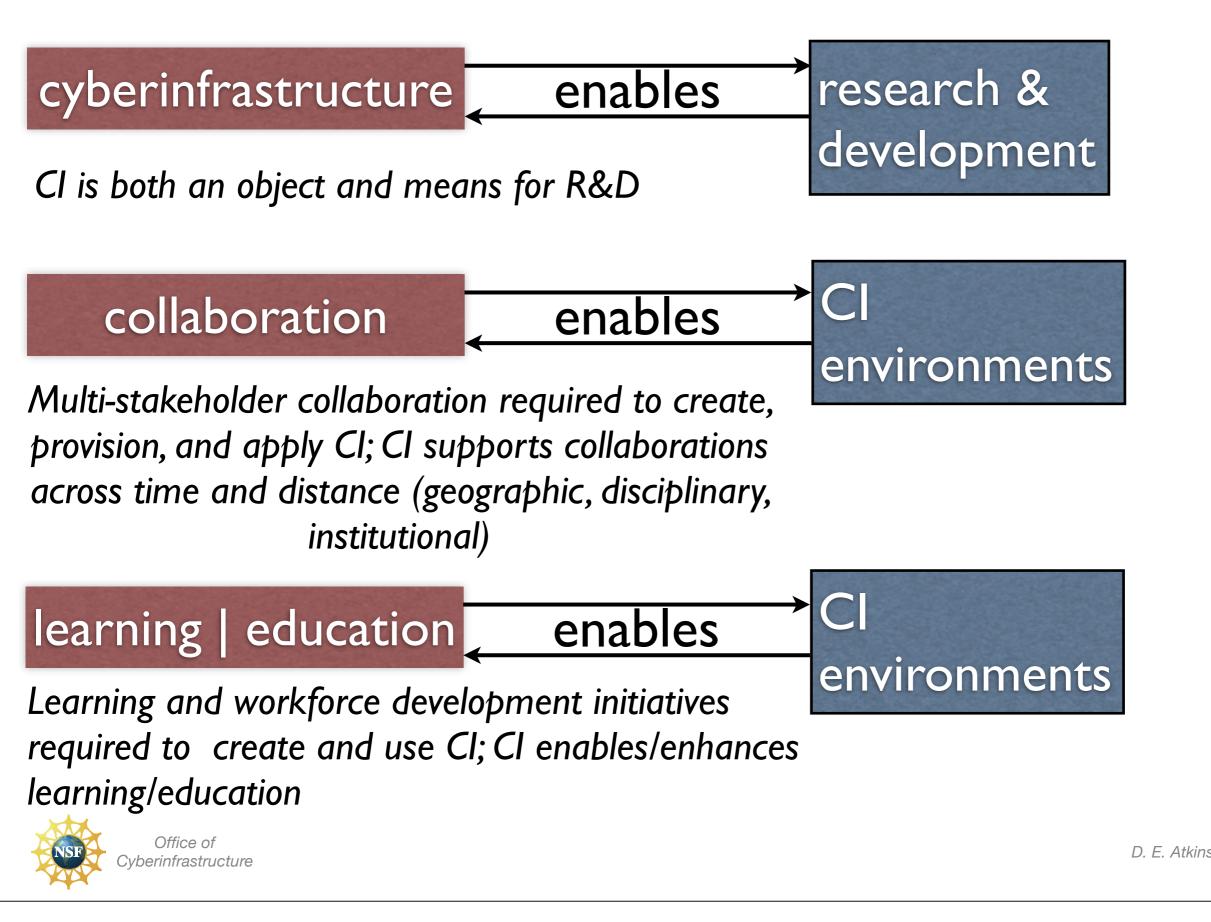
- Support for authentic, inquiry-based learning at UG and pre-college level.
- Distributed workshops for post-campaign data analysis.
- Session re-play for delayed participation and capture of process/provenance.
- Data-theory closure.
- A "living specification". Experiencing the new.



Current Nomenclature



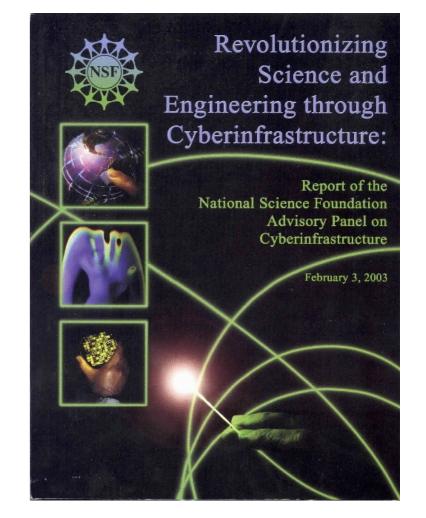
Dualities



NSF Blue Ribbon Advisory Panel on Cyberinfrastructure

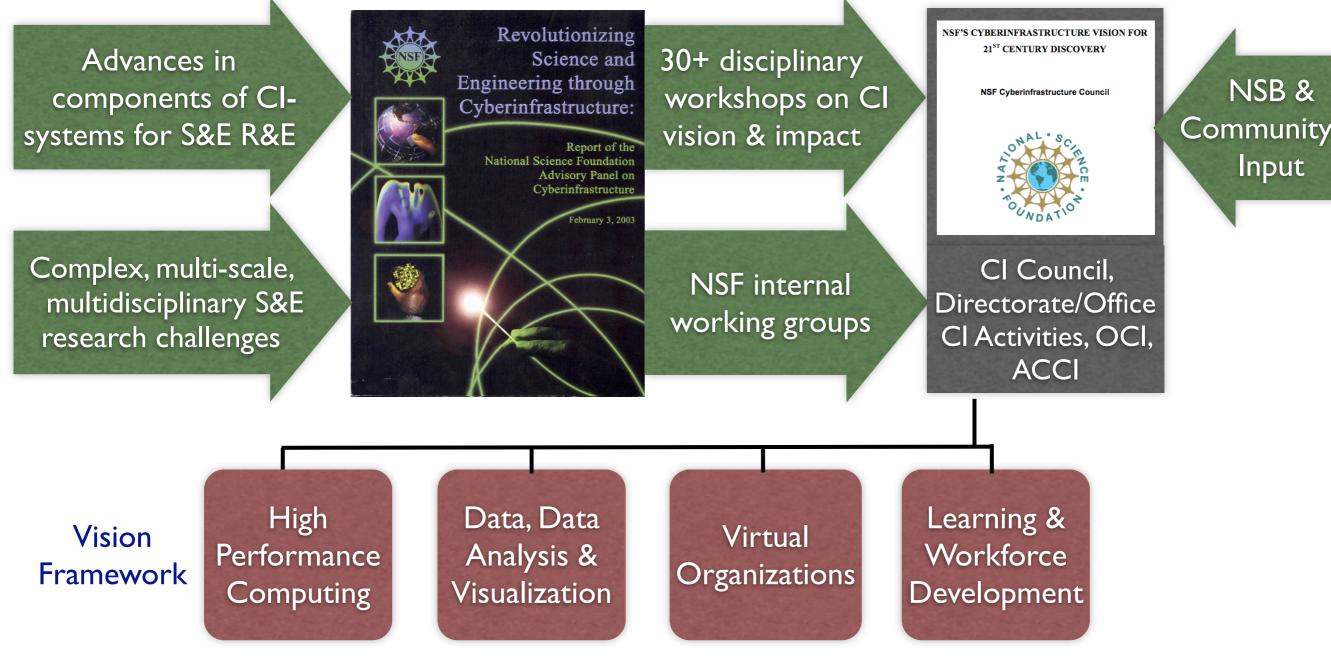
"a new age has dawned in scientific and engineering research, pushed by continuing progress in computing, information, and communication technology, and pulled by the expanding complexity, scope, and scale of today's challenges. The capacity of this technology has crossed thresholds that now make possible a comprehensive Text "cyberinfrastructure" on which to build new types of scientific and engineering knowledge environments and organizations and to pursue research in new ways and with increased efficacy."

http://www.nsf.gov/od/oci/reports/toc.jsp



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

Vision and Activities Based on Broad and Diverse Community Engagement



- •All directorates and offices support cyberinfrastructure.
- •Science-driven partnerships between creation, provisioning and use of CI



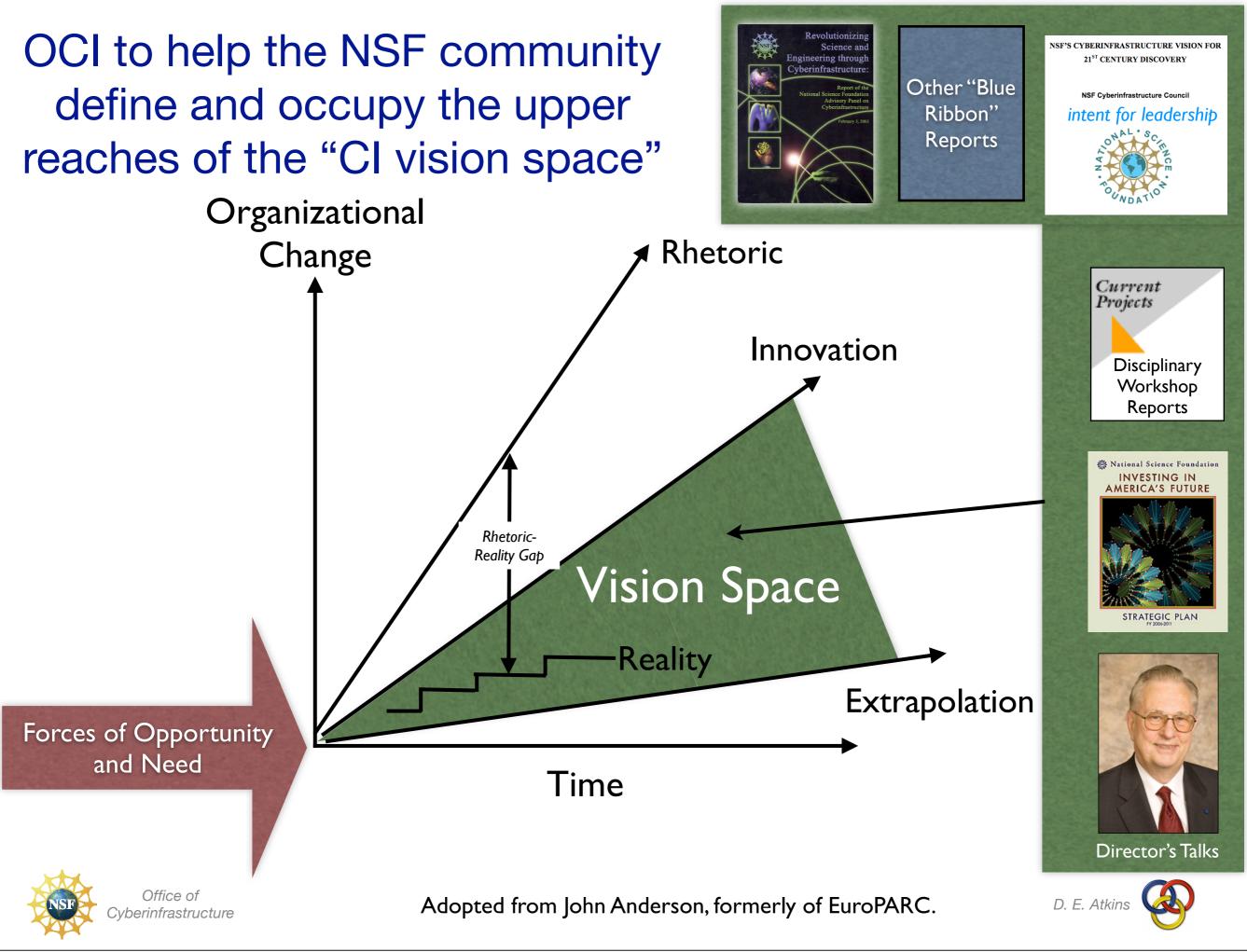
•Supports integrated research and education and broadened access and participation.

Cyberinfrastructure

www.nsf.gov/oci/

HOME FUNDING AWARD	DS DISCOVERIES M	NEWS PUBLICATIONS	STATISTICS	ABOUT	FastLane	
National Science	e Foundation		SEARCH			
NSF OFFICE OF			NSF Web S	ite	;	
Cyberinfrastruct	ure				٢	
OCI Home OCI Funding	OCI Awards	OCI Discoveries	OCI New	s	About OCI	
Cyberinfrastructure - stir advances in 21st century science and engineering	<u> </u>	2467				
About OCI Special Announcements			Quick Lin	iks		
View OCI Staff Directory					nd Workshops	
Search OCI Staff Directory	Career Opportunities - Dear Colleague Letter			Relating to Cyberinfrastructure and		
٢	The Office of Cyberinfrastructure (OCI) announces a nationwide search to fill a number of Program Director positions. For more information, click on the link http://www.nsf.gov/publications/vacancy.jsp?org=OCI&nsf_org=OCI			Its Impact		
Seneral Information About OCI				Publicati	ons <u>See A</u>	
Career Opportunities					Blue-Ribbon	
Advisory Committee	Petascale Acquisition Forum, Mar 24, '06				Advisory Panel on Cyberinfrastructure	
Budget Excerpt	NSF Invites Prospective Proposing Institutions and Vendors to a			Other Sit	e Features	
How to Decease Your Dropocal	Discussion of Plans for a Petascale HPC Acquisition		1	Special Re	ports	
	As indicated in the President's FY 2007 Budget Request,			Research	Overviews	
computing ()		the acquisition of a petascale high-performance HPC) system. Subject to the availability of funds, I		Multimedia	a Gallery	
Frequently Asked Questions	expects to begin funding the resulting multi-year acquis in FY07. The petascale HPC system to be acquired will			Classroom	Resources	
Other Types of Proposals	science and engineer		NSF-Wide	Investments		
Regional Grants Conferences	computationally cha					
How to Manage Your Award	HPC system vendors and potential resource provider organizations (organizations who, either separately or in collaboration with others, wish to propose to manage the development, deployment, and energiation of a potencial system on headle of the science and					
Srant Policy Manual						
Grant General Conditions	operation of a petascale system on behalf of the science and engineering research community) are invited to meet with each					
Cooperative Agreement Conditions	other and with NSF staff to discuss the time-line and strategy for this petascale system acquisition on Friday, March 24, 2006, from					
Special Conditions	9:00a.m 11:00a.r	undation, 4201				
Federal Demonstration Partnership	Wilson Blvd., Arlingt this meeting should					
Policy Office Website	HPC-Input@nsf.gov,					





Some Science Drivers

- Inherent complexity and multi-scale nature of todays frontier science challenges.
- Requirement for multi-disciplinary, multiinvestigator, 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.

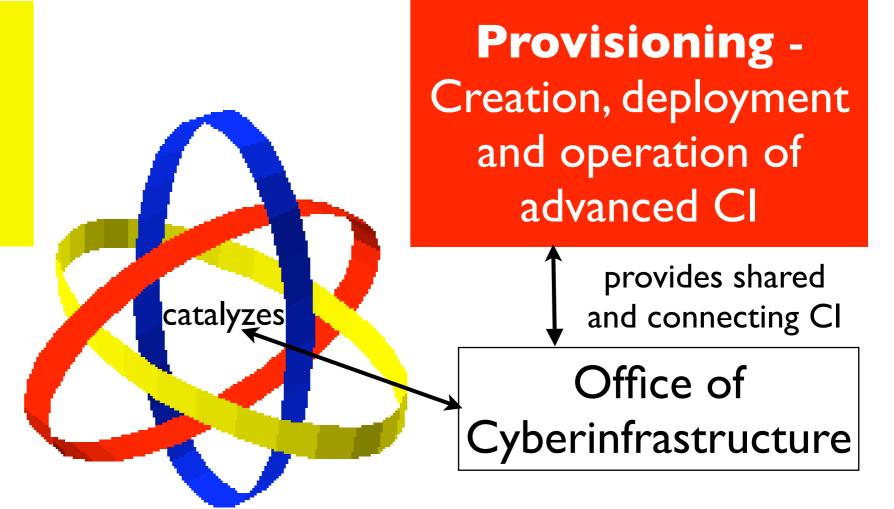




Achieving the NSF CI (e-science) Vision requires synergy between 3 types of activities

Transformative Application - to enhance discovery & learning

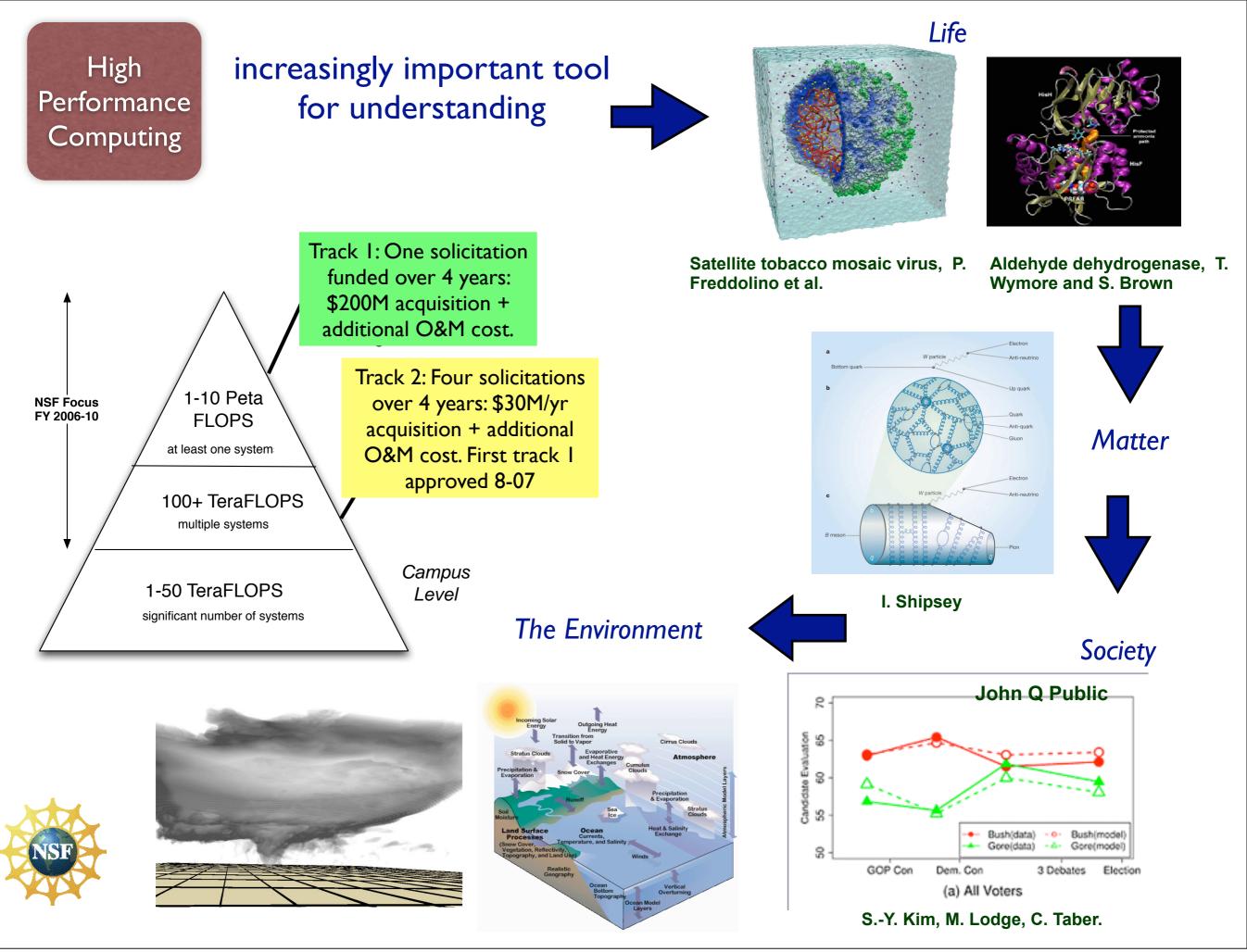
Borromean Ring: The three rings taken together are inseparable, but remove any one ring and the other two fall apart. See <u>www.liv.ac.uk/</u> ~spmr02/rings/



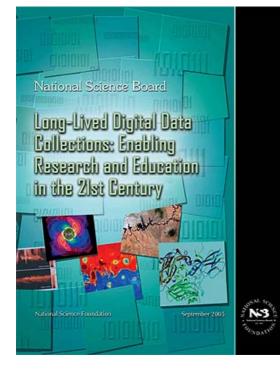
R&D to enhance technical and social effectiveness of future CI environments



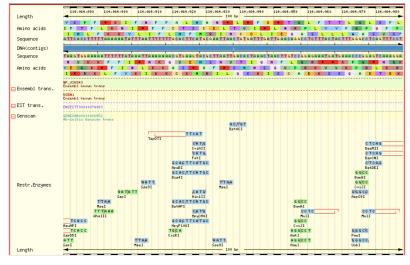
Office of Cyberinfrastructure

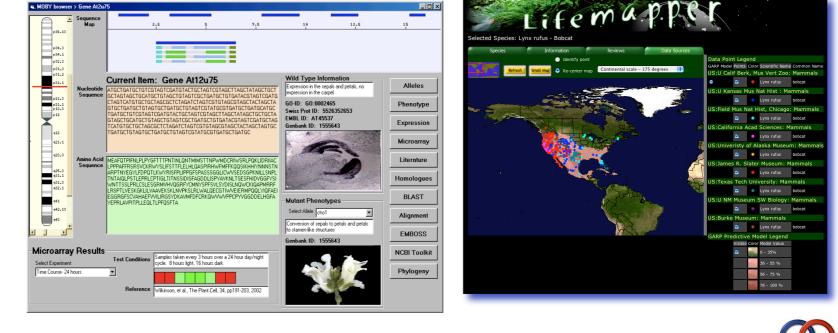


Data, Data Analysis & Visualization



- 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







New Report: To Stand the Test of Time

To Stand the Test of Time

Long-term Stewardship of Digital Data Sets in Science and Engineering

A Report to the National Science Foundation from the ARL Workshop on New Collaborative Relationships: The Role of Academic Libraries in the Digital Data Universe

September 26–27, 2006 Arlington, VA

> "Nature, to be commanded, must be obeyed." Attributed to Francis Bacon (1561–1626) Novum Organum, bk.1, aph. 129 (1620)



Available online at <u>http://</u> www.arl.org/info/events/ digdatarpt.pdf

Lead NSF Program Officer for Data Initiatives, Chris Greer is at this meeting.



Office of Cyberinfrastructure



Virtual Organizations



NVO







iVDgL







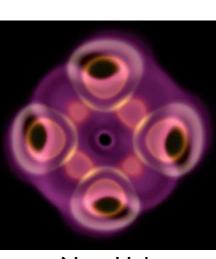
Open Science Grid

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

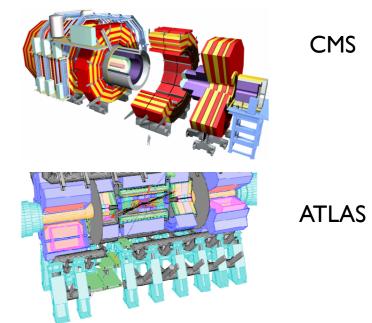
- To promote and support the establishment of world-class VOs that are secure, efficient, reliable, accessible, usable, pervasive, persistent and interoperable, and that are able to exploit the full range of research and education tools available at any given time
- To support the development of common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



NEES



NanoHub



Instances of Virtual Organizations (VOs)

People*

People*

People*

Interfaces for interaction, workflow, visualization and collaboration for distributed teams in domain/project specific and potentially functionally-complete VOs.

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

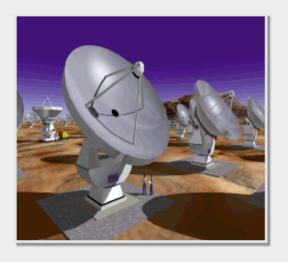


Office of *** P** Cvberinfrastructure

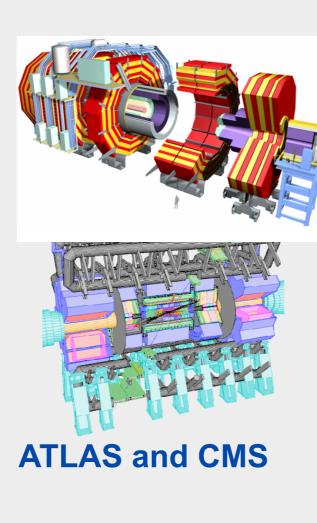
* People engaged in discovery and learning as individuals and in teams. \Box

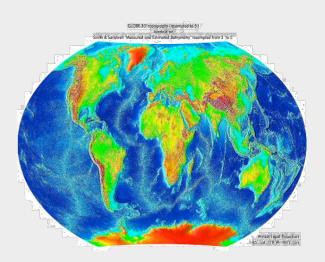


CI/VO Enabled Science



NVO and ALMA

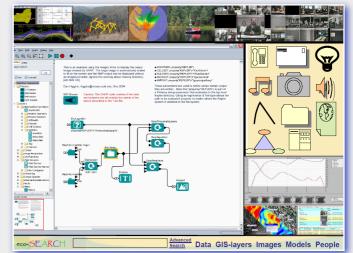




Climate Change



LIGO



NEON

The number of international-scale projects is growing rapidly!





Alternate Generic Names for Virtual Organizations

- Distributed Knowledge Communities
- Virtually-augmented Communities
- Participation Environments built upon Participation Architectures
- Learning and Discovery Ecosystems





Virtually Augmented Organizations offer additional modes of interaction between People, Information, and **Facilities** Time Same Different (synchronous) (asynchronous) DT-SP **ST-SP** Same **Geographic Place** P: Physical mtgs **P**: Shared **I**: Print-on-paper notebook books, journals : Library reserves Physical + F: Physical labs, **F**:Time-shared Virtual. studios, shops physical labs, ... Not Physical Different **ST-DP DT-DP** vs.Virtual **P**:AV conference P: Email :Web search I: Knowbots **F**: Online **F**:Autonomous observatories instruments

P: people, I: information, F: facilities, instruments

Need a socio-technical approach: Realizing the potential of e- | cyber science to support effective VOs requires attention to *institutional infrastructure*

- Need more than good ICT systems and tools for individuals and organizations. (Technology determinism alone is not enough).
- No less important is the *institutional contexts* (i.e. norms of practice and rules) to facilitate collaboration within science and technical research communities.
- The institutional and organizational environment of e-science encompasses a wide and diverse array of interrelated social, economic, and legal factors that
 - create incentives for, and constraints upon individual and collective action; and
 - thereby shape the production, utilization, consumption, and governance of e-science capabilities and products.



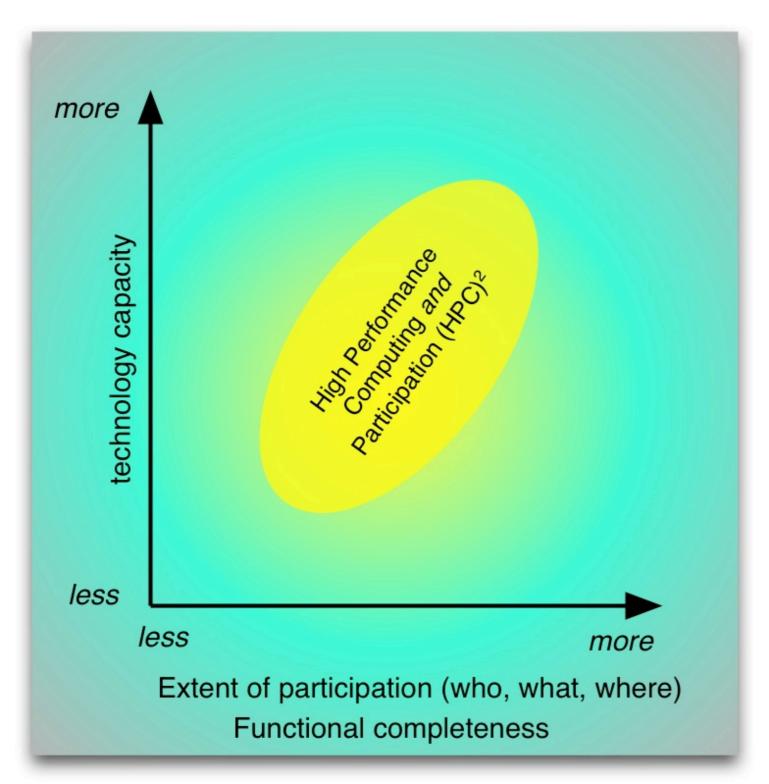
Office of

vberinfrastructure

Adapted from Paul David, see <u>www.oii.ox.ac.uk/resources/publications/RR2.pdf</u>



Optimizing CI Investment







The Grid Movement Search The Site:



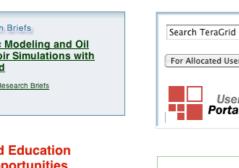
Science Gateways

TeraGrid Related Education **Resources & Opportunities**

The Science Gateways program is designed to enable entire communities of users with a common scientific goal to use the TeraGrid through a common interface. More on Science Gateways

ASTA

The Advanced Support for TeraGrid Applications (ASTA) Program aims to help ensure that the resources of TeraGrid are optimally utilized for important scientific coveries and technological innovation.



Home

TeraGrid offer a variety of workshops, institutes, seminars and on-line learning resources to engage the community in making effective use of TeraGrid resources. A list of these learning opportunities across all of the Resource Provider sites is posted on the Education, Outreach and Training web pages at http://www.teragrid.org/eot/workshops.php

We would like to bring your attention to the following workshops:

October 16-17, 2006: Introduction to & Optimization for SDSC Sytems

October 29, 2006: Remote/Collaborative TeraScale Visualization on the TeraGrid TeraGrid is Voted GRIDtoday Top

eeee

Enabling Grids

for E-sciencE

EGEE PROJECT

EGEE & INDUSTRY

USERS AND EGEE

WORLDWIDE GRID

EGEE EVENTS

NEWSROOM

Research Grid





09/28/2006: Texas Advanced Computing Center receives \$59 million high-performance computing award from National Science Foundation

09/22/2006: Call for Papers - The 8th LCI International Conference on High-Performance Clustered Computing

09/20/2006: RENCI Bioportal Enhanced with First Workflow

09/18/2006: Katrina: After the Storm – Civic Engagement Through Arts. Humanities and Technology -September 28-30, 2006 09/18/2006: Call For Participation - What To Do With A Million Books: Chicago



About the OSG

Getting Started

News and Events

Contacts

OSG at Work

OSG at Work

Monitoring

Security

Support

Tools for Collaborators

Virtual Data Toolkit

Document Database

Logos and Templates

Science on the OSG

Home



Science on the Open Science Grid



search ...

Simulating Supersymmetry with ATLAS

One of the discoveries eagerly anticipated by particle physicists working on the world's next particle collider is that of supersymmetry, a theoretical lost symmetry of nature. Physicist Sanjay Padhi from the University of Wisconsin-Madison has used Open Science Grid resources to show that there is a good possibility of discovering supersymmetry in the first few months of operation of the new collider, if the new symmetry exists in nature. Read more ...

OSG NEWS

- <u>Open Science Grid Receives \$30 Million Award to</u> Empower Scientific Collaboration and Computation
- There will be a joint EGEE/OSG session on Security at the EGEE Conference, September 25-29, 2006.
- September 11, 2006 article in GRIDToday: <u>Running an</u> Effective Distributed Facilit
- Press Release, September 7, 2006: DOE Announces \$60 Million in Projects to Accelerate Scientific Discovery through Advanced Computing.

Search

Register as a Community Member | Log-in | RSS

Welcome to EGEE (Enabling Grids for E-sciencE).

The Enabling Grids for E-sciencE project brings together scientists and engineers from more than 90 institutions in 32 countries world-wide to provide a seamless Grid infrastructure for e-Science that is available to scientists 24 hours-a-day. Conceived from the start as a four-year project, the second two-year phase started on 1 April 2006, and is funded by the European Commission.

Expanding from originally two scientific fields, high energy physics and life sciences, EGEE now integrates applications from many other scientific fields, ranging from geology to computational chemistry. Generally, the EGEE Grid infrastructure is ideal for any scientific research especially where the time and resources needed for running the applications are considered impractical when using traditional IT infrastructures.

Latest News First EELA Bulletin - EELA News -

Year 1, Issue 1 now available

. Read more

Try the GRID

Click here

Become a User

Want to become a user of the EGEE Grid? Click here

Collaborating Projects

E. Atkins If your project is related to EGEE, please register it here



Search OSG at Work:

View Live Grid Status

search...

Offi Svberinfr



Let's look at a few real example Grid Science Gateways

(about a dozen ... many more exist!)

These example slides courtesy of D. Gannon

NEESGrid

G · D

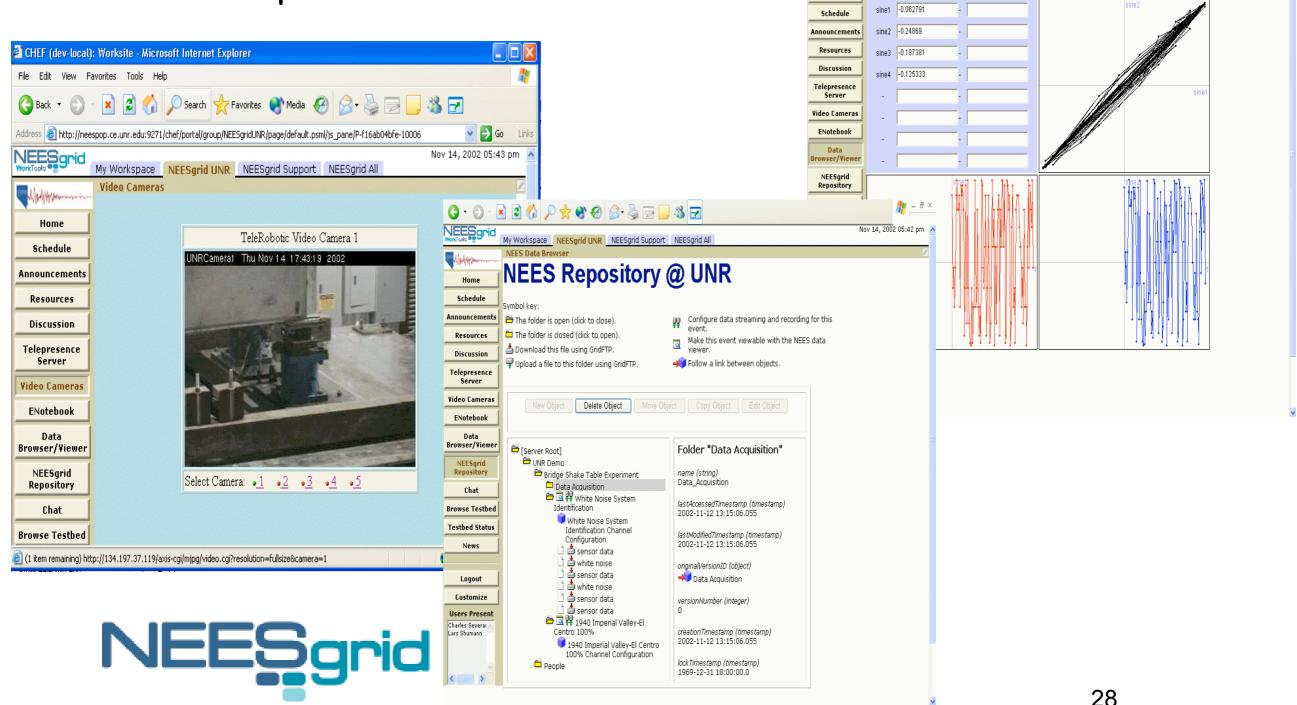
Home

- 🖹 🙆 🏠 🔎 👷 🛞 🏵 🍰 🖂 📃 🦓 🔂

Event: "core: ex2 sine1-4"

My Workspace NEESgrid UNR NEESgrid Support NEESgrid All

Realtime access to earthquake Shake table experiments at remote sites.



Į _ ð ×

Nov 14, 2002 05:40 pm

BIRN – Biomedical Information

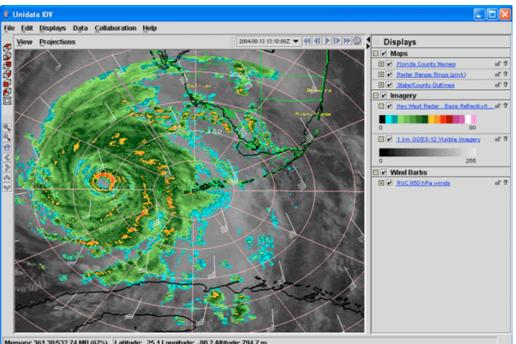


Mesoscale Meteorology

NSF LEAD project - making the tools that are needed to make accurate predictions of tornados and hurricanes.

- Data exploration and Grid workflow





S WELCOME TO THE LEAD PORTAL



Linked Environments for Atmospheric Discovery (LEAD) makes meteorological data, forecast models, and analysis and visualization tools available to anyone who wants to interactively explore the weather as it evolves. The LEAD Portal brings together all the necessary resources at one convenient access point ... read more

> FEATURES FOR ANYONE INTERESTED IN THE WEATHER

Researchers	With university, government, or industry affiliations	GET FEATURES
Educators	At college and university level, high school, or middle schools	GET FEATURES
Students	At graduate, undergraduate, middle and high school levels	GET FEATURES
Visitors	Newcomers and the curious	GET FEATURES

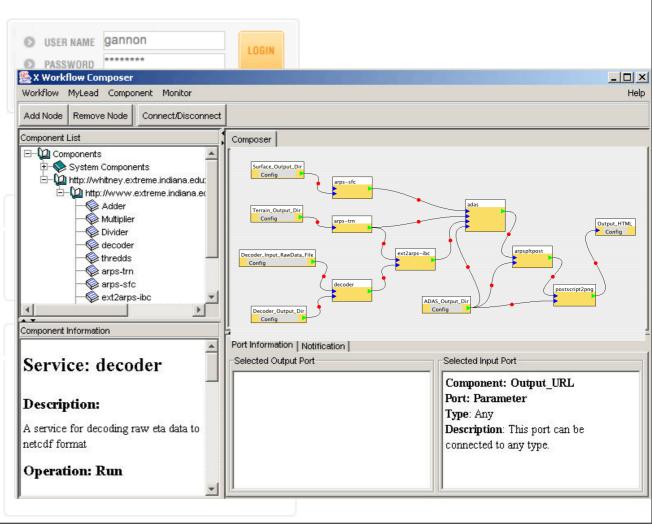
> POPULAR TOOLS



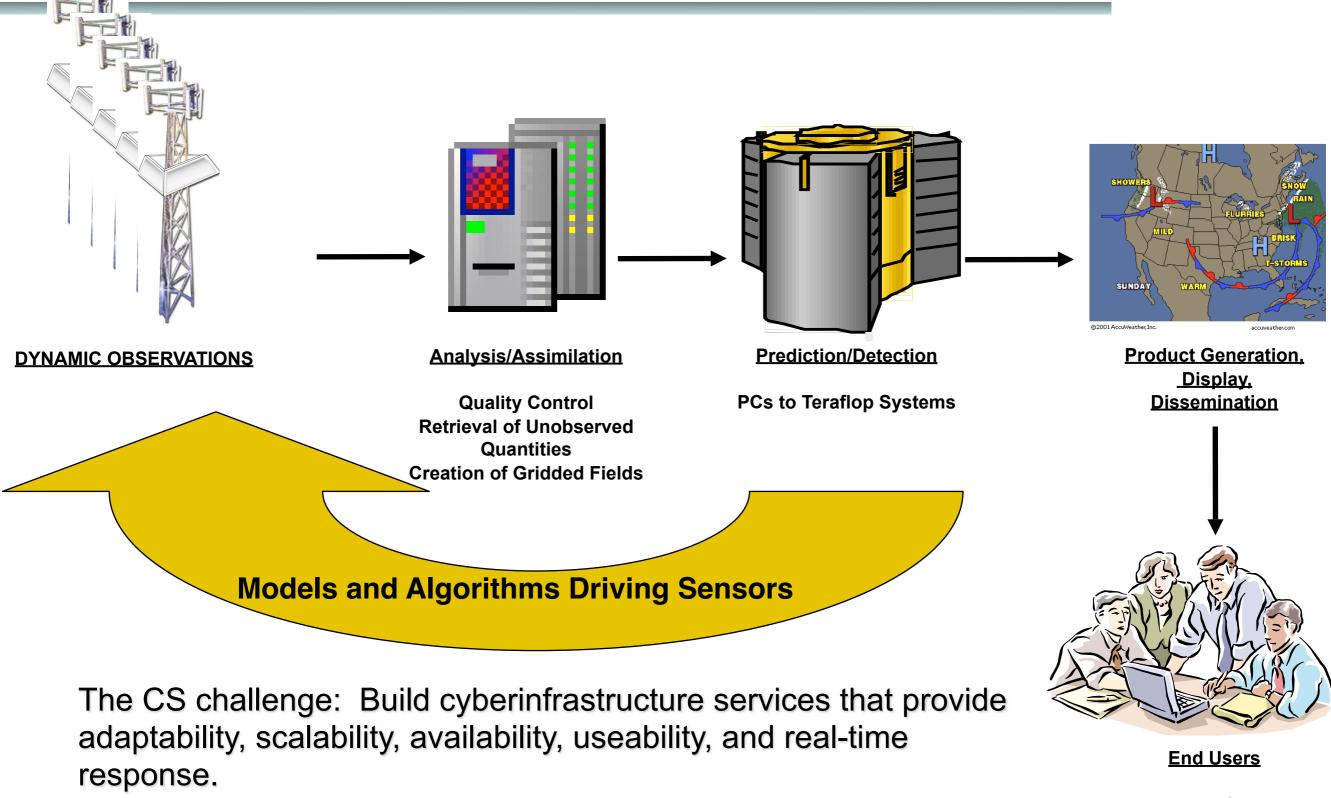
Make a Forecast or Analysis
Experiment Builder







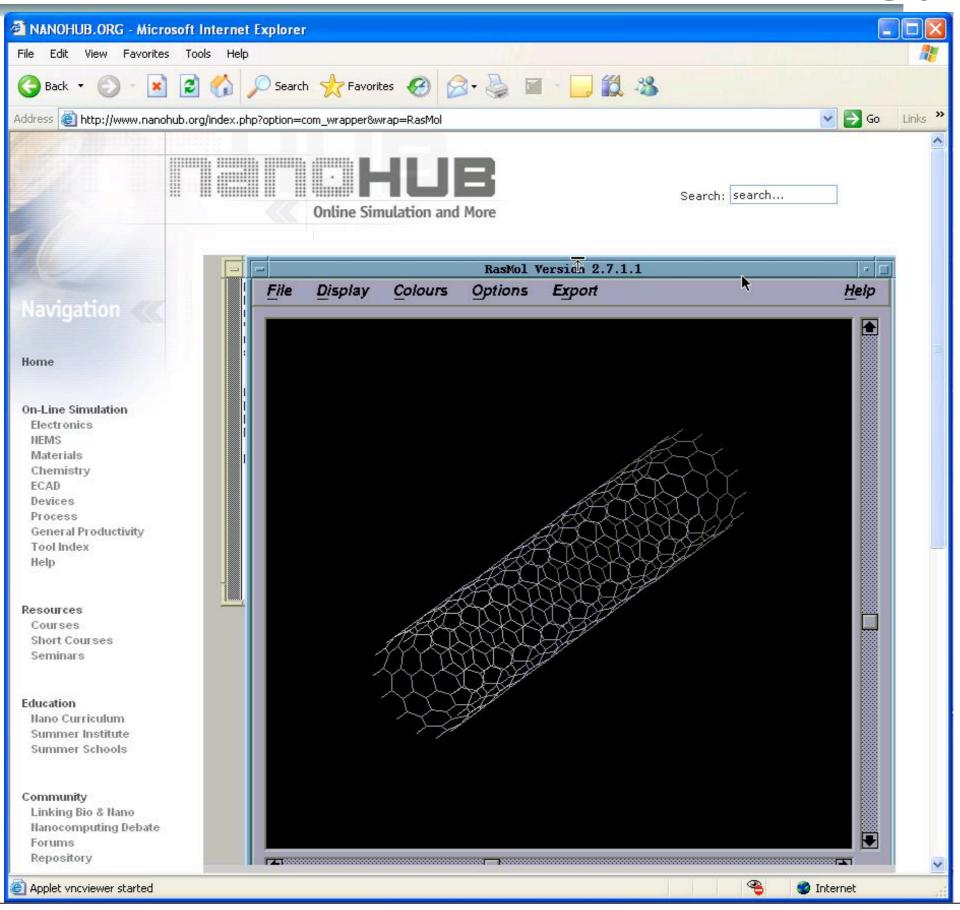
The LEAD Vision: Adaptive Cyberinfrastructure



NWS Private Companies Students

From D. Gannon

Nanohub - nanotechnology



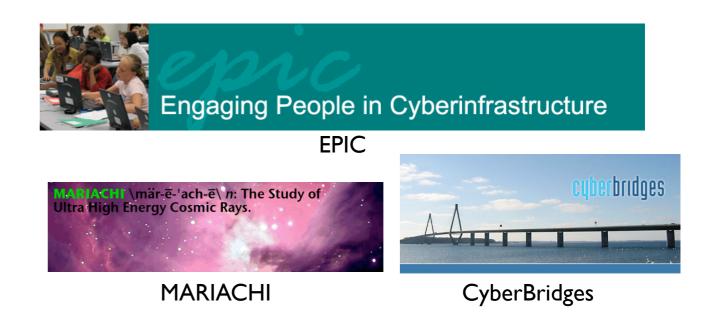
Learning & Workforce Development

- 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.





BIOINFORMATICS CI INSTITUTE



FP7 - Putting the knowledge triangle at work

To be a genuinely competitive knowledge economy, Europe must be better

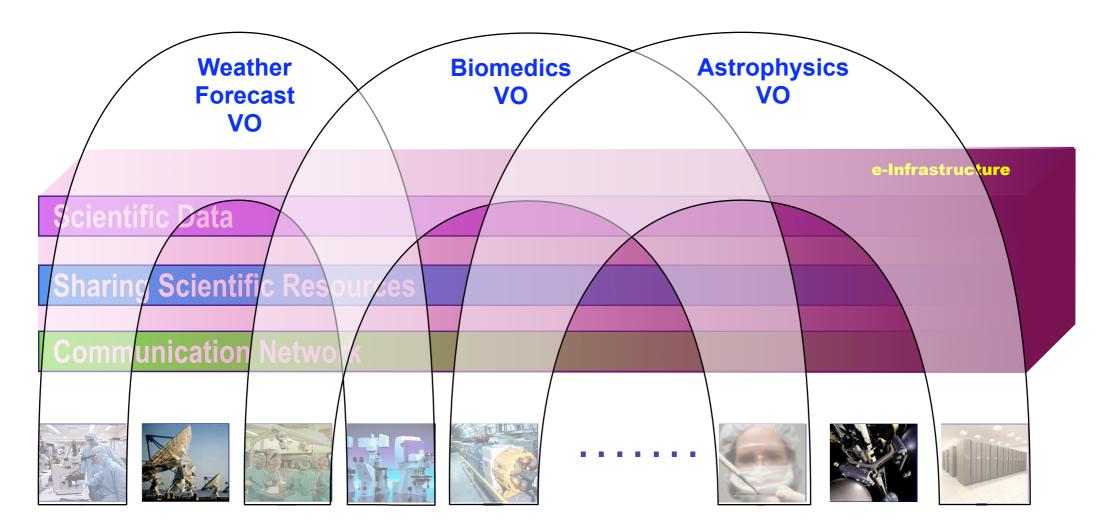
- In producing knowledge through research
- In diffusing it through education
- In applying it through innovation



e-Infrastructures in FP7 - strategy - Virtual Organizations (VO)

Bringing the best brains together

Sharing the best scientific resources

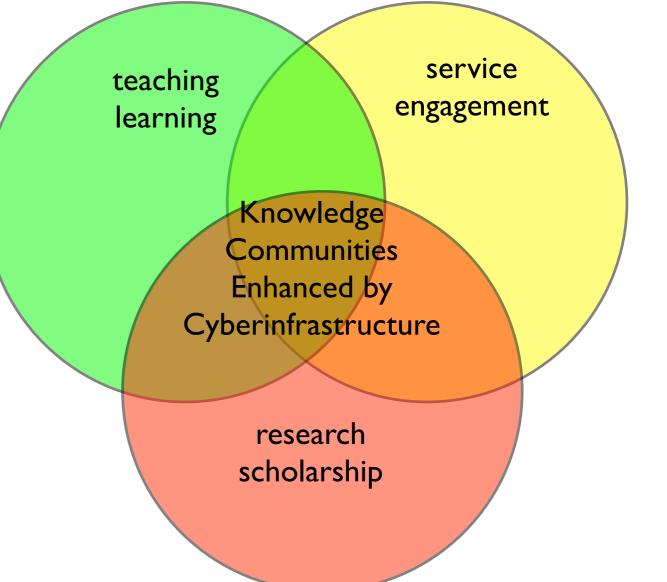


Producing the best science



And while we are at it...

- can we create Cl platforms in support of research, learning, and societal engagement in ways that exploit complementarity between them?
 - "Better than being there."
 - Pasteur's Quadrant research
 - Ubiquitous learning ecologies



The opportunity is more than distributed computing and data storage. It is about using ICT/CI as an enhancement platform for knowledgebased communities to learn, discover, and engage in more "barrier free" ways.



The Meta University A Personal View

What we are observing is the early emergence of a Meta University -- a transcendent, accessible, empowering, dynamic, communally-constructed framework of open materials and platforms on which much of higher education worldwide can be constructed or enhanced.

From Charles M. Vest, President Emeritus of MIT 37

The Meta University

- Will enable -- not replace -- residential campuses
- Will bring cost efficiencies to institutions through shared development;
- Will be adaptable -- not prescriptive;
- Will serve both teachers and learners;
- Will speed the propagation of high-quality education and scholarship;
- Will build capacity for economic development;
- Will build bridges across cultures and political boundaries; and
- Will be particularly important to the developing world.

From Charles M. Vest, President Emeritus of MIT 38

Something to ponder.....

Global Needs

Half of the world's population is under 20 years old.

Today, there are over 30 million people who are fully qualified to enter a university, but there is no place available. This number will grow to over 100 million during the next decade.

To meet the staggering global demand for advanced education, a major university would need to be created every week.

"In most of the world, higher education is mired in a crisis of access, cost, and flexibility. The dominant forms of higher education in developed nations—campus based, high cost, limited use of technology—seem ill-suited to addressing global education needs of the billions of young people who will require it in the decades ahead."

Sir John Daniels

From James J. Duderstadt, President Emeritus, U. of Michigan

International e-science (cyberinfrastructure) movement





Openness Movement

Open Courseware Consortium

Open Software: middleware,toolkits, apps, services Creative Commons Licensing

Open scholarly communication

Open Repositories, Digital Libraries, Archives

Web 2.0 Collaborative contribution environments

Openness Movement

Culture of Contribution





Concepts, Visions of the Future of Learning & Discovery

Meta-University, Universities in and of the World

> JSB Culture of Learning

National Academies ITFRU Studies

NSF CIVision,

ACLS CI &

Humanities

OECD and other international visioning activities

OTHER?

Studies, Concepts, Visions

Office of Cyberinfrastructure



Multiple Stakeholders

Federal and State Government	Philanthropy	Industry	Stakeholders
Academia	Civil Society	International version of all these	





Gaming, collaboration, virtual, immersive and augmented reality





Visions, Frameworks

Openness Movement

Stakeholders

Opportunity for Innovation and Leadership in Learning and Discovery

CI Movement

Ganning....

Critical Societal Needs





Office of Cyberinfrastructure

Questions

- The LHC community has made massive investments in high-capacity, globally-distributed grid computing to support the data analysis mission. But to what it extent is it leveraging this investment to reduce the barriers of geography for the thousands of people involved in LHC projects?
 - Being at CERN when on your campus. Being on your campus when at CERN.
 - Broadened participation.
 - Research and learning, including participation by students, citizen scientists, general public.





- Could at least a subset (microcosm?) of the LHC community develop a prototype of a "global meta university;" an open ecosystem of discovery and learning.
 - Primary goal: more discovery and learning in the field of physics.
 - Secondary goal a large scale prototype of an open, global scale, "meta-university"?
 - Possibility of broadened set of sponsors.
- Why not leadership in new CI-enabled platforms for participation in learning and discovery from the particle physics community?





- Perhaps such global-scale CI-enabled platforms for participation will be essential for success with the next generation linear accelerator.
- Furthermore perhaps eventually it can mitigate against the growing concern about the expense of creation, operation, and sustaining very expensive research infrastructure, and the attendant erosion of research funding for the use of such facilities.





Questions & Discussion



