

## APPENDIX A: ACRONYMS

2MASS	Two Micron All Sky Survey
CAIDA	Cooperative Association for Internet Data Analysis
CALIT2	California Institute for Telecommunications and Information Technology
CAVE	Cave Automatic Virtual Environment
CCSDS	Consultative Committee for Space Data Standards
CHARMM	Chemistry at Harvard Macromolecular Mechanics
CI	Cyberinfrastructure
CIC	Cyberinfrastructure Council
CIO	Chief Information Officer
CMS	Compact Muon Solenoid
CODATA	Committee on Data for Science and Technology
CPU	Central Processing Unit
CSNET	Computer Science Network
CT	Computed Tomography
DANSE	Data Analysis for Neutron Scattering Experiments
DARPA	Defense Advanced Research Projects Agency
DNA	Deoxyribonucleic Acid
DOD	Department of Defense
DOE	Department of Energy
ETF	Extensible Terascale Facility
FLOPS	Floating point operations/sec
GPS	Global Positioning System
HBCU	Historically Black Colleges and Universities
HPC	High Performance Computing
HPCMOD	DOD's High-Performance Computing Modernization program
HPCS	DARPA's High Productivity Computing Systems program
HPCC	High-Performance Computing and Communications
HPWREN	High Performance Wireless Research and Education Network
GEON	Geosciences Network
GLORIAD	Global Ring Network for Advanced Applications Development
GriPhyN	Grid Physics Network
ICPSR	Inter-university Consortium for Political and Social Research
ICSU	International Council for Science
ICSTI	International Council for Scientific and Technical Information
IRIS	Incorporated Research Institutions for Seismology
ISO	International Organization of Standardization
IT	Information Technology
ITR	Information Technology Research
IVDGL	International Virtual Data Grid Laboratory
LIGO	Laser Interferometer Gravitational Wave Observatory
LTER	Long Term Ecological Research
LWD	Learning and Workforce Development
MPI	Message Passing Interface
MREFC	Major Research Equipment and Facilities Construction
NARA	National Archives and Records Administration
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research

*The simulation on the opposite page, run by Louisiana State University astrophysicists on the NCSA Tungsten system, shows two stars in a binary system during a late phase of their evolution. The more massive star (top) has transferred a significant fraction of its mass to its companion and the two stars are undergoing a catastrophic merger.*

NCSA	National Center for Supercomputing Applications
NEES	Network for Earthquake Engineering Simulation
NEON	National Ecological Observatory Network
NIH	National Institutes of Health
NITRD	Networking and Information Technology Research and Development
NNSA	National Nuclear Security Administration
NREL	National Renewable Energy Laboratory
NRL	Naval Research Laboratory
NSFNET	NSF Network
NRC	National Research Council
NSB	National Science Board
NSF	National Science Foundation
OAIS	Open Archival Information System
OOI	Ocean Observatories Initiative
OS	Operating System
PACI	Partnership for Advanced Computational Infrastructure
PDB	Protein Data Bank
PITAC	President's Information Technology Advisory Committee
RCSB	Research Collaboratory for Structural Bioinformatics
RLG	Research Library Group
RNA	Ribonucleic Acid
SciDAC	Scientific Discovery through Advanced Computing
SDSC	San Diego Supercomputer Center
SNS	Spallation Neutron Source
SSP	Software Services Provider
TESS	Time-sharing Experiments for the Social Sciences
TFLOPS	Teraflops: Trillion floating point operations/sec
USNC/ CODATA	U.S. National Committee for CODATA
vBNS	very high speed Backbone Network Service
VO	Virtual Organization
WDC	World Data Center

## APPENDIX B: REPRESENTATIVE REPORTS AND WORKSHOPS

Building a Cyberinfrastructure for the Biological Sciences; workshop held July 14-15, 2003; information available at [http://research.calit2.net/cibio/archived/CIBIO\\_FINAL.pdf](http://research.calit2.net/cibio/archived/CIBIO_FINAL.pdf) and <http://research.calit2.net/cibio/report.htm>

CHE Cyber Chemistry Workshop; workshop held October 3-5, 2004; information available at [http://bioeng.berkeley.edu/faculty/cyber\\_workshop/](http://bioeng.berkeley.edu/faculty/cyber_workshop/)

Commission on Cyberinfrastructure for the Humanities and Social Sciences; sponsored by the American Council of Learned Societies; seven public information-gathering events held in 2004; report in preparation; information available at <http://www.acls.org/cyberinfrastructure/cyber.htm>

Community Climate System Model Strategic Business Plan (2003), June 2003, 28 pp; information available at <http://www.ccsm.ucar.edu/management/busplan2004-2008.pdf>

Community Climate System Model Science Plan 2004-2008 (2003), June 2003, 76 pp; information available at <http://www.ccsm.ucar.edu/management/sciplan2004-2008.pdf>

Computation as a Tool for Discovery in Physics; report by the Steering Committee on Computational Physics; information available at <http://www.nsf.gov/pubs/2002/nsf02176/start.htm>

Cyberinfrastructure for the Atmospheric Sciences in the 21st Century; workshop held June 2004; information available at [http://gladiator.ncsa.uiuc.edu/PDFs/nsf/cyrdas\\_report\\_final.pdf](http://gladiator.ncsa.uiuc.edu/PDFs/nsf/cyrdas_report_final.pdf)

Cyberinfrastructure for Engineering Research and Education; workshop held June 5 – 6, 2003; information available at <http://www.nsf.gov/eng/general/Workshop/cyberinfrastructure/index.jsp>

Cyberinfrastructure for Environmental Research and Education (2003); workshop held October 30 – November 1, 2002; information available at <http://www.ncar.ucar.edu/cyber/cyberreport.pdf>

Cyberinfrastructure (CI) for the Integrated Solid Earth Sciences (ISES) (June 2003); work-

shop held on March 28-29, 2003; June 2003; information available at <http://www.paleostrat.org/Documents/ises-ci%202003.pdf>

Cyberinfrastructure and the Next Wave of Collaboration; D.E. Atkins, Keynote for EDU-CAUSE, held April 2005; information available at <http://www.educause2005.auckland.ac.nz/interactive/presentations/Atkins.pdf>

Cyberinfrastructure needs for environmental observatories; information available at <http://www.orionprogram.org/office/NSFCyberWkshp.html>

Cyberinfrastructure Research for Homeland Security; February 26-27, 2003; information available at [web.calit2.net/RiskReduction/crhdraft.pdf](http://web.calit2.net/RiskReduction/crhdraft.pdf)

Cyberlearning Workshop Series; workshops held Fall 2004 – Spring 2005 by the Computing Research Association (CRA) and the International Society of the Learning Sciences (ISLS); information available at <http://www.cra.org/Activities/workshops/cyberlearning>

Data Management for Marine Geology and Geophysics: Tools for Archiving, Analysis, and Visualization (2001); information available at [http://hummm.who.edu/DBMWorkshop/data\\_mgt\\_report.hi.pdf](http://hummm.who.edu/DBMWorkshop/data_mgt_report.hi.pdf)

Environmental Cyberinfrastructure Needs For Distributed Sensor Networks; workshop held August 12-14, 2003; information available at [http://www.lternet.edu/sensor\\_report](http://www.lternet.edu/sensor_report)

Establishing a Petascale Collaboratory for the Geosciences: Scientific Frontiers Report, July 2005, information available at [http://www.geoprose.com/projects/petascale\\_science.html](http://www.geoprose.com/projects/petascale_science.html)

EXchanging CyberInfrastructure Themes in Engineering Design; Report for NSF EXCITED workshop, held February 28 – March 1, 2005; information available at <http://www.mne.psu.edu/simpson/NSF/EXCITED/>

Federal Plan for High-End Computing (2004); 72 pp; information available at [http://www.ostp.gov/nstc/html/HECRTF-FINAL\\_051004.pdf](http://www.ostp.gov/nstc/html/HECRTF-FINAL_051004.pdf)

Final Report: NSF SBE-CISE Workshop on Cyberinfrastructure and the Social Sciences, F. Berman and H. Brady, held May 12, 2005; information available at <http://vis.sdsc.edu/sbe/reports/SBE-CISE-FINAL.pdf>

From Cyberinfrastructure to Cyberdiscovery in Materials Science: Enhancing outcomes in materials research, education and outreach, August 3-5, 2006; information available at [http://www.mcc.uiuc.edu/nsf/ciw\\_2006/](http://www.mcc.uiuc.edu/nsf/ciw_2006/)

Geoinformatics: Building Cyberinfrastructure for the Earth Sciences (2004); workshop held May 14 – 15, 2003; Kansas Geological Survey Report 2004-48; information available at <http://www.geoinformatics.info>

Geoscience Education and Cyberinfrastructure, Digital Library for Earth System Education, (2004); workshop held April 19-20, 2004; information available at <http://www.dlese.org/documents/reports/GeoEd-CI.pdf>

Getting Up to Speed: The Future of Supercomputing (2004). 308 pp; information available at <http://www.nap.edu/books/0309095026/html/> or [http://www7.nationalacademies.org/cstb/pub\\_supercomp.html](http://www7.nationalacademies.org/cstb/pub_supercomp.html)

High-Performance Computing Requirements for the Computational Solid Earth Sciences (2005); 96 pp; information available at [http://www.geo-prose.com/computational\\_SES.html](http://www.geo-prose.com/computational_SES.html)

High-Performance Computing in the Geosciences Workshop Report. September 25-27, 2006, information available at [http://www.ncar.ucar.edu/Director/dcworkshop/HPCGEO\\_workshop\\_report\\_FINAL.pdf](http://www.ncar.ucar.edu/Director/dcworkshop/HPCGEO_workshop_report_FINAL.pdf)

Identifying Major Scientific Challenges in the Mathematical and Physical Sciences and their CyberInfrastructure Needs, workshop held April 21, 2004; information available at <http://www.nsf.gov/attachments/100811/public/CyberscienceFinal4.pdf>

Improving the effectiveness of U.S. Climate modeling, Commission on Geosciences, Environment and Resources (2001). National Academy Press, Washington, D.C., 144 pp; information available at <http://www.nap.edu/books/0309072573/html/>

GEON: Developing the Cyberinfrastructure for the Earth Sciences A Workshop Report on Intrusive Igneous Rocks, Wilson Cycle and Concept Spaces; information available at <http://geon.geol.vt.edu/geon/pubreps/workshop.pdf>

An Information Technology Infrastructure Plan to Advance Ocean Sciences (2002). 80 pp; information available at <http://www.geo-prose.com/oiti/index.html>

Materials Research Cyberscience enabled by Cyberinfrastructure; workshop held June 17 – 19, 2004; information available at <http://www.nsf.gov/mps/dmr/csci.pdf>

Multi-disciplinary Workshop at the Interface of Cyberinfrastructure, and Operations Research, with “Grand Challenges” in Enterprise-wide Applications in Design, Manufacturing and Services; workshop held August 31 - September 1, 2004; information available at <https://engineering.purdue.edu/PRECISE/CI-OR/index.html>

Multiscale Mathematics Initiative: A Roadmap; workshops held May 3-5, July 20-22, September 21-23, 2004; information available at [http://www.si.umich.edu/InfrastructureWorkshop/documents/NSF\\_2004\\_CIMultiscaleMath.pdf](http://www.si.umich.edu/InfrastructureWorkshop/documents/NSF_2004_CIMultiscaleMath.pdf)

NIH/NSF Spring 2005 Workshop on Visualization Research Challenges; workshop held on May 2-3, 2005; information available at <http://www.sci.utah.edu/vrc2005/index.html>

NSF Conceptual Design Review Panel for the Ocean Observatories Initiative, August 14-17, 2006; information available at <http://www.orionprogram.org/capabilities/cdr/default.html>

NSF EPSCoR Cyberinfrastructure Workshop, May 10-12, 2006; information available at [http://ciworkshop.utsi.edu/NSF\\_CI\\_WkshpReport\\_Final.pdf](http://ciworkshop.utsi.edu/NSF_CI_WkshpReport_Final.pdf)

NSF Workshop for a Plant Cyberinfrastructure Center, October 17-18, 2005; information available at [http://www.arabidopsis.org/portals/masc/masc\\_docs/masc\\_wk\\_rep.jsp](http://www.arabidopsis.org/portals/masc/masc_docs/masc_wk_rep.jsp)

An Operations Cyberinfrastructure: Using Cyberinfrastructure and Operations Research to Improve Productivity in American Enterprises; workshop held August 30 – 31, 2004; information available at <http://www.optimization-online.org/OCI/OCI.doc>; <http://www.optimization-online.org/OCI/OCI.pdf>

Petascale Computing in the Biological Sciences, Workshop Report, Edited by: Allan Snavelly, Gwen Jacobs, and David A. Bader, August 29-30, 2006; information available at [http://www.sdsc.edu/PMaC/BioScience\\_Workshop/Publications/PetascaleBIOworkshopreport.pdf](http://www.sdsc.edu/PMaC/BioScience_Workshop/Publications/PetascaleBIOworkshopreport.pdf)

Planning for Cyberinfrastructure Software (2005); workshop held October 5 – 6, 2004; information available at [http://www.nsf.gov/od/oci/ci\\_workshop/workshopreport\\_final\\_rev2a.pdf](http://www.nsf.gov/od/oci/ci_workshop/workshopreport_final_rev2a.pdf)

Preparing for the Revolution: Information Technology and the Future of the Research University (2002); NRC Policy and Global Affairs, 80 pp; information available at <http://www.nap.edu/catalog/10545.html>

Polar Science and Advanced Networking; workshop held on April 24 - 26, 2003; sponsored by OPP/CISE; information available at <http://www.polar.umcs.maine.edu>

Recurring Surveys: Issues and Opportunities; workshop held March 28-29, 2003; information available at [www.nsf.gov/sbe/ses/mms/nsf04\\_211a.pdf](http://www.nsf.gov/sbe/ses/mms/nsf04_211a.pdf) (2004)

Report of the High Performance Computing Town Hall Meeting: Science, Requirements, and Benchmarks; October 17, 2005; information available at <http://www.mcc.uiuc.edu/nsfhpc/>

Report on the NSF Workshop on Cyber-Based Combustion Science, April 19-20, 2006; information available at [http://www.nsf-combustion.umd.edu/final\\_report.php](http://www.nsf-combustion.umd.edu/final_report.php)

Research Opportunities in CyberEngineering/CyberInfrastructure; workshop held April 22 - 23, 2004; information available at <http://thor.cae.drexel.edu/~workshop/>

Revolutionizing Science and Engineering Through Cyberinfrastructure: report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure; Daniel E. Atkins (Chair), January 2003; information available at <http://www.nsf.gov/cise/sci/reports/atkins.pdf>

Roadmap for the Revitalization of High-End Computing (June 16-18, 2003); information available at <http://www.cra.org/reports/supercomputing.pdf>

Science-Based Case for Large-Scale Simulation; workshop held June 24-25, 2003; information available at [http://www.pnl.gov/scales/docs/volume1\\_72dpi.pdf](http://www.pnl.gov/scales/docs/volume1_72dpi.pdf); [http://www.pnl.gov/scales/docs/SCaLeS\\_v2\\_draft\\_toc.pdf](http://www.pnl.gov/scales/docs/SCaLeS_v2_draft_toc.pdf)

Summit on Digital Tools for the Humanities Report; workshop held September 28-30, 2005; information available at <http://www.iath.virginia.edu/dtsummit/SummitText.pdf>

Supplement to the President's Budget for FY 2007; Report by the Subcommittee on Networking and Information Technology Research and Development (NITRD), February 2006; information available at <http://www.nitrd.gov/pubs/2007supplement/>

Trends in IT Infrastructure in the Ocean Sciences (2004); workshop held May 21-23, 2003; information available at [http://www.geo-prose.com/oceans\\_iti\\_trends/oceans\\_iti\\_trends\\_rpt.pdf](http://www.geo-prose.com/oceans_iti_trends/oceans_iti_trends_rpt.pdf)

Understanding Infrastructure: Dynamics, Tensions, and Design, Report of An NSF Workshop on History and Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructure, held September 2006; information available at [http://www.si.umich.edu/cyber-infrastructure/UnderstandingInfrastructure\\_FinalReport25jan07.pdf](http://www.si.umich.edu/cyber-infrastructure/UnderstandingInfrastructure_FinalReport25jan07.pdf)

Workshop on Challenges of Scientific Workflows; held May 1-2, 2006; information available at [http://vtcpc.isi.edu/wiki/index.php/Main\\_Page](http://vtcpc.isi.edu/wiki/index.php/Main_Page)

Workshop on Cyberinfrastructure in Chemical and Biological Process Systems: Impact and Directions; workshop held September 25-26, 2006; information available at <http://www.oit.ucla.edu/nsfci/NSFCIFullReport.pdf>

## APPENDIX C: CHRONOLOGY OF NSF INFORMATION TECHNOLOGY INVESTMENTS

NSF's early investments in what has now become known as cyberinfrastructure date back almost to the agency's inception. In the 1960s and 1970s, the agency supported a number of campus-based computing facilities. As computational methodologies became increasingly essential to the research endeavor, the science and engineering community began to call for NSF investments in specialized, higher capability computing facilities that would meet the computational needs of the broad national community. As a consequence, NSF's Supercomputer Centers program was initiated in 1985 through the agency's support of five academic-based supercomputer centers.

During the 1980s, academic-based networking activities also flourished. Networking technologies were expected to improve the effectiveness and efficiency of researchers and educators, providing enhanced, easy access to computer resources and more effective transfer and sharing of information and knowledge. After demonstrating the potential of Computer Science Network (CSNET) in linking computer science departments, NSF moved on to develop the high-speed backbone, called NSF Network (NSFNET), with the five supercomputer centers supported under the Supercomputer Centers program and the National Center for Atmospheric Research becoming the first nodes on the backbone. NSF support also encouraged the development of regional networks to connect with the backbone NSFNET, thereby speeding the adoption of networking technologies on campuses around the country. In 1995, in partnership with MCI, NSF catalyzed support of the very high speed Backbone Network Service (vBNS), permitting advanced networking research and the development of novel scientific applications. A few years later, NSF established the NSF Middleware Initiative, focused on the development of advanced networking services to serve the evolving needs of the science and engineering community.

In the early to mid-1990s, informed by both the Branscomb and the Hayes Reports, NSF consolidated its support of national computing facilities in the establishment of the Partnerships for Advanced Computational Infrastructure (PACI)

program. Two partnerships were established in 1997, together involving nearly 100 partner institutions across the country in efforts to make more efficient use of high-end computing in all areas of science and engineering. The partnerships have been instrumental in fostering the maturation of cyberinfrastructure and its widespread adoption by the academic research and education community, and by industry.

Also in the early 1990s, NSF, as part of the U.S. High Performance Computing and Communications (HPCC) program, began to support larger-scale research and education-focused projects pursuing what became known as "grand challenges." These HPCC projects joined scientists and engineers, computer scientists and state-of-the-art cyberinfrastructure technologies to tackle important problems in science and engineering whose solution could be advanced by applying cyberinfrastructure techniques and resources. First coined by the HPCC program, the term "grand challenge" has been widely adopted in many science and engineering fields to signify an overarching goal that requires a large-scale, concerted effort.

During the 1990s, the penetration of increasingly affordable computing and networking technologies on campuses was also leading to the creation of what would become mission-critical, domain-specific cyberinfrastructure. For example, in the mid 1990s the earthquake engineering community began to define what would become the Network for Earthquake Engineering Simulation, one of many significant cyberinfrastructure projects in NSF's portfolio today.

In 1999, the President's Information Technology Advisory Committee (PITAC) released the seminal report Information Technology Research (ITR)-Investing in our Future, prompting new and complementary NSF investments in CI projects, such the Grid Physics Network (GriPhyN) and international Virtual Data Grid Laboratory (iVDGL) and the Geosciences Network, known as GEON. Informed by the PITAC report, NSF also created a Major Research Equipment and Facilities

Construction (MREFC) project entitled Terascale Computing Systems that began its construction phase in FY 2000 and ultimately created the Extensible Terascale Facility – now popularly known as the Teragrid. Teragrid entered its production phase in October 2004 and represents one of the largest, fastest, most comprehensive distributed cyberinfrastructures for science and engineering research and education.

In 2001, NSF charged an Advisory Committee for Cyberinfrastructure, under the leadership of Dr. Dan Atkins, to evaluate the effectiveness of PACI and to make recommendations for future NSF investments in cyberinfrastructure. The Atkins Committee, as it became popularly known,

recommended support for the two Partnership lead sites through the end of their original PACI cooperative agreements. In October 2004, following merit review, the National Science Board (NSB) endorsed funding of those sites through the end of FY 2007.

Through 2005, in addition to the groups already cited, a number of prestigious groups have made recommendations that continue to inform the agency's cyberinfrastructure planning including the High-End Computing Revitalization Task Force, the PITAC Subcommittee on Computational Science, and the NRC Committee on the Future of Supercomputing.

## APPENDIX D: MANAGEMENT OF CYBERINFRASTRUCTURE

NSF has nurtured the growth of what is now called cyberinfrastructure for a number of decades. In recent years, the Directorate for Computer and Information Science and Engineering (CISE) has been responsible for the provision of national supercomputing infrastructure for the academic community. In addition, the Directorate was instrumental in the creation of what ultimately became known as the Internet. During this incubation period, the management of CI was best provided by those also responsible for the research and development of related CI technologies.

Over the years, the penetration and impact of computing and networking on campuses has been extensive, and has led to the creation of many disciplinary-specific or community-specific CI projects and activities. Today, CI projects are supported by all NSF Directorates and Offices. Because of the growing scope of investment and variability in needs among users in the broad science and engineering community, it has become clear that effective CI development and deployment now requires the collective leadership of NSF senior management. This leadership will be

provided by the CIC chaired by the NSF Director and comprised of the NSF Deputy Director, the Assistant Directors of NSF's Research Directorates and the Heads of the Office of International Science and Engineering, Office of Polar Programs, and the recently established Office of Cyberinfrastructure (OCI). The CIC has been meeting regularly since May 2005, and OCI was established in the Office of the Director on July 22, 2005. In January 2007, the CIC authorized a standing CI Coordinators Committee (CICC) to augment its activities with representatives from all NSF Directorates and Offices and chaired by the Director of OCI. A foundation-wide external Advisory Committee for Cyberinfrastructure (ACCI) has also been established and is now active.

CISE will continue to be responsible for a broad range of programs that address the administration's priorities for fundamental research and education in computing, representing more than 85% of the overall federal investment in university-based basic research.

## APPENDIX E: REPRESENTATIVE DISTRIBUTED RESEARCH COMMUNITIES (VIRTUAL ORGANIZATIONS)

### *Collaboratories:*

- **CoSMIC:** Combinatorial Sciences and Materials Informatics Collaboratory <http://mse.iastate.edu/cosmic/> An international research and education center promoting the use of informatics and combinatorial experimentation for materials discovery and design. Based at Iowa State University with domestic partners at Florida International University and the University of Maryland, CoSMIC is composed of an international consortium of universities and laboratories.

**DANSE:** The Data Analysis for Neutron Scattering Experiments [http://wiki.cacr.caltech.edu/danse/index.php/Main\\_Page](http://wiki.cacr.caltech.edu/danse/index.php/Main_Page) Prompted by the development of the Spallation Neutron Source (<http://www.sns.gov>) (SNS), under construction in Oak Ridge, Tennessee, DANSE goals are to build a software system that 1) enables new and more sophisticated science to be performed with neutron scattering experiments, 2) makes the analysis of data easier for all scientists, and 3) provides a robust software infrastructure that can be maintained in the future.

- **DISUN:** Data Intensive Scientific University Network <http://www.disun.org/> A grid-based facility of computational clusters and disk storage arrays distributed across four institutions each of which serves as a High Energy Physics Tier-2 site. The Compact Muon Solenoid (CMS) experiment is primarily supported by DISUN.
- **Ecoinformatics.org** <http://www.ecoinformatics.org/> A community of ecologists and informatics specialists operating on the principles of open-source software communities. Collaborative projects such as software and standards development that cross institutional and funding boundaries are managed through contributed personnel in an open and democratic process. Ecological Metadata Language is a product of this community.
- **GriPhyN:** Grid Physics Network <http://www.griphyn.org/> A large ITR project focused on computer science and grid research applied to the distributed computing and storage requirements

of the high energy physics community. GriPhyN is a foundational element of the Open Science Grid (OSG).

- **ICPSRC:** Inter - University Consortium for Political and Social Research <http://www.icpsr.umich.edu/> Maintains and provides access to a vast archive of social science data for research and instruction, and offers training in quantitative methods to facilitate effective data use. ICPSR preserves data, migrating them to new storage media as changes in technology warrant. ICPSR also provides user support to assist researchers in conducting their projects and identifying relevant data for analysis.
- **iVDGL:** International Virtual Data Grid Laboratory <http://www.ivdgl.org/> A global data and compute grid serving physics and astronomy, funded as a large ITR project by MPS. While its sister project, GriPhyN, funded the grid research, iVDGL supports the operational grid and associated services supporting projects like CMS, ATLAS, and LIGO.
- **LHC:** Large Hadron Collider and **LCG:** Large Hadron Collider Computing Grid The Large Hadron Collider, <http://lhc.web.cern.ch/lhc/> being built at CERN near Geneva, is the largest scientific instrument on the planet. The mission of the LHC Computing Project (LCG) is to build and maintain a data storage and analysis infrastructure for the entire high-energy physics community that will use the LHC. See also the World LHC Computing Grid (WLCG) <http://lcg.web.cern.ch/LCG/>, a distributed production environment for physics data processing.
- **LTER:** Long-Term Ecological Research Network <http://lternet.edu/> A networked collaboratory of 26 field sites and network office focused on answering questions about the long-term dynamics of ecosystems, ranging from near pristine to highly engineered sites. In addition to site-based science, education and outreach, five common research themes ensure that multi-site projects and synthesis are integral to the program.



- NCEAS: National Center for Ecological Analysis and Synthesis <http://www.nceas.ucsb.edu/fmt/doc?frames.html> NCEAS and the National Evolutionary Synthesis Center (NESCent) <http://www.nescent.org/> serve as data centers and collaboratories for the ecology and evolutionary biology communities. Both are involved in the development of software, house databases, and sponsor collaborative activities aimed at synthesizing research.
- NEES: George E. Brown, Jr. Network for Earthquake Engineering Simulation <http://www.nees.org/> A shared national network of 15 experimental facilities, collaborative tools, a centralized data repository, and earthquake simulation software, all linked by the ultra-high-speed Internet2 connections of NEESgrid. These resources provide the means for advanced collaborative research based on experimentation and computational simulations of the ways buildings, bridges, utility systems, coastal regions, and geomaterials perform during seismic events.
- NNIN: National Nanotechnology Infrastructure Network <http://www.nnin.org/> Integrated networked partnership of 13 user facilities that serves the resource needs of nanoscale science, engineering and technology. NNIN provides users with shared open access, on-site and remotely, to leading-edge tools, instrumentation, and capabilities for fabrication, synthesis, characterization, design, simulation and integration for the purpose of building structures, devices, and systems from atomic to complex large-scales.
- OSG: Open Science Grid <http://www.opensciencegrid.org/> The national physics grid arising from the GriPhyN, iVDGL, and DOE's PPDG projects. As an operational grid spread across dozens of institutions and DOE sites, OSG primarily supports high energy physics grid applications and some non-physics applications in an opportunistic grid dominated by commodity clusters whose CPU counts total into the thousands. The OSG Consortium builds and operates the OSG, bringing resources and researchers from universities and national laboratories together and cooperating with other national and international infrastructures to give scientists from many fields access to shared resources worldwide.
- QuarkNet Cosmic Ray eLAB Project <http://www11.i2u2.org:8080/elab/cosmic/project.jsp> A distributed learning lab for collaborating high school physics students and teachers who collect and analyze cosmic ray data. The e-LAB participants work with computer scientists to provide cutting edge tools that use grid techniques to share data, graphs, and posters and to encourage collaboration among students nationwide.
- SCEC CME : Southern California Earthquake Center Community Modeling Environment <http://epicenter.usc.edu/cmeportal/> A recent geophysics and IT collaboratory targeted at seismic hazard analysis and geophysical modeling. The computational test-bed under development will allow users to assemble and run highly complex seismological and geophysical simulations utilizing Teragrid with the goal of forecasting earthquakes in Southern California.
- UltraLight <http://ultralight.caltech.edu/web-site/ultralight/html/index.html> A collaboration of experimental physicists and network engineers to enable petabyte-scale analysis of globally distributed data. Goals include: 1) developing network services that broaden existing Grid computing systems by promoting the network as actively managed component; 2) testing UltraLight in Grid-based physics production and analysis systems; and 3) engineering a trans- and intercontinental optical network testbed, including high-speed data caches and computing clusters.
- Veconlab: The Virginia Economics Laboratory <http://veconlab.econ.virginia.edu/admin.htm/> Focuses on game theory and social interactions in economics and related fields. The Veconlab server provides a set of about 40 web-based programs that can be used to run interactive, social science experiments for either teaching or research purposes.
- Vlab: The Virtual Laboratory for Earth and Planetary Materials <http://www.vlab.msi.umn.edu/>, An interdisciplinary consortium for development and promotion of the theory of planetary materials. Computational determination of geophysically important materials properties at extreme conditions provides accurate information to a) interpret seismic data in the context of likely geophysical processes and b) be used as input for more sophisticated and reliable modeling of planets.

*Observatories:*

- **CHES:** Cornell High Energy Synchrotron Source <http://www.chess.cornell.edu> Provides users with state-of-the-art synchrotron radiation facilities for research in Physics, Chemistry, Biology, and Environmental and Materials Sciences. A special NIH Research Resource, called MacCHES, <http://www.macchess.cornell.edu/> supports special facilities for protein crystallographic studies.
- **CHRNS:** Center for High Resolution Neutron Scattering <http://www.ncnr.nist.gov/programs/CHRNS/> Develops and operates state-of-the-art neutron scattering instrumentation with broad applications in materials research for use by the general scientific community. Combined, CHRNS instruments provide structural information on a length scale of 1 nm to ~10 microns, and dynamical information on energy scales from ~30 neV to ~100 meV, the widest ranges accessible at any neutron research center in North America.
- **EarthScope** <http://www.earthscope.org/> Applies modern observational, analytical and telecommunications technologies to investigate the structure and evolution of the North American continent and the physical processes controlling earthquakes and volcanic eruptions. Efforts involve data transmission from numerous seismographs and GPS receivers to two data centers, serving the raw data in real time and making available derived data products via a web portal.
- **IceCube** <http://www.icecube.wisc.edu/> A one-cubic-kilometer international high-energy neutrino observatory being built and installed in the clear deep ice below the South Pole Station. IceCube will open unexplored bands for astronomy, including the PeV (10<sup>15</sup> eV) energy region, where the Universe is opaque to high energy gamma rays and where cosmic rays do not carry directional information because of their deflection by magnetic fields.
- **IRIS:** Incorporated Research Institutions for Seismology <http://www.iris.edu/> A university research consortium dedicated to exploring the Earth's interior through the collection and distribution of seismographic data. IRIS programs contribute to scholarly research, education, earthquake hazard mitigation, and the verification of a Comprehensive Test Ban Treaty. The IRIS Data Management System manages and disseminates time series data from a variety of worldwide seismic instruments to provide basic data in support of earthquake studies and research into the structure of the Earth's crust, mantle and core.
- **LIGO:** Laser Interferometer Gravitational Wave Observatory <http://www.ligo.caltech.edu> Dedicated to the detection of cosmic gravitational waves and harnessing of these waves for scientific research. The facility consists of two widely separated installations within the United States — one in Hanford Washington and the other in Livingston, Louisiana — operated in unison as a single observatory.
- **Microbial Observatories** <http://www.nsf.gov/bio/pubs/awards/mo.htm> A network of sites or “microbial observatories” in different habitats to study and understand microbial diversity over time and across environmental gradients. Supported projects must establish or participate in an established, Internet-accessible knowledge network to disseminate information resulting from these activities.
- **NEON:** National Ecological Observatory Network <http://www.neoninc.org/> A national research platform and virtual laboratory for studying the role of the biosphere in earths systems by examining the structure, function, and evolution of biological systems at regional to continental scales. A CI backbone will link field instrumentation and mobile instrument platforms, sensor networks, laboratory and field instrumentation, natural history archives, and analytical and modeling capabilities, to facilities for archival, computation, visualization, and forecasting.
- **NHMFL:** National High Magnetic Field Laboratory <http://www.magnet.fsu.edu/index.aspx> Develops and operates high magnetic field facilities that scientists use for research in physics, biology, bioengineering, chemistry, geochemistry, biochemistry, materials science, and engineering. High magnetic fields are a critical link in the development of new materials that impact nearly every modern technology.

- ORION and OOI: Ocean Research Interactive Observatory Networks <http://orionprogram.org/> and <http://www.orionocean.org/OOI/default.html> An integrated observatory network for oceanographic research and education with interactive access to ocean observing systems that track a wide range of episodicity and temporal change phenomena. OOI has three elements: 1) a global-scale array of relocatable deep-sea buoys, 2) a regional-scaled cabled network consisting of interconnected sites on the seafloor, and 3) network of coastal observatories.
- Other Virtual Organizations:*
- AToL: The Tree of Life initiative <http://atol.sdsc.edu/> Supports the reconstruction of the evolutionary history of all organisms, seen as a grand challenge. The AToL research community works toward describing the evolutionary relationships of all 1.7 million described species.
  - CASA: Center for Adaptive Sampling of the Atmosphere <http://www.casa.umass.edu/> Integrates advancements in radar technology, networking and atmospheric sciences to develop a new, low cost network of weather radars that can adapt sampling procedures in real time in order to optimize information and provide unprecedented details of severe storms.
  - CCMC: The Community Coordinated Modeling Center <http://ccmc.gsfc.nasa.gov/> Provides access to modern space science simulations and supports the transition to space weather operations of modern space research models. To support community research CCMC adopts state of the art space weather models that are developed by outside researchers, executes simulation runs with these models, offers a variety of visualization and output analysis tools, and provides access to coupled models and existing model frameworks.
  - CEDAR: Coupling, Energetics and Dynamics of Atmospheric Regions <http://cedarweb.hao.ucar.edu/cgi-bin/ion-p?page=cedarweb.ion> For characterization and understanding of the atmosphere above ~60 km, with emphasis on the processes that determine the basic composition and structure of the atmosphere. Activities include collaborative research projects, multi-site field campaigns, workshops, and participation of graduate and undergraduate students.
  - CIG: Computational Infrastructure for Geodynamics <http://www.geodynamics.org/> Membership-governed organization that supports Earth science by developing and maintaining software for computational geophysics and related fields. CIG consists of: (a) a coordinated effort to develop reusable and open-source geodynamics software; (b) an infrastructure layer of software for assembling state-of-the-art modeling; (c) extension of existing software frameworks to interlink multiple codes and data through a superstructure layer; (d) strategic partnerships with computational science and geoinformatics; and (e) specialized training and workshops.
  - CLIVAR: Climate Variability and Predictability <http://www.clivar.org/> An international research program addressing many issues of natural climate variability and anthropogenic climate change, as part of the wider World Climate Research Programme (WCRP). It makes available a wide variety of climate-related data and models via a web interface and data management system.
  - CUAHSI HIS: The Consortium of Universities for the Advancement of Hydrologic Science, Inc. <http://www.cuahsi.org/> Represents about 100 U.S. universities and develops infrastructure and services that support advancement of hydrologic science and education. The CUAHSI Hydrologic Information System (HIS) project is conducted by a group of academic hydrologists collaborating with the San Diego Supercomputer Center as a technology partner to produce a prototype Hydrologic Information System.
  - C-ZEN: Critical Zone Exploration Network <http://www.czen.org/> Promotes interdisciplinary research on the zone defined by the outer limits of vegetation and the lower boundary of ground water – identified by the National Research Council as the Critical Zone (NRC, 2001). CZEN cyberinfrastructure consists of a critical zone ontology, development of tools for knowledge discovery and management, and data and metadata standards development.
  - DEISA: Distributed European Infrastructure for Supercomputing Applications <http://www.deisa.org/> A consortium of leading national supercomputing centers that currently deploys and operates a persistent, production quality, distributed supercomputing environment with

European continental scope. DEISA's purpose is to enable discovery across a spectrum of science and technology through deep integration of existing national high-end computational platforms, coupled by a dedicated network and supported by innovative system and grid software.

- EGEE: The Enabling Grids for E-science <http://public.eu-egee.org/> Funded by the European Commission to build on recent advances in grid technology and develop a service grid infrastructure which is available to scientists 24 hours-a-day. Spanning over 30 countries and 150 sites across Europe, EGEE supports applications in Earth Sciences, High Energy Physics, Bioinformatics, Astrophysics, and other domains.
- GCAT: Genomic Consortium for Active Teaching <http://www.bio.davidson.edu/projects/gcat/gcat.html> A consortium of institutions whose purpose is to integrate genomic techniques in the undergraduate life sciences curricula. GCAT provides faculty with resources and training to utilize the technology and means to share data across institutions.
- GEON: The Geosciences Network <http://www.geongrid.org/> Advances geoinformatics by training geoscience researchers, educators, students and practitioners in the use of cyberinfrastructure. GEON provides a service-oriented architecture (SOA) for support of "intelligent" search, semantic data integration, visualization of 4D scientific datasets, and access to high performance computing platforms for data analysis and model execution, via the GEON Portal.
- GLOBEC: U.S. GLOBEC (GLOBAL ocean ECosystems dynamics) <http://www.pml.ac.uk/globec/> Looks at how climate change and variability translate into changes in the structure and dynamics of marine ecosystems, marine animal populations and in fishery production. Research approaches use inter-related modeling, process-oriented studies, broad scale observations, and retrospective studies.
- IPBIR: Integrated Primate Biomaterials and Information Resource <http://www.ipbir.org/> To assemble, characterize, and distribute high-quality DNA samples of known provenance with accompanying demographic, geographic, and behavioral information in order to stimulate and facilitate research in primate genetic diversity and evolution, comparative genomics, and population genetics.
- LEAD: Linked Environments for Atmospheric Discovery: <http://lead.ou.edu/> Seeks to mitigate the impacts of extreme weather events by accommodating the real time, on-demand, and dynamically-adaptive needs of mesoscale weather research. A major underpinning of LEAD is dynamic workflow orchestration and data management in a web services framework with grid-enabled systems.
- MARGINS: <http://www.nsf-margins.org/Home.html> Focuses on understanding the complex interplay of processes that govern creation and destruction of continental margins and result in the concentration of both resources and geohazards where land and oceans meet. Data from collaborative research projects are available via a Data Management System and web-based services.
- NCAR: National Center for Atmospheric Research <http://www.ncar.ucar.edu/> A federally funded research and development center that, together with partners at universities and research centers, studies Earth's atmosphere and its interactions with the Sun, the oceans, the biosphere, and human society. NCAR hosts numerous projects that span the full range of VO activities and it recently launched a set of integrative initiatives designed to explore the Earth system from a 21st-century vantage point: <http://www.ncar.ucar.edu/stratplan/initiatives.html>
- NCED: National Center for Earth-surface Dynamics <http://www.nced.umn.edu/> Fosters an integrated, predictive science of the skin of the Earth – "critical zone" - where interwoven physical, biological, geochemical, and anthropogenic processes shape Earth's surface. NCED's main cyberinfrastructure activities relate to a California field site at Angelo Coast Range Reserve with a steep, relatively rapidly eroding landscape where an advanced environmental observatory with a wireless network and automated environmental sensors are under construction.
- NCN: Network for Computational Nanotechnology <http://www.ncn.purdue.edu/> Advances approaches and simulation tools that allow engineers to design new nanoelectronic and NEMS technologies. The fields include: i) nanoelectron-

ics, ii) NEMS, and iii) nano-bioelectronics, with the overall goal of connecting dry electronic and mechanical nanosystems to wet biological nanosystems. NanoHUB <http://www.nanohub.org/> is a web-based initiative spearheaded by NCN that provides online simulation tools for nano-electronics and molecular dynamics, along with semiconductor devices, processes and circuits.

- PEATNET: Peatland Ecosystem Analysis and Training Network <http://www.peatnet.siu.edu/> An international interdisciplinary collaborative effort focused on northern peatland ecosystems. Web site serves as a mechanism to coordinate PEATNET activities and communicate findings and outcomes, and as the location of a peatland Web-discussion group and the PEATNET Virtual Resource Center.
- PRAGMA: The Pacific Rim Application and Grid Middleware Assembly [www.pragma-grid.net/mission.htm](http://www.pragma-grid.net/mission.htm) Formed to establish sustained collaborations and advance the use of grid technologies in applications among a community of investigators working with leading institutions around the Pacific Rim. In PRAGMA, applications are the key, integrating focus that brings together necessary infrastructure and middleware to advance the application's goals.
- RIDGE <http://ocean-ridge.ldeo.columbia.edu/general/html/home.html> A long-term research program to study Earth's oceanic spreading ridge system as an integrated whole, from its inception in the mantle to its manifestations in the biosphere and water column. Ridge 2000 supports interdisciplinary research and availability of data through web-based services.
- SAHRA: The Center for Sustainability of semi-Arid Hydrology and Riparian Areas <http://www.sahra.arizona.edu/> Promotes sustainable management of semiarid and arid water resources, with three components: the SAHRA Geo-database (SGD) for data storage and sharing, hydrologic observations, and integrated modeling at three resolutions.
- SBN: Seamount Biogeosciences Network <http://www.earthref.org/SBN/> Brings together the diverse science disciplines involved in seamount research for improved communication and scientific collaboration, data archiving and integration, and sharing of seagoing logistical operations. Key SBN components include regular workshops and the development of a community website and database.
- Suominet: Real Time Integrated Atmospheric Water Vapor and TEC from GPS <http://www.suominet.ucar.edu/> Measures phase delays induced in GPS signals by the ionosphere and neutral atmosphere with high precision and converts these delays into integrated water vapor and total electron content (TEC). These measurements contribute to our understanding of the Earth's weather and climate system and TEC data addresses topics in upper atmospheric research.
- TeraGrid [www.teragrid.org](http://www.teragrid.org) An open scientific discovery infrastructure combining leadership class resources at eight partner sites to create an integrated, persistent computational resource. Deployed in September 2004, TeraGrid brings over 40 teraflops of computing power and nearly 2 petabytes of rotating storage, and specialized data analysis and visualization resources into production, interconnected at 10-30 gigabits/second via a dedicated national network.
- TESS: Time-sharing Experiments for the Social Sciences <http://www.experimentcentral.org/> Permits original data collection through national telephone surveys to which researchers can add their original questions and through arrangements that allow researchers to run their own studies on random samples of the population that are interviewed via the Internet.
- UNAVCO <http://www.unavco.org/> A non-profit, membership-governed consortium that supports and promotes high-precision techniques for the measurement and understanding of crustal deformation. The data processing center analyzes and distributes GPS data from continuous GPS installations globally, including the dense array of stations dedicated to the EarthScope project.

## IMAGE CREDITS

PAGE	CREDIT
Cover and inside cover	© 2002 The Regents of the University of California
i	Sam Kittner/kittner.com
iv	Courtesy of Ronald Cohen, Carnegie Institution of Washington
2	Blake Harvey, NCSA
4	Marcus Thiebaut, Information Sciences Institute, University of Southern California
6	Aleksei Aksimentiev and Klaus Schulten, Theoretical and Computational Biophysics Group, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign
7	Christy Bowe, ImageCatcher News Service
8 (top)	© 2007 JupiterImages Corporation
8 (bottom)	Marcelo Garcia and Mariano Cantero, University of Illinois at Urbana-Champaign, and S. Balachandar, University of Florida
9	Leroy N. Sanchez, Los Alamos National Laboratory
11	Peter West, National Science Foundation
12	© University Corporation for Atmospheric Research
13	Bob Wilhelmson, NCSA and the University of Illinois at Urbana-Champaign; Lou Wicker, National Oceanic and Atmospheric Administration's National Severe Storms Laboratory; Matt Gilmore and Lee Counce, University of Illinois Atmospheric Science Department. Visualization by Donna Cox, Robert Patterson, Stuart Levy, Matt Hall and Alex Betts, NCSA
14	Courtesy of Theoretical and Computational Biophysics Group, Beckman Institute, University of Illinois at Urbana-Champaign; Mercury Cluster background image courtesy of NCSA and the Board of Trustees of the University of Illinois
15	Gary Strand, NCAR
16	Scientific contact Ed Seidel (eseidel@aci.mpg.de); simulations by Max Planck Institute for Gravitational Physics (Albert-Einstein-AEI); visualization by Werner Bengel, Zuse Institute, Berlin (ZIB) and AEI. The computations were performed on NCSA's It
17	Peter Virnau, Massachusetts Institute of Technology
19	Kyle McQuilkin and Ryan Ridgely, Ohio University

<u>PAGE</u>	<u>CREDIT</u>
20	James Matthews, Linghao Zhong and John Brady, Cornell University; Mike Himmel and Mark Nimlos, NREL; Tauna Rignall, Colorado School of Mines; Mike Crowley, The Scripps Research Institute. Work was funded by an NREL subcontract funded by DOE Office of the Biomass Program
21	Jason C. Fisher, © 2005 American Institute of Biological Sciences
22	National Virtual Observatory and California Institute of Technology ( <a href="http://us-vo.org/">http://us-vo.org/</a> )
23	National Science Foundation
24	Courtesy of Montage, IPAC, Caltech, SDSC. This research made use of Montage, funded by NASA's Earth Science Technology Office, Computation Technologies Project, under Cooperative Agreement Number NCC5-626 between NASA and the California Institute of Technology. Montage is maintained by the NASA/IPAC Infrared Science Archive
25	IRIS
26, 27	© CERN
28	Wolfgang Bluhm
30	EarthScope
32	Curtis Suplee, National Science Foundation
33	Ethan Dicks, National Science Foundation
34	John Orcutt, Scripps Institute of Oceanography
35	HPWREN
36	Peggy Greb, Agricultural Research Service Image Gallery, USDA
38	Prairie View Solar Observatory
39	The JASON Project
40	Gary Meek, Georgia Institute of Technology
41	Theoretical and Computational Biophysics Group, University of Illinois at Urbana-Champaign
42	Patrick Motl, Mario D'Souza, Joel E. Tohline, and Juhan Frank, Louisiana State University

NATIONAL SCIENCE FOUNDATION  
4201 WILSON BOULEVARD  
ARLINGTON, VA 22230

NSF 07-28

