Large Facilities & CyberInfrastructure NSF Introduction

Bill Miller

Science Advisor
Division of Advanced Cyberinfrastructure, (ACI)
Directorate for Computer & Information Science & Engineering (CISE)
National Science Foundation
WLMiller@nsf.gov

On behalf of facility and CI colleagues across NSF

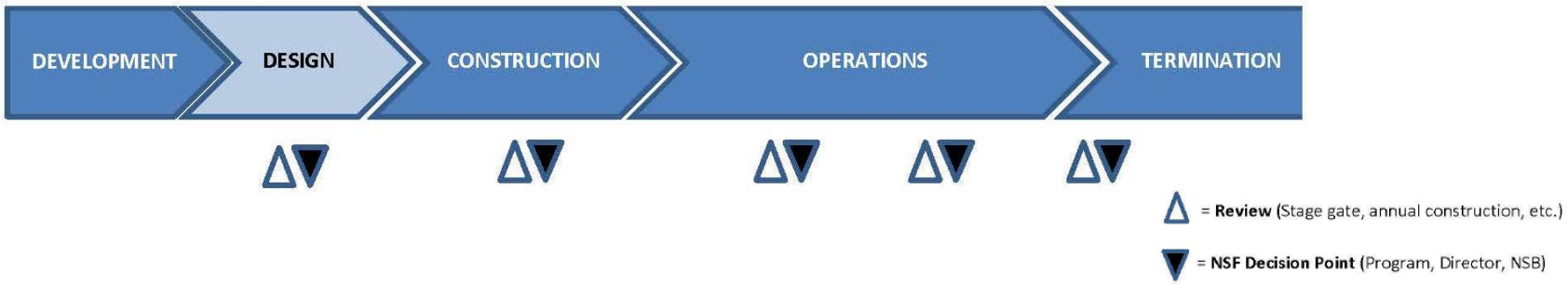
NSF Facility CI Workshop, Dec 1-2, 2015

NSF Motivations

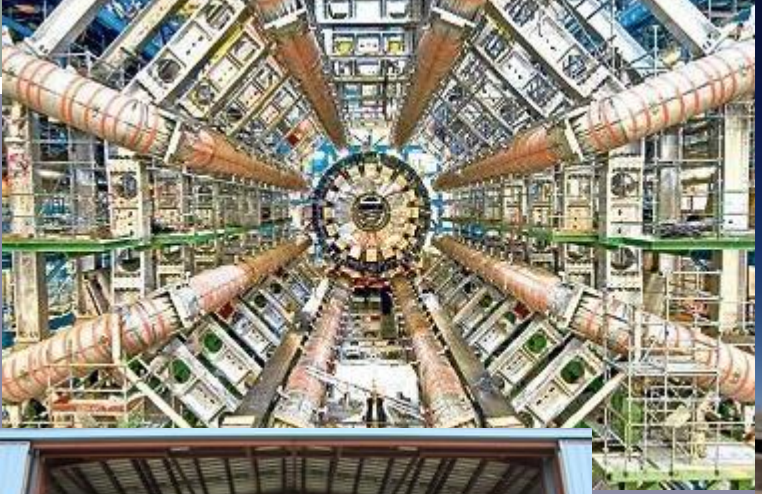
- NSF invests in a complex array of **large facilities** and **research cyberinfrastructure** to enable research at national and global scales.
- These two spheres have been largely independent but are increasingly interacting, as
  - Facilities become cyber-driven to support their missions and users.
  - Team-science and multi-disciplinary collaboration proliferate to address complex questions, integrate facility and other data, and validate computational models.
  - The scientific process is being revolutionized via technology for communication, computing, data management…
- We want **to understand current and future CI needs of facilities**, and tune and leverage our CI investments to maximize the scientific impact of facilities and research *writ large*.

NSF Large Facilities

- A unique, diverse array of national and global platforms that enable large scale science and engineering research.
- Largest (and longest) sustained NSF investments:
  - Construction: 5 MREFC projects, ~$300-600M each.
  - Operations: 20+ (~$1 B/yr R&RA). Lifetime in decades.
- Fundamentally different from other NSF activities: Long gestation; complex approval, budgeting, oversight; high on agency, community, Executive and Congressional radars.

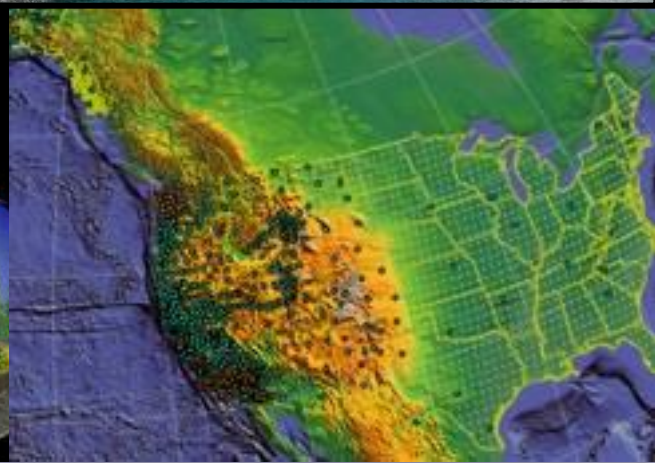


www.nsf.gov/bfa/lfo/large-facilities-list.pdf



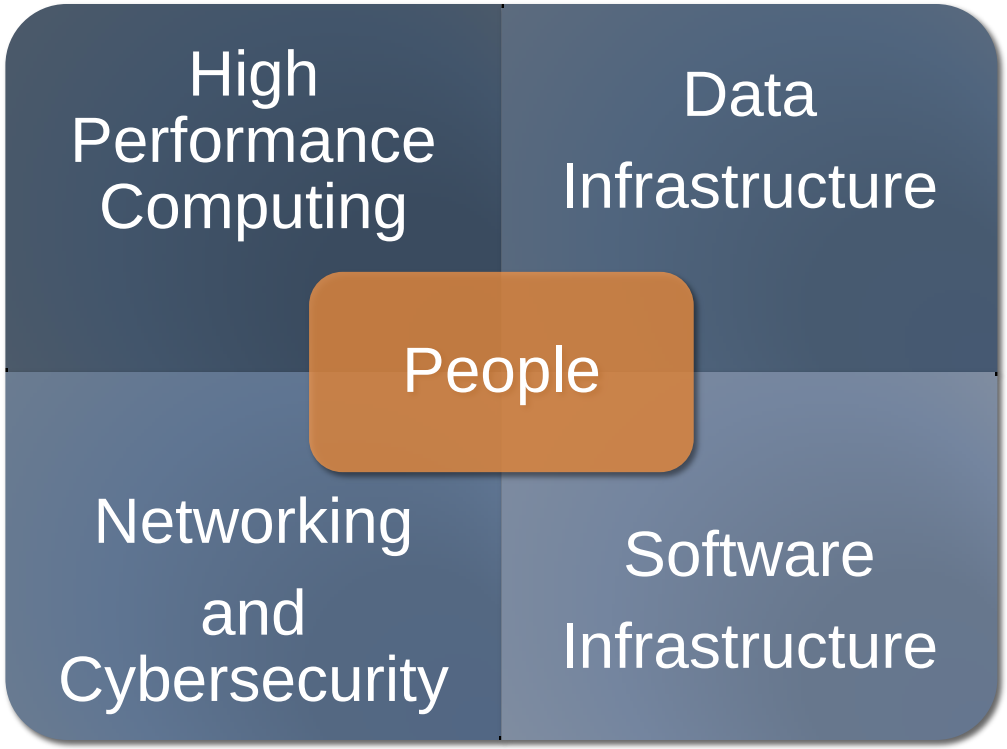
NSF


IRMINGER SEA
60°N
30°W
50°N
145°W
46°N
127°W
REGIONAL SCALE NODES ENDURANCE ARRAY
PIONEER ARRAY
40°N
70°W
ARGENTINE BASIN
42°S
42°W
SOUTHERN OCEAN
55°S
90°W



NSF Support for Research Cyberinfrastructure

- Directorates are all heavily invested to connect researchers to data, analysis platforms, and advanced computing.
- Dedicated units: CISE Advanced Cyberinfrastructure (ACI), BIO Biological Infrastructure (DBI)

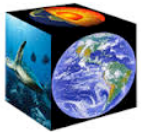


- Collaboration
- Virtual control
- Gateways & Hubs
- Data Management
- Analysis
- Simulation
- Integration
- Workflows
- …

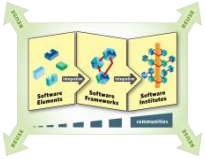
NSF Support for Research Cyberinfrastructure

Program and Project Examples

- **HubZero** (support from ENG, CISE). Open source platform for creating sites that support scientific discovery, collaboration and learning, e.g. NanoHub.
- **iPlant** (BIO, with connectivity to ACI-funded resources). Extensible CI platform for data-intensive life sciences research, including access to sequencing, data, storage, analysis tools, supercomputing, cloud, and more.



- **EarthCube** (ACI, GEO). Major community-based effort to create new capabilities for research, data and knowledge sharing in the geosciences. Involves academia, government, industry, CI, CS and social sciences.



- **Data Infrastructure Building Blocks** and **Software Institutes** programs (ACI with co-funding from all NSF directorates). Fund projects to deploy shared CI resources within and across science domains.

- **Extreme Digital and Leadership High Performance Computing** programs (ACI), and Yellowstone HPC (GEO). Provide world-class performance computational science resources and services.

- **Open Science Grid** (OSG) (ACI, MPS, DOE). Shared, distributed high-throughput computing resources for the whole science community.

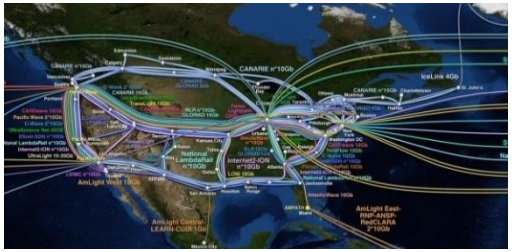
Transformation in how science is done

An ecosystem of shared science and cyber infrastructure resources

Large Scale
Science Facilities

National Supercomputing
& Distributed Computing
And Coordination Services

Social Collaboration
and Data Networks

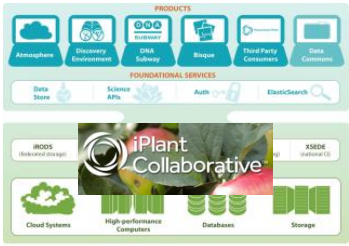


Researcher(s)

Software

Data

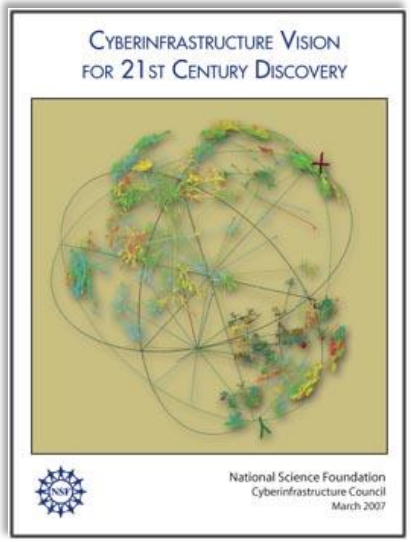
Commercial
Cloud
Services



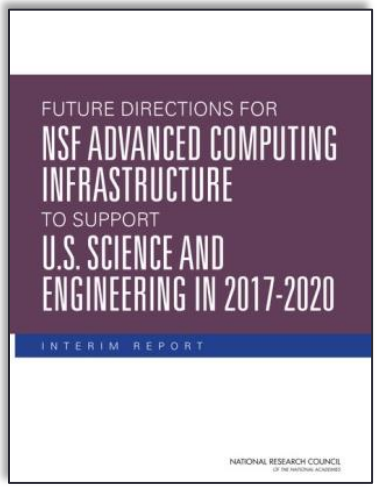
Shared Data/Software,
Gateway, & Hub Resources
And User Services

NSF's strategy for cyberinfrastructure is informed by many forms of community input

Initial Vision (2007-2010)



National Academies Study (On going)



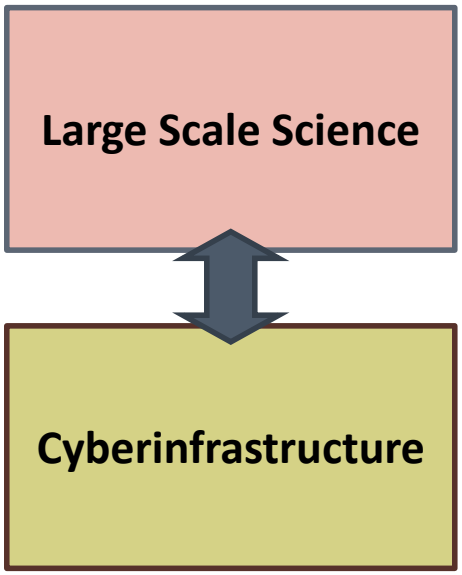
Interim Report, Oct 2014
Final Report expected Fall 2015

NSF Advisory Committee for Cyberinfrastructure (ACCI)

Directorate Advisory Committees

RFI on Science Drivers Requiring Capable Exascale HPC

Desired outcomes for the workshop



- **Gain a deeper understanding of the CI landscape within and external to large facilities.**
  - Identify related needs, gaps and trends that can guide current and future NSF CI programs, and suggest new funding emphases.
- **Establish a direct dialog between the NSF large facility community and the CI community.**
  - Facilitate exchange of ideas and practices, success cases, and awareness of technology developments and trends.
  - Maximize use of existing CI resources, minimizing duplication of effort (and funds).
  - Develop new partnerships among facilities and CI developers to address challenges, and inspire new R&D to accelerate discovery.
- **Ultimately: Foster a dynamic, interconnected *national cyberinfrastructure ecosystem*** that evolves to support the changing needs of the whole research community.

Thanks!

NSF Large Facilities List

In Operations (20+)

- Atacama Large Millimeter Array (ALMA), MPS
- United States Antarctic Program (ANT), GEO
- Arecibo Observatory (AO), MPS/GEO
- Academic Research Fleet (ARF), GEO
- Cornell High Energy Synchrotron Source (CHESS), MPS
- Seismological & Geodesy Facilities and EarthScope (SAGE, GAGE), GEO
- GEMINI Observatory (GEMINI), MPS
- Ice Cube (Ice Cube), GEO/MPS
- International Ocean Discovery Program (IODP), GEO
- Large Hadron Collider, ATLAS, CMS (LHC), MPS
- Laser Interferometer Gravitational-wave Observatory (LIGO), MPS
- National Center for Atmospheric Research (NCAR), GEO
- Natural Hazards Engineering Research Infrastructure (NHERI), ENG
- National High Magnetic Field Laboratory (NHMFL), MPS
- National Nanotechnology Coordinated Infrastructure (NNCI), ENG
- National Optical Astronomy Observatory (NOAO), MPS
- National Radio Astronomy Observatory (NRAO), MPS
- National Superconducting Cyclotron Laboratory (NSCL), MPS
- National Solar Observatory (NSO), MPS

Under construction (5)

- Alaska Region Research Vessel (R/V Sikuliaq), GEO
- Daniel K. Inouye Solar Telescope (DKIST), MPS
- Large Synoptic Survey Telescope (LSST), MPS
- National Ecological Observatory Network (NEON), BIO
- Ocean Observatories Initiative (OOI), GEO

Latest list with links:
www.nsf.gov/bfa/lfo/large-facilities-list.pdf