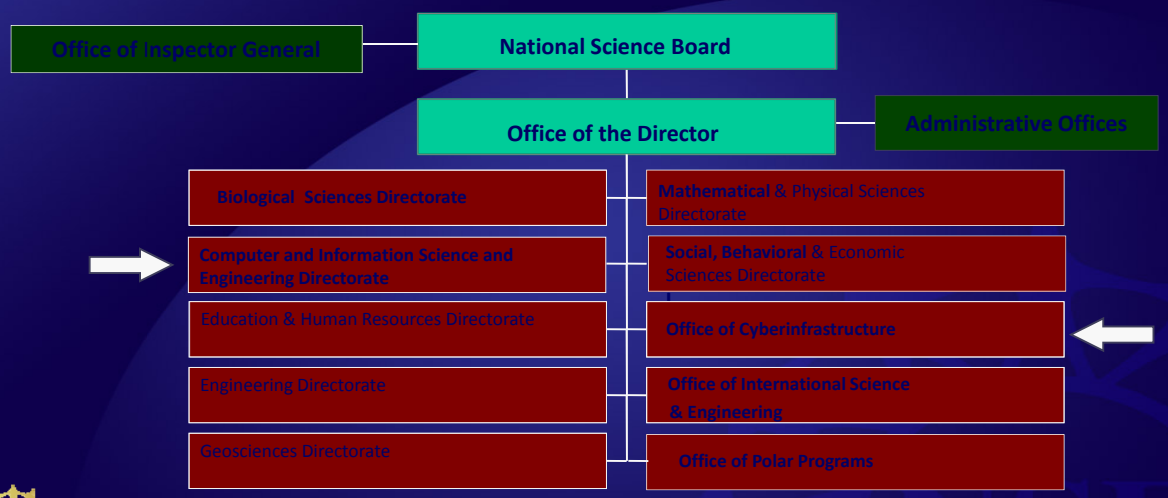
National Science Foundation
Office of Cyberinfrastructure

Program Overview

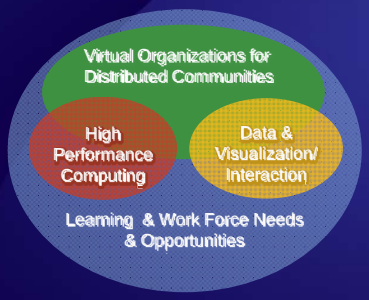
National Science Foundation

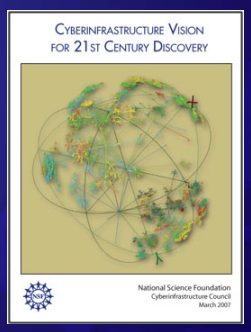


OCI – An Office in Transition

NSF Vision for Cyberinfrastructure circa 2007

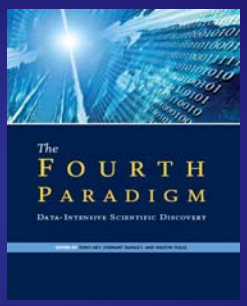
- "National-level, integrated system of hardware, software, data resources & services… to enable new paradigms of science"

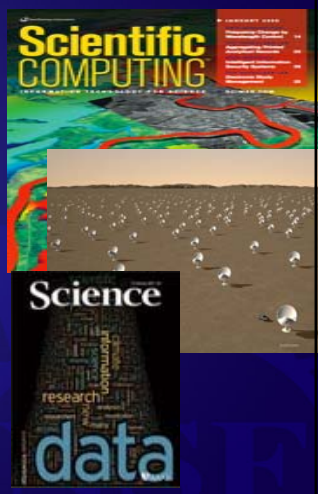




Framing the Challenge: Science and Society Transformed by Data

- Modern science
  - Data- and compute-intensive
  - Integrative, multiscale
- Multi-disciplinary Collaborations for Complexity
  - Individuals, groups, teams, communities
- Sea of Data
  - Age of Observation
  - Distributed, central repositories, sensor-driven, diverse, etc





CISE/ACI (OCI) Network Infrastructure Programs

- CC-NIE (Campus Cyberinfrastructure – Network Infrastructure and Engineering) – NSF 12-541
  - network infrastructure improvements at campus level (re-architecting for large data science flows, upgrades within campus and external connectivity)
  - Network integration and applied innovation (e.g. experimental deployment of SDN/openflow)
  - $21.6M, 39 awards ($3M in CISE/CNS co-funding) – see Wed. afternoon talk on CC-NIE here at I2
- IRNC – International Research Network Connections
  - Continues NSF's ~20 year commitment to directly supporting international network connectivity dedicated to research and education
  - mid-way through 5-year awards supporting multi-gig connectivity to Europe, Asia, Americas, Australia – see www.irnclinks.net for more info
  - New approach under consideration for a globally coordinated International Research Network Backbone to meet science and education demands into 2020, stay tuned

Campus Cyberinfrastructure – Network Infrastructure and Engineering (CC-NIE)

- NSF 12-541 – solicitation released March 1, 2012
- http://www.nsf.gov/publications/pub_summ.jsp?WT.z_pims_id=504748&ods_key=nsf12541
- 1st area: Data Driven Networking Infrastructure for the Campus and Researcher
- 2nd area: Network Integration and Applied Innovation
- This is a networking solicitation
- Proposals due May 30, 2012

Software as Infrastructure

- Software is an integral enabler across all science and engineering disciplines of computation, experiment, and theory and central component of new cyberinfrastructure
- Software enables education; education enables software
- Applications becoming more complex: multiscale/multimodel simulation codes, new data analytics and statistics, etc.
- Environments and applications becoming more complex: heterogeneous, distributed, massively parallel, etc.
- Software engineering: robustness, usability, reliability, V & V, etc.
- Research environment's challenges: motivation, credit, funding, licensing, multiple disciplines, international, etc.

SCIENCE

CDSE and SOFTWARE

INFRASTRUCTURE

Science drives technology innovation and vice versa

NSF/Office of CyberInfrastructure

Create and maintain a software ecosystem providing new **capabilities** that advance and accelerate scientific inquiry at unprecedented complexity and scale

Support the foundational **research** necessary to continue to efficiently advance scientific software

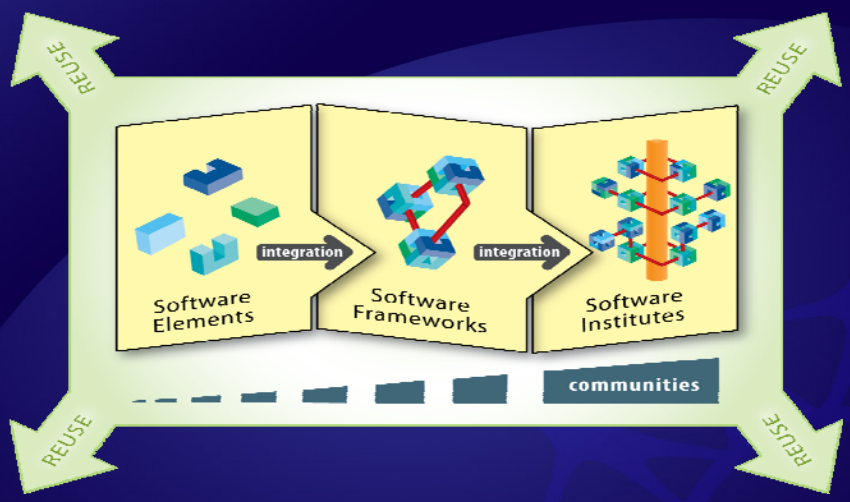
Enable transformative, interdisciplinary, collaborative, **science and engineering** research and education through the use of advanced software and services

Transform practice through new **policies** for software addressing challenges of academic culture, open dissemination and use, reproducibility and trust, curation, sustainability, governance, citation, stewardship, and attribution of software authorship

Develop a next generation diverse workforce of scientists and engineers equipped with essential skills to use and develop software, with software and services used in both the research and **education** process

NSF/Office of CyberInfrastructure

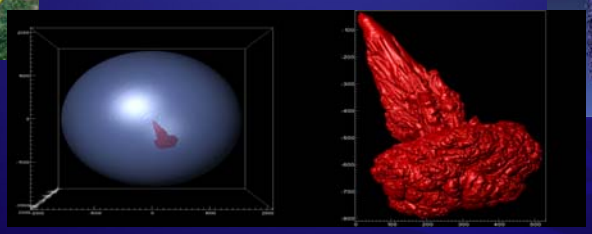
Putting it together

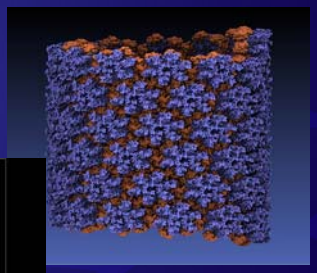


Look at NSF 11-539 and 11-589 for past solicitations;
current NSF 12-576; new solicitations coming;
consider supplements, EAGERs, etc.

OCI contribution to NSF's Advanced Computational Infrastructure (ACI)

Vision: *Support a comprehensive portfolio of advanced computing infrastructure, programs and other resources to facilitate cutting-edge foundational research in Computational and Data Enabled Science and Engineering (CDS&E) and its applications to all disciplines.*



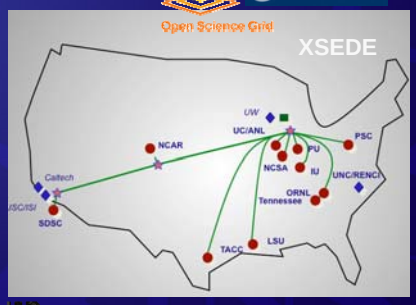


Advanced Computational Infrastructure

OCI ACI Strategy - Advancing CI for Science and Engineering Research and Education

- *Anticipate and invest in diverse and innovative national scale shared resources, outreach and education complementing campus and other national investments*
- *Leverage and invest in collaborative flexible "fabrics" dynamically connecting scientific communities with computational resources and services at all scales (campus, regional, national, international)*

CIPRES – Cyberinfrastructure for Phylogenic Research



Significant OCI FY2013 Projects: Blue Waters/UIUC

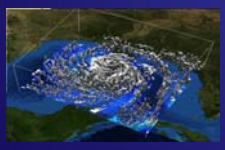
National resource offering large allocations for a small number of diverse and significant research projects across the U.S.

Highly Scalable Heterogeneous System to enable investigations of computationally challenging problems that require sustained PetaFlops (10^{15}) performance and/or large data and large memory

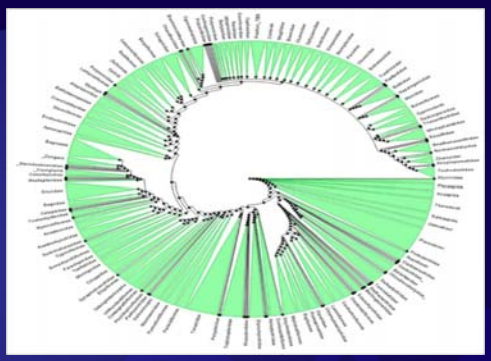
Significant FY2013 Projects: Stampede/UT at Austin

- Expands the range of data intensive computationally-challenging science and engineering applications that can be tackled with current national resources
- Introduces new heterogeneous architecture based on Intel MIC to science and engineering research communities

Stampede will accommodate larger simulations (both in fidelity and number of ensemble members) producing more accurate forecasts, and permit more research groups during critical response efforts



Hurricane Ike tracking predictions using the WRF program and 30 ensemble members (Courtesy F. Zhang, PSU and Y. Weng)



Phylogenic Trees: Stampede will allow us to approach the full tree for all green plant species (~500,000) on Earth to gain insights into the origins of drought resistance or nitrogen efficiency in plants, which could then be bred into future food crops

Crosscutting and NSF-wide Opportunities

- Big Data Core Technologies
- Computational and Data-Enabled Science and Engineering (CDS&E)
- **Cyber-enabled Sustainability Science and Engineering (CyberSEES)**
- Innovation Corps (I-Corps)
- INSPIRE
- Research Coordination Networks
- **Secure and Trustworthy Cyberspace (SaTC)**

Core Techniques and Technologies for Advancing Big Data Science & Engineering (BIG DATA)

Foundational research to extract knowledge from data

- Foundational research for managing, analyzing, visualizing, and extracting useful information from large, diverse, distributed, and heterogeneous data sets.

Collection, Storage, and Management of "Big Data"

- Data representation, storage, and retrieval
- New parallel data architectures, including clouds
- Data management policies, including privacy and access
- Communication and storage devices with extreme capacities
- Sustainable economic models for access and preservation

Data Analytics

- Computational, mathematical, statistical, and algorithmic techniques for modeling high dimensional data
- Learning, inference, prediction, and knowledge discovery for large volumes of dynamic data sets
- Data mining to enable automated hypothesis generation, event correlation, and anomaly detection
- Information infusion of multiple data sources

Data Sharing and Collaboration

- Tools for distant data sharing, real time visualization, and software reuse of complex data sets
- Cross disciplinary information and knowledge sharing
- Remote operation and real time access to distant data sources and instruments

Cross-Directorate Program: NSF Wide

Multi-agency Commitment: NSF and NIH

CyberSEES (NSF 13-500)
Cyber-Enabled Sustainability Science & Engineering

Goal: To advance sustainability science via interdisciplinary research enabled by innovations in computing & information sciences.

The proposed research under this solicitation must be well-grounded in at least one concrete sustainability problem facing the planet, but applicable more generally. The sustainability challenges associated with computing itself are also covered by the solicitation.

Award Types:

Type 1: Small, proof-of-concept, capacity building, or exploratory research projects (up to $300K, 2 yrs)

Type 2: Regular projects (up to $1.2M, 4 yrs)

Deadlines:

Letters of Intent (required): Dec 4, 2012;

Proposals: Feb 5, 2013 (at most 2 per PI)

Funding: Total $12M, 12-20 awards (including Type 1 & 2)

Secure and Trustworthy Cyberspace (SaTC)

NSF 12-503 involves CISE, MPS, OCI, SBE

- SaTC supports projects that promise fundamental scientific advances and technologies to protect cyber-systems (including host machines, the internet and other cyber-infrastructure) from malicious behavior, while preserving privacy and promoting usability
  - Trustworthy Computing Systems: Design, building or operation or analysis of attack resistant cyber-infrastructure
  - Social, Behavioral and Economic Sciences: Develops new theories and methods in social, behavior, and economic sciences that advance cyber-security
  - Transitions To Practice: later stage activities in the research and development lifecycle such as prototyping and experimental deployment
- Small (500k), Medium (1.2 Million), Frontier ($10 Million) awards

Questions?

Marc Rigas
Staff Associate
Office of Cyberinfrastructure

mrigas@nsf.gov