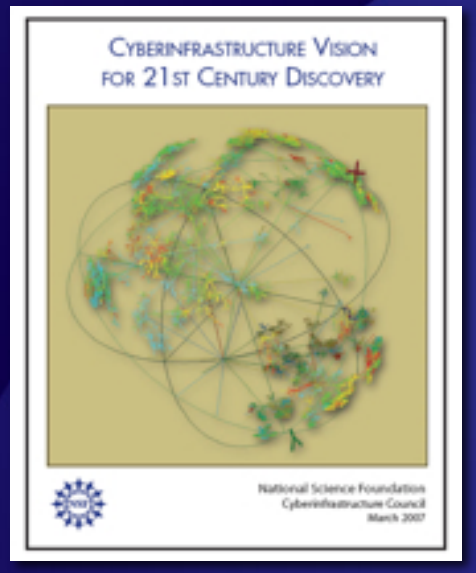
(Exascale) Software at NSF

IESP Meeting
Oxford, April 2010

Rob Pennington
Program Officer
Office of CyberInfrastructure (OCI)

What is Needed?

An ecosystem, not components…

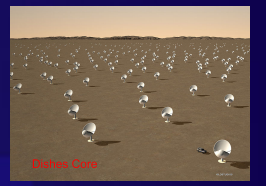


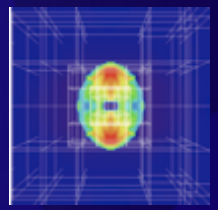
NSF-wide CI *Framework* for 21st Century Science & Engineering

People, Sustainability, Innovation, Integration

CF21: Cyberinfrastructure Framework…

- High-end computation, data, visualization transformative science; *sustainability, extensibility*
  - Facilities/centers as *hubs of innovation*
- MREFCs and collaborations including large-scale NSF collaborative facilities, international partners
- Software, tools, science applications, and VOs critical to science, integrally connected to instruments
- Campuses fundamentally linked end-to-end; grids, clouds, loosely coupled campus services, policy to support
- People - Comprehensive approach workforce development for 21st century science and engineering





NSF Office of Cyberinfrastructure (OCI) and CF21

- Driver for integrative CI activity via CF21
  - Working with all units, community
    - Develop vision and implementation plan
    - OCI budget ¼ NSF CI
- Catalyst for *coordinated, linked* investments
  - CI in all forms: campus, centers, MREFC
    - Leadership in R&D for prototypes, pilots, best practices
    - Looking for coherence, re-use of CI
  - Science applications enabled by CI
  - People: supporting next generation of CI researchers
- Steward for NSF-wide computational science
  - Working with all NSF units to provide sustainable home

Software Drivers

- Software infrastructure is an important component of Cyberinfrastructure
  - Perhaps least coordinated and systematically funded
  - Recognized in the PITAC reports as the "Grand Challenge" of Cyberinfrastructure
- Software critical for simulation and data synthesis/analysis
  - Used in science discovery, engineering design, decision support, etc.
  - It is not the computer science community that bemoans lack of investment in software, so much as the application community at large
  - Scientists expect new hardware capabilities and observational facilities to yield new scientific results and software is often the bottleneck

Initiatives

- Task Forces for strategic input from the community(s)
- PetaApps and Petascale Computing Resource Allocation (PRAC) program
- Software Infrastructure for Sustained Innovation (SI^2)
- G8 Research Councils Initiative on Software for Exascale Computing

ACCI Task Forces

Campus Bridging: Craig Stewart, IU (BIO)

Data & Viz: Shenda Baker, Harvey Mudd (MPS); Tony Hey, (CISE)

- Timelines: 12-18 months
- Advising NSF
- Workshop(s)
- Recommendations
- Input to NSF informs CF21 programs, 2011-2 CI Vision Plan

Software: David Keyes, Columbia/ KAUST (MPS)

Computing: Thomas Zacharia, ORNL/ UTK (DOE)

Education & Workforce: Alex Ramirez, CEOSE

Grand Challenge Communities/VOs: Tinsley Oden, Austin (ENG)

Questions and More Questions

- What are the new applications that will drive CI use that are emerging or likely to emerge in the coming decade? What are examples of the research questions that will require the use of CI to address?
- Key Challenge "SCALE THE SCIENCE -- NOT THE CODES!" Verification, validation, uncertainty quantification --> PREDICTION!
- How can NSF best stimulate development of these applications?
- How can useful application and systems software be developed, disseminated and sustained beyond the development period?

More Questions

- What application support environments will be needed? Numeric and non-numeric library packages, problem-solving environments?
  - How do we make these sustainable?

- How can NSF catalyze developments that make it possible to use good tools -- compilers, debuggers and performance tools, on system scales all the way down to the typical researcher's laptop or desktop?

- What education and training actions should be considered to prepare researchers, students and educators for future cyberinfrastructure?

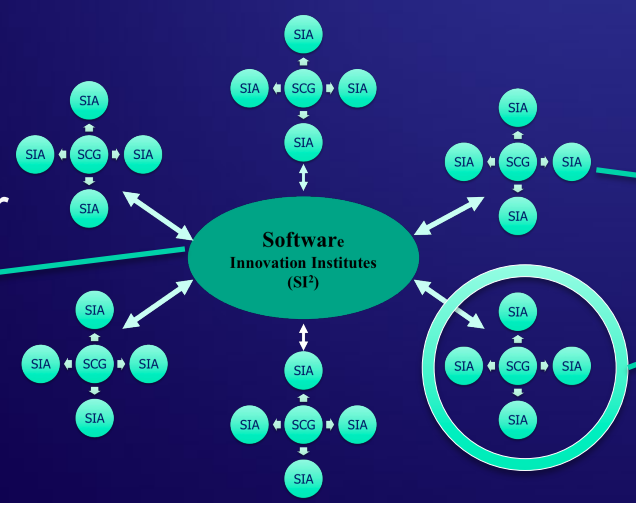
Sustained Long-Term Investment in Software

- Transform innovations into sustainable software
  - Robust, efficient, resilient, repeatable, manageable, sustainable, community-based, etc.
- Catalyze software as a symbiotic "process"
- Address all aspects, layers and phases of software
  - Systematic approaches
  - Theory validated by empirical trials
  - Tools that embody and support processes
  - Metrics, validation mechanisms, governance structures
  - Amortised over large (global) user communities
  - Support for maintenance and user support

CF21 Software Institutes and Innovators (SI^2)

- Transform innovation into sustainable software
- Significant multiscale, long-term program
  - Perhaps $200M+ over a decade
    - $10M+ identified in FY10
    - Significant ramp up in future years FY11 and beyond
  - Connected institutes, teams, investigators
  - Integrated into CF21 framework

3-6 centers, long term, for critical mass, sustainability



Many individuals w/short term grants, funded by OCI and directorates

Numerous teams of scientists and computational and computer scientists with longer term grants

Software Infrastructure for Sustained Innovation (SI^2) – More Information

- Dear Colleagues Letter
  - http://www.nsf.gov/pubs/2010/nsf10029/nsf10029.jsp?WT.mc_id=USNSF_179
- Solicitation
  - http://www.nsf.gov/si2/
- Point of Contact: Manish Parashar

  mparasha@nsf.gov

  www.nsf.gov/oci

Software: PetaApps

- OCI in partnership with the other NSF Directorates and Offices
- Develop the future simulation, optimization and analysis tools that use emerging petascale computing
- Will advance frontiers of research in science and engineering with a high likelihood of enabling transformative research
- Areas examined include:

-Climate Change

-Earthquake Dynamics

-Storm Surge Models

-Supernovae simulations

G8 HORCS Exascale Software Pilot Collaboration

- **G8 Research Councils Initiative on Multilateral Research**
  - Interdisciplinary Program on Application Software towards Exascale Computing for Global Scale Issues
- Participating funding agencies
  - U.S., Canada, France, Germany, Japan, Russia, UK
- Deutsche Forschungsgemeinschaft (DFG) is the Call Secretariat
  - Official website at http://www.dfg.de/g8-initiative

Research Projects

- Specifically focused on computational science research for globally significant science or engineering which requires exascale capabilities
- Collaborative research projects
  - Support for collaborations between experts in research areas related to global challenges and developers of future exascale platforms
  - Researchers from at least three of the partner countries
  - Jointly reviewed by the participating funding organizations
  - Successful projects are expected to demonstrate added value through multilateral collaboration
  - Support for U.S.-based researchers will be provided through awards made by the National Science Foundation

Timeline

- Preliminary Proposal Due Date to the Call Secretariat
  - 7 May 2010
- Notification for Submission of Full Proposals
  - 28 June 2010
- Full Proposals Due Date to Call Secretariat and NSF
  - 25 August 2010
- Official Funding Decisions/Award
  - February 2010

Budgets/Funding

- Funding of the participating researchers provided by their respective national funding organization
  - Standard award terms and conditions.
  - Funding is meant for collaborative research
  - May include clearly justified travel and workshops
- The total budget is approximately 10 million € over three years
  - Funding can be provided for projects lasting for two or three years
  - Anticipated that 8-10 research consortia will be funded in this call
  - NSF anticipates making awards at a level of approximately $150,000 per award per year

Next Priorities

- Integrating all activities into a much more comprehensive cyberinfrastructure… a cyberinfrastructure framework for 21st century science
- New programs in software: life-cycle, all layers
- Creating deeper partnerships with DoE and other agencies; international partnerships
- Cyber-learning
- Creating a computational science research agenda that crosses NSF and reaches out to other agencies and other countries