Challenges of Software as Infrastructure at NSF/CISE/ACI

Daniel S. Katz
Program Director, Division of Advanced Cyberinfrastructure
Cyberinfrastructure Framework for 21st Century Science and Engineering (CIF21)

- Cross-NSF portfolio of activities to provide integrated cyber resources that will enable new multidisciplinary research opportunities in all science and engineering fields by leveraging ongoing investments and using common approaches and components (http://www.nsf.gov/cif21)

- ACCI task force reports (http://www.nsf.gov/od/oci/taskforces/index.jsp)
  - Campus Bridging, Cyberlearning & Workforce Development, Data & Visualization, Grand Challenges, HPC, Software for Science & Engineering
  - Included recommendation for NSF-wide CDS&E program

- Vision and Strategy Reports

- Implementation
  - Implementation of Software Vision
    http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504817
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.
ACI Software Cluster Programs

• Exploiting Parallelism and Scalability (XPS)
  – CISE (including ACI): for foundational groundbreaking research leading to a new era of parallel (and distributed) computing
  – Proposals submitted in Feb.

• Computational and Data-Enabled Science & Engineering (CDS&E)
  – Virtual program (ENG, MPS, ACI) for science-specific proofing of algorithms and codes
  – Identify and capitalize on opportunities for major scientific and engineering breakthroughs through new computational and data analysis approaches

• Software Infrastructure for Sustained Innovation (SI2)
  – Transform innovations in research and education into sustained software resources that are an integral part of the cyberinfrastructure
  – Develop and maintain sustainable software infrastructure that can enhance productivity and accelerate innovation in science and engineering
Challenges

• What fraction of funds should be spent on support of existing infrastructure vs. development of new infrastructure?
  – How do we decide when to stop supporting a software element?

• Does the open source model work for all science?
  – For some science? For underlying tools?
  – How many users/developers are needed for success?

• How do we encourage reuse and discourage duplication?
  – Credit & attribution?

• How do we make science reproducible
  – CI tools?, policies?, V&V?

• How do we more effectively support career paths for software developers (with universities, labs, etc.)