Software Infrastructure for Sustained Innovation (SI²)

http://www.nsf.gov/si2/
Science is Revolutionized by CI

- **Modern science**
  - Data- and compute-intensive
  - Integrative

- **Multiscale Collaborations for Complexity**
  - Individuals, groups, teams, communities

- **Must Transition** NSF CI approach to support
  - Integrative, multiscale
  - 4 centuries of constancy, 4 decades $10^9$-$10^{12}$ change!

- **Multiple crisis**
  - Hardware, Data, Education/WFD, ...
Software is Critical

- CI – Unprecedented complexity, challenges

- Software is essential to every aspect of CI – “the glue”
  - Drivers, middleware, runtime, programming systems/tools, applications, ...

- This software is different …? 
  - In its natures, who builds it, how is it built, where it runs, its lifetime, etc.

- Software crisis?
  - Software complexity is impeding the use of CI
    - Science apps have $10^3$ to $10^6$+ lines, have bugs
    - Developed over decades – long lifecycles (~35 years)
  - Software/systems design/engineering issues
    - Emergent rather than by design
  - Quality of science in question
Software Crisis?

THE SOFTWARE, IT NO WORKY!!!
Unlike HW/Instruments, it is easy to ignore software....
Software Grand Challenge

- SW as the modality for CF21 and Computational Science in the 21st Century
- Sustainable SW as a CI resource
  - What SW to sustain?
  - How to sustain it?
- Fundamental Grand Challenge: Robust, Sustainable and Manageable Software at CI-Scale
  - Repeatability, Reliability, Performance, Usability, Energy efficiency, ....

- Sustainability, manageability, etc., are NOT add-ons – it has to be integrated into the design
Many complex aspects....

- **Building the right software** – application involvement, understanding requirements
  - scales, types of software, target user communities

- **Building software right** – teams, reward structures, processes, metrics, verification/testing

- **Protecting investments** – active management, sustainability, leverage/reuse, ownership, business models

- **Building trust** – user community must be able to depend on the availability of a robust and reliable software infrastructure!
Cycles of Innovation: The Current State

- Independent cycles of innovations in relevant disciplines
  - Out of phase; different timescale
- Coupling (if any) is loose and asynchronous
  - Incorrect and/or in-efficient solutions
- Few synergies; Plenty of repetition and re-invention

Loose coupling
("Throw it over the fence…")

Methods/Formulations

Challenge Problem

Codes/Software

Scientific Insight

Computational Technology Innovation

Software/Systems

Algorithms/Processes

Challenge Problems

NSF
Cyber-Science: Synergies & Symbiosis

S2I2:
- Processes, mechanism, frameworks
- Foster, nurture partnerships
- Build trust
Sustained Long-Term Investment in Software

- Transform innovations into sustainable software that is an integral part of a comprehensive cyberinfrastructure
  - robust, efficient, resilient, repeatable, manageable, sustainable, community-based, etc.

- Catalyze and nurture multidisciplinary software as a symbiotic “process” with ongoing evolution
  - Domain and computational scientists, software technologists

- 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
Sustained Long-Term Investment in Software

- Significant multiscale, long-term program
  - Envisions $200-300M over a decade
  - Connected institutes, teams, investigators
  - Integrated into CF21 framework

3-6 centers, 5+5 years, for critical mass, sustainability

Many individuals w/short term grant

Numerous teams of scientists and computational and computer scientists with longer term grants
Software Infrastructure for Sustained Innovations (SI²) - Mechanisms

- Create a software ecosystem that scales from individual or small groups of software innovators to large hubs of software excellence
- 3 interlocking levels of funding

Scientific Software Elements (SSE): 1–2 PIs
- $0.2 – 0.5M, 3 years

Scientific Software Integration (SSI): Focused Groups
- ~$1M per year, 3 – 5 years

Scientific Software Innovation Institutes (S²I²): Large Multidisciplinary Groups
- $6–8M per year, 5 (+) years
- Planning Activities
- FY 11 and beyond only

Focus on innovation

Focus on sustainability
Software Infrastructure for Sustained Innovation (SI²): FY10 First round

- **Letters of Intent (Required) – May 10, 2010**
  - Title, Team, Synopsis (science/engr. drivers, target user community, specific software elements)

- **Full Proposals – June 14, 2010**
  - SSE: ~2 PIs + 2 GAs, 3 years
  - SSI: ~3-4 PIs, 3-4 GAs, 1-2 senior personnel/developers, 3-5 years
  - No S2I2 in FY 10

- Proposals from all parts of NSF were received
  - 200 projects were submitted
  - ~10% overall funding rate is anticipated

- Now we look to the future of this program!!!!
Scientific Software Innovation Institutes (S2I2) – Call for Exploratory Workshop Proposals

Goals:

- Inform NSF on what should be included in the solicitation
- Inform the community as it responds to the solicitation in FY11
- Provide a forum of discussions about the SI2 vision, and S2I2 models and structures within and across communities.
Scientific Software Innovation Institutes (S2I2) – Call for Exploratory Workshop Proposal

Questions

- What scientific areas have significant challenges that can benefit, in terms of scientific innovation/discovery as well as productivity, from an S2I2

  - Is there an need for such an Institute and if so what would be the appropriate focus area(s) and scale?
  - What communities would it serve, who would participate, what interconnections would it have to the larger community of computational scientists, experimentalists, and beyond.

- What are the key attributes of an S2I2? What are appropriate organizational, personnel and management structures, as well as operational processes?
Scientific Software Innovation Institutes (S2I2) – Call for Exploratory Workshop Proposals

Questions

- What expertise and capabilities should an S2I2 provide and how should it interface and interact with science communities? What education and outreach functionalities are meaningful in an S2I2?

- What are the critical linkages between an S2I2 and other components of a community cyberinfrastructure (i.e., software tools, databases, instruments, etc.)? What is the unique role of an S2I2 in the broader cyberinfrastructure ecosystem (e.g., TeraGrid/XD, DataNet, MREFC, etc.)?
Scientific Software Innovation Institutes (S2I2) – Call for Exploratory Workshop Proposals

Questions

- What are meaningful metrics, evaluation mechanisms and governance structures for an S2I2? What are appropriate approaches to sustainability of the S2I2?

- How would an S2I2 impact the science and engineering community and impacts its practices, capabilities and productivity?
Software Infrastructure for Sustained Innovation (SI²): Metrics of Success

- Buy-in from the broader community
- Demonstrated leverage and reuse
- Emergence of successful models, processes, architectures, metrics for S&E software – empirically validated
- Emergence of models and mechanisms for community sustainability of software institutes
- Accepted research agenda by academic community
Software Infrastructure for Sustained Innovation (SI²) – More Information

- DCL
- Solicitation
- S2I2 DCL
- SI² POC: Manish Parashar
  - mparasha@nsf.gov
Summary

- Science is being revolutionized through CI
  - Compute, data, networking advance suddenly 9-12 orders of magnitude after 4 centuries
  - All forms of CI—integrated—needed for complex science
- NSF responsive: developing much more comprehensive, integrated CF21 initiative
  - Community deeply engaged in planning
  - Activities begin FY10, ramp up FY11-12 and beyond
- Focus on sustainability, people, innovation
  - Longer term programs, better linked, hubs of innovation
  - Support development of computational scientists who develop and/or use advanced CI
- Robust, reliable, sustainable software is critical!
Thank You!

Long-lived: enduring, tenacious, robust

Sustainable System
“meets the needs of the present without compromising the ability of future generations to meet their own needs”

[UN Brundtland Report 1987, of sustainable development]