Software Infrastructure for Sustained Innovation (SI^2)

http://www.nsf.gov/si2/

Talk Outline

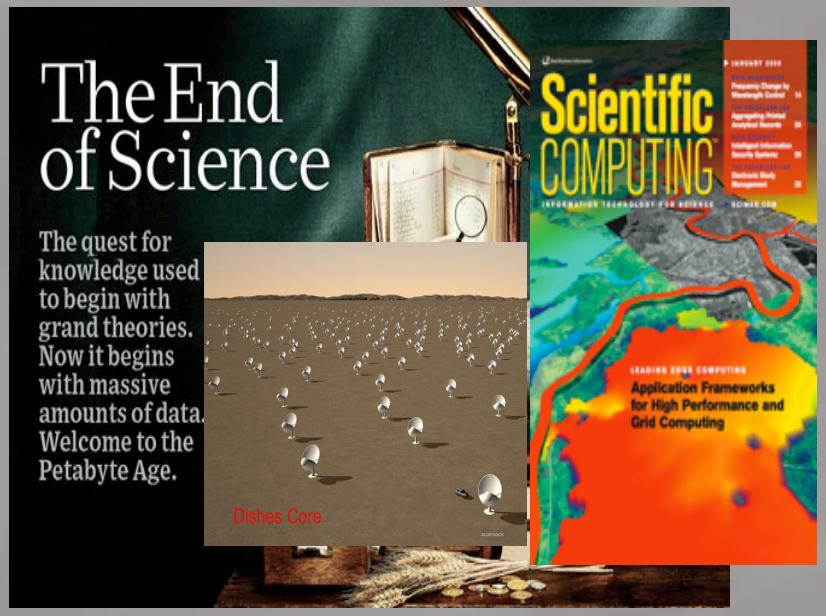
- Major NSF theme - CF21
  - Software an integral part of the investment
- Rethinking the software challenge
- Elements of SI2
- Output from workshop as input to NSF
- Summary

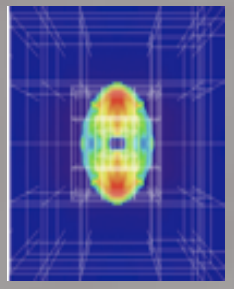
Cyberinfrastructure Framework for 21st Century Science and Engineering (CF21)

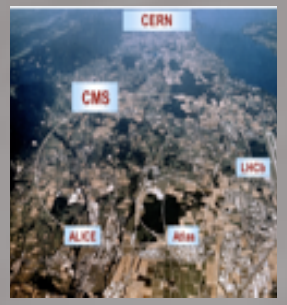
"it is imperative that NSF develop a strategic long term vision"

The Driver for CF21
Science is Revolutionized by CI

- Modern science is often
  - Data- and compute-intensive
  - Integrative (cross-disciplinary)
- Collaborative by necessarily to address complexity
  - Individuals, groups, teams, communities
- NSF CI approach must evolve to meet the needs of the community
  - Effective investments will mitigate the community from confronting major impediments in hardware, data education and workforce development
- On the back of a galloping pony
  - 4 centuries of constancy
  - 4 decades 10^{9-12} change!







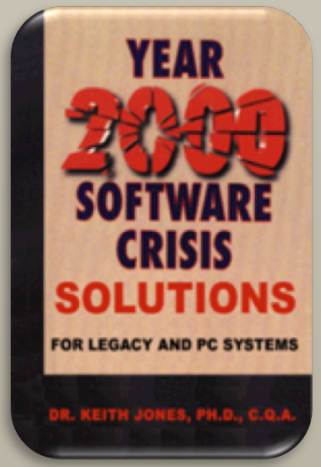
Concepts and Capabilities of CF21
NSF response

- High end infrastructure and services with focus on sustainability and extensibility
- Networking including focus on linking cyberinfrastructure framework into campuses
- Major national and international research facilities and collaborations including large-scale NSF collaborative facilities and projects
- A comprehensive plan for education and outreach in computational science

http://www.nsf.gov/pubs/2010/nsf10015/nsf10015.pdf

Software: Critical Aspect of CF21

- 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 Development and the NSF Modes of Support

- Computational Science & SW is an essential elements of CF21
- Growing recognition of software as infrastructure
- Software must have characteristics of infrastructure
  - must be sustainable, robust, manageable, and extendable
    repeatability, reliability, performance, usability, energy efficiency, ….
- These characteristics must be integrated into the design of software systems, not after thoughts
- Often the NSF modes of support for the development of software do not promote software systems with the most designable characteristics

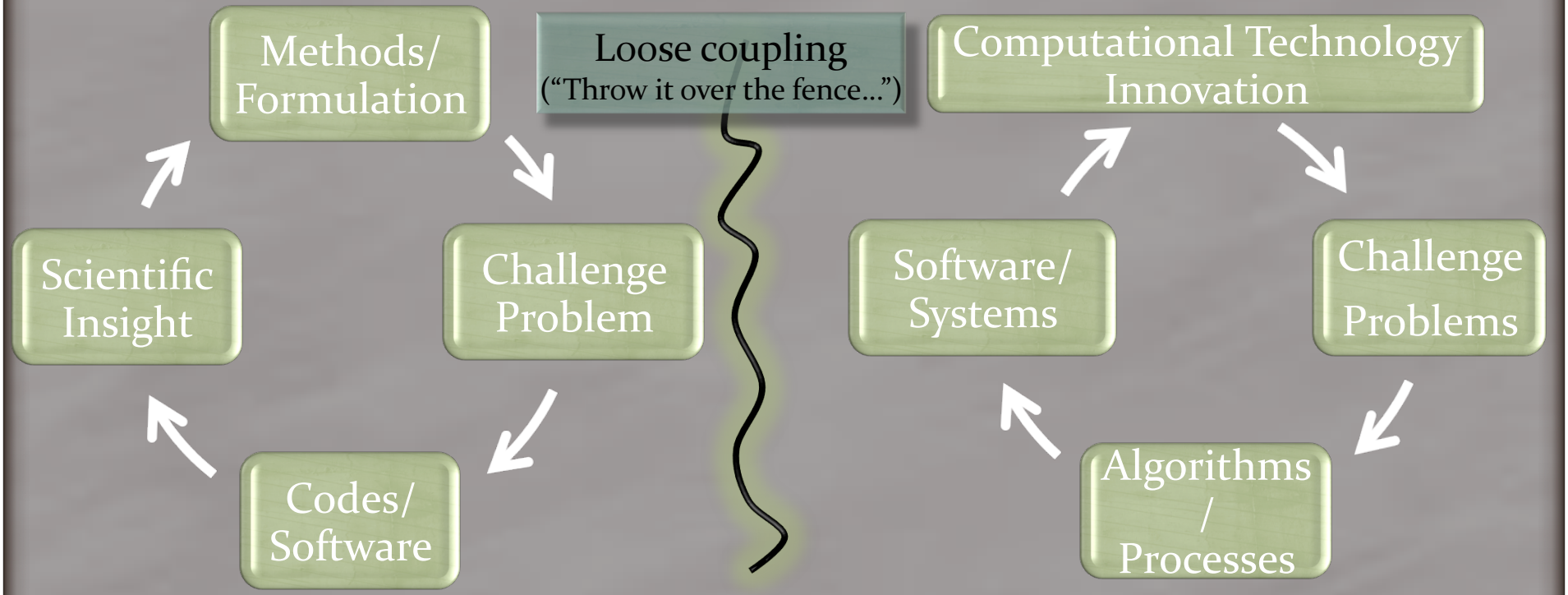
Aspects to be considered in software development

- Building the right software – discipline scientist/engineers 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!

Rethinking the Software Challenge

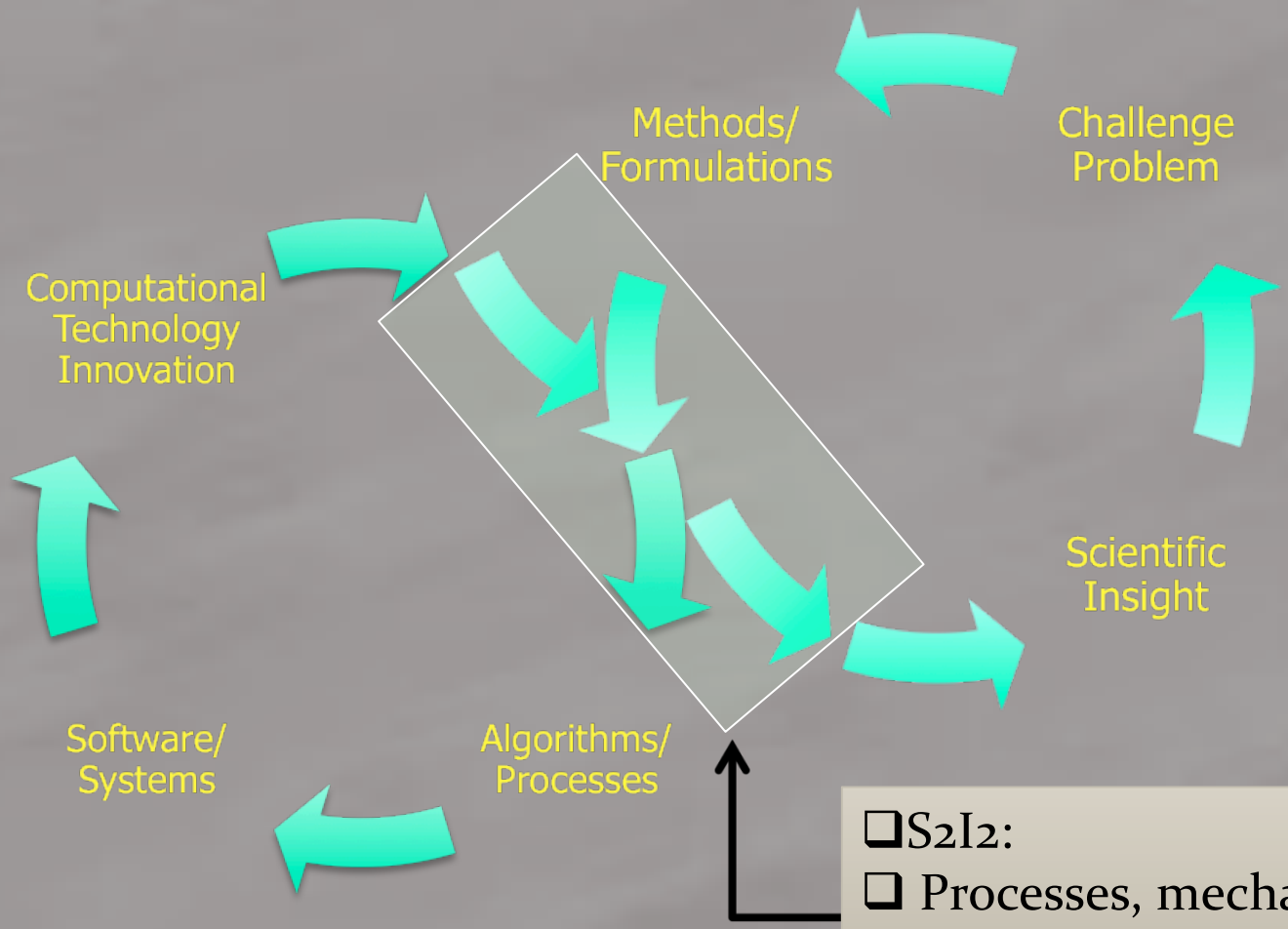
A better approach to investing

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

Cyber-Science: Synergies & Symbiosis



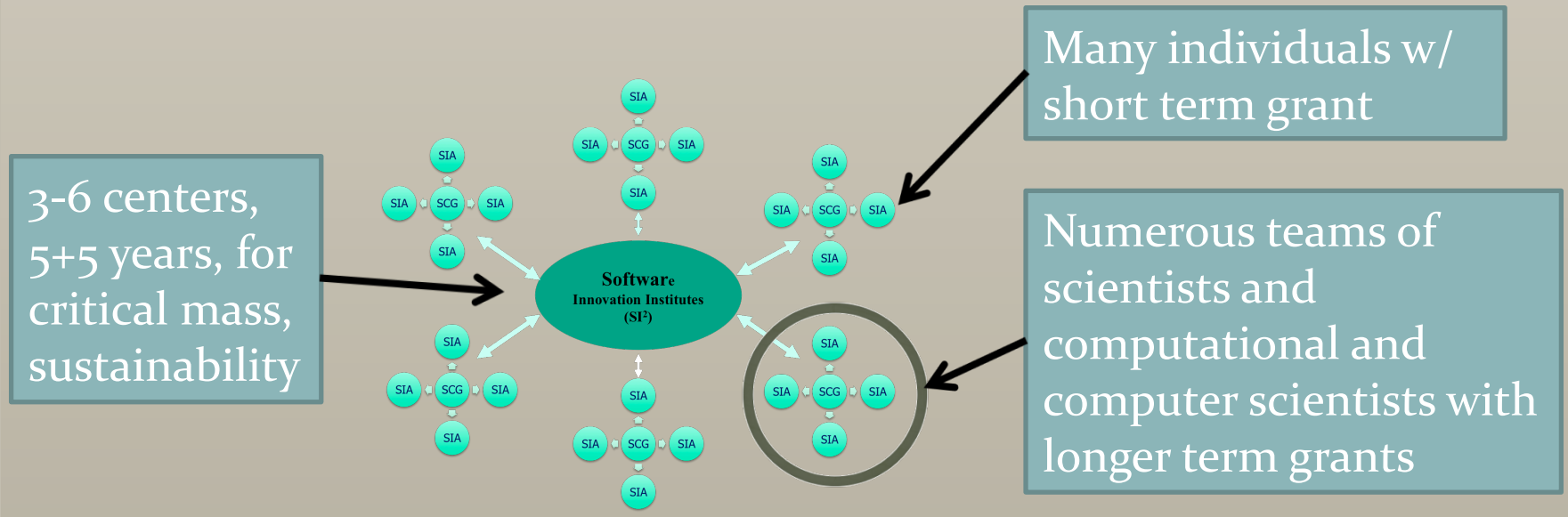
- S2I2:
  - Processes, mechanism, frameworks
  - Foster, nurture partnerships
  - Build trust

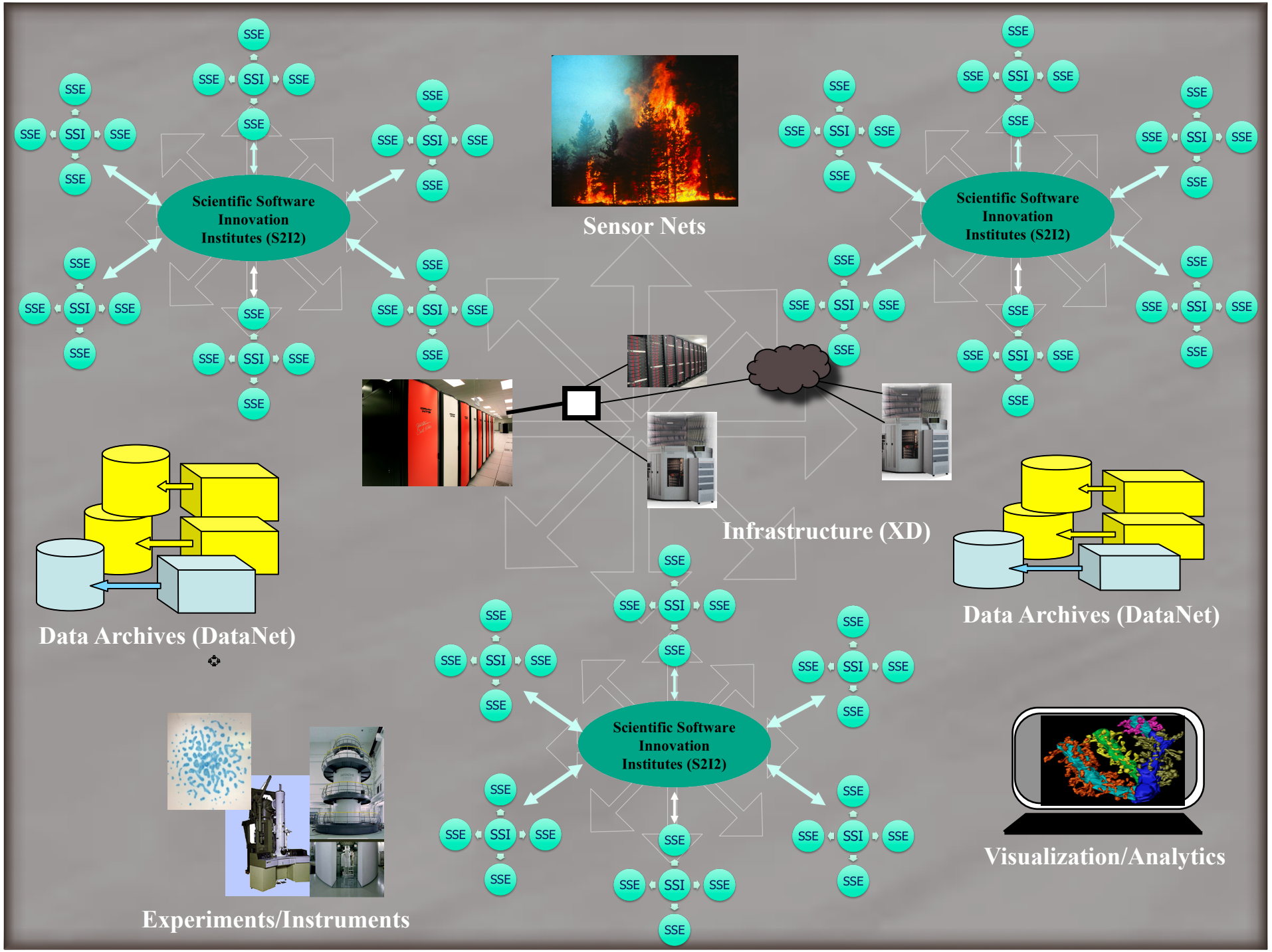
Sustained Long-Term Investment in Software

- 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 transdisciplinary 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



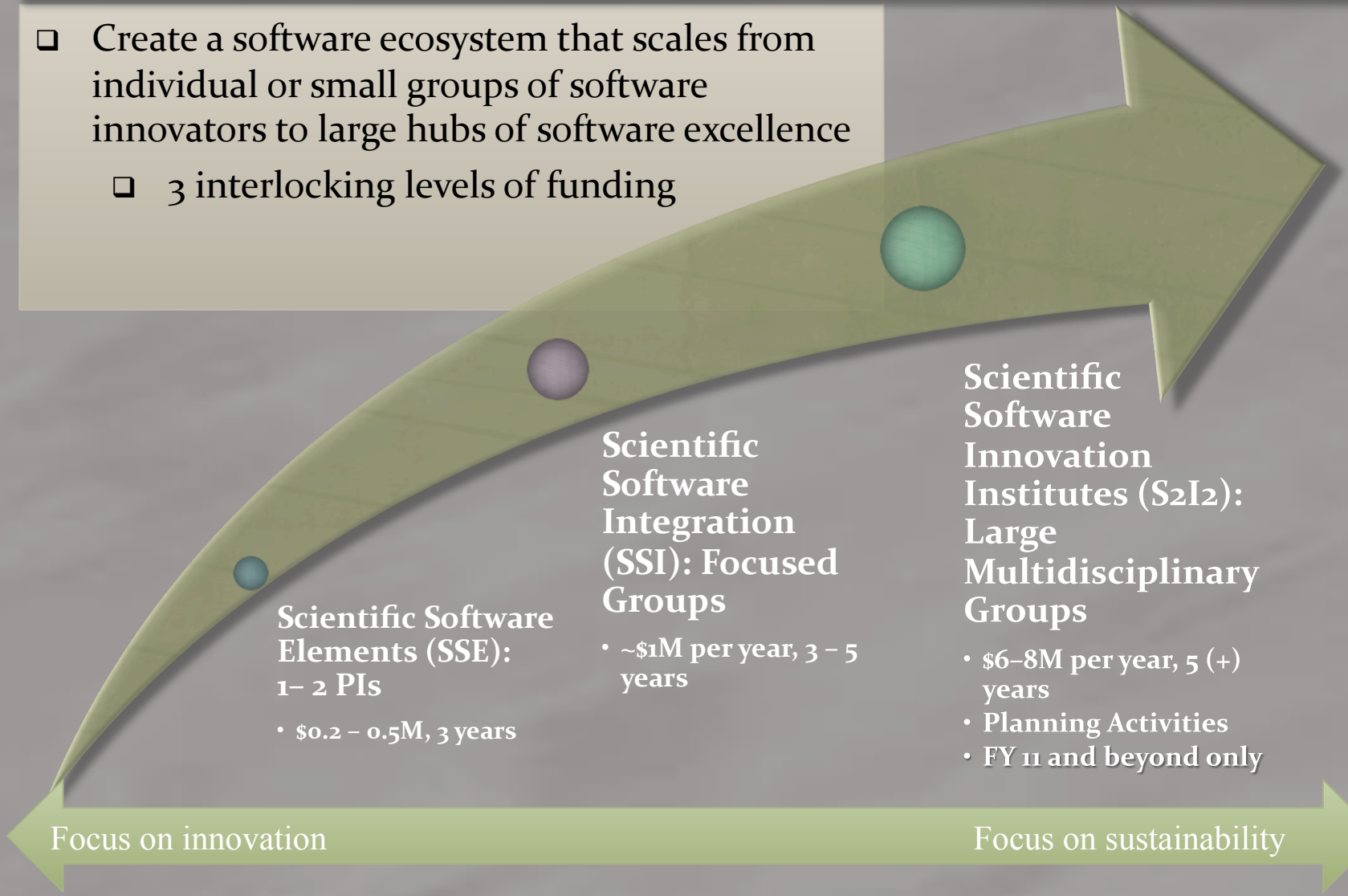

SSE
SSI
Scientific Software Innovation Institutes (S2I2)
Sensor Nets
Infrastructure (XD)
Data Archives (DataNet)
Data Archives (DataNet)
Experiments/Instruments
Visualization/Analytics

Elements of SI2

Experience from FY 10 and planning for FY 11

Software Infrastructure for Sustained Innovations (SI^2) - 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



Software Infrastructure for Sustained Innovation (SI^2): 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
  - Funds allocated $11.4M and fund award $18.7M
- **Now we look to the future of this program!!!!**

Goals and Output from Workshop

Influencing future direction of SI2

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 Proposals

- 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 (SI^2): 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 (SI^2) – More Information

- DCL
  - http://www.nsf.gov/pubs/2010/nsf10029/nsf10029.jsp?WT.mc_id=USNSF_179
- Solicitation
  - http://www.nsf.gov/si2/
- S2I2 DCL
  - http://www.nsf.gov/pubs/2010/nsf10050/nsf10050.jsp?org=NSF
- SI^2 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!
Questions and/or Comments

Voyager Spacecraft (1977 -):

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]