

# Cyberinfrastructure: Enabling New Research Frontiers

Sangtae "Sang" Kim

Division Director – Division of Shared Cyberinfrastructure  
Directorate for Computer and Information Science  
and Engineering

National Science Foundation

CUDI Meeting, Oct. 15, 2004 – via video conference

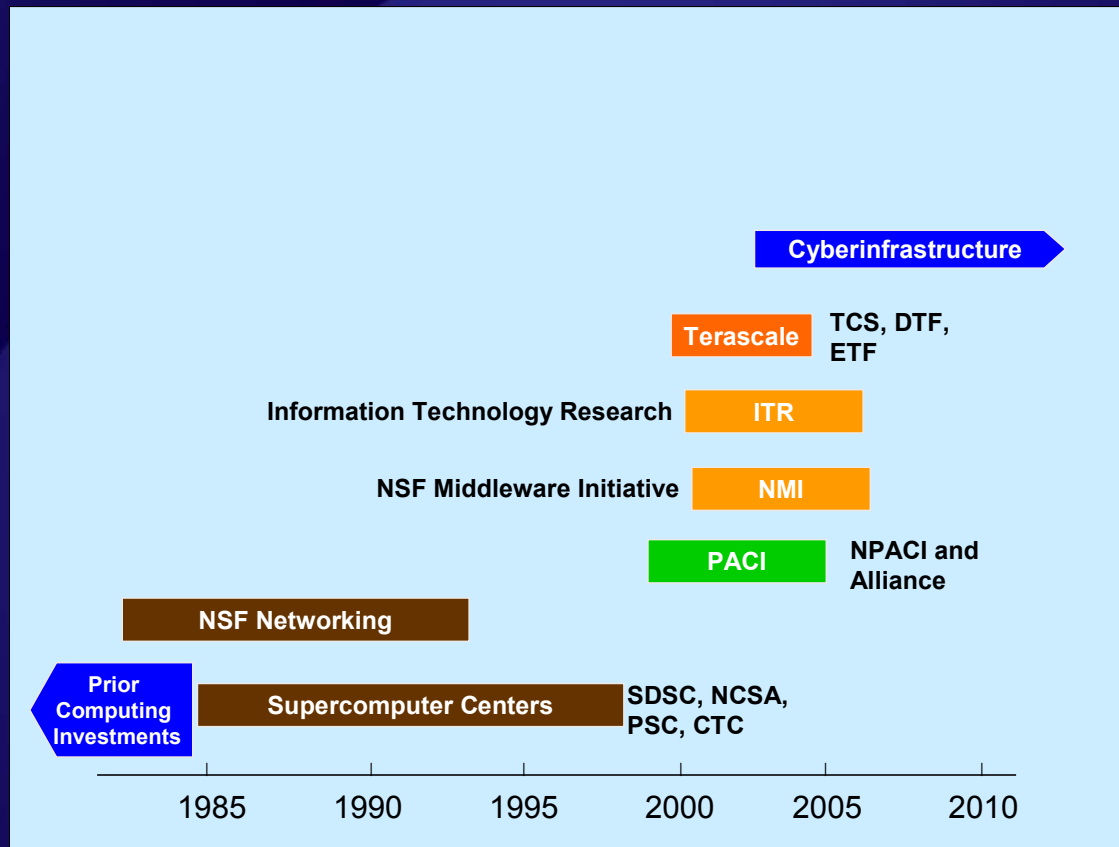
# Topics Covered Today

- Guiding Principles for Shared Cyberinfrastructure at NSF
- Enabling role of Cyberinfrastructure:
  - *Molecular Architecture as a New Frontier*
  - *Computational Steering as a New Capability*
- Looking to the Future
  - *Tipping Point: Information flow reversal*

# Guiding Principles for SCI at NSF

- Serve all of science & engineering
- Firm and continuing commitment to providing the most advanced cyberinfrastructure (CI), with high-end computing (HEC) at the core
- Encourage emerging CI while maintaining and transitioning extant CI
- Provide balance in CI equipment
- Strong links to ongoing fundamental research to create future generations of CI

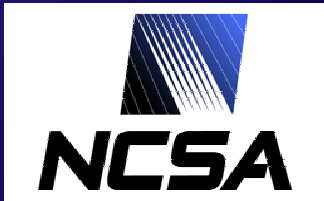
# History of NSF CI Investments



# Looking to the Future

- Science frontiers as the drivers
- Balance capability *and* capacity:
  - the Extensible Terascale Facility (ETF)
- Emerging CPU-intensive and data-intensive paradigms for molecular architecture as an illustrative example
- The next wave

# TeraGrid Partners



National Center for  
Supercomputing Applications



Center for Advanced  
Computing Research



*SDSC: A National Laboratory  
for Computational Science  
and Engineering*



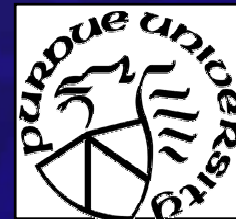
Argonne National Laboratory



Indiana University



Texas Advanced Computing Center

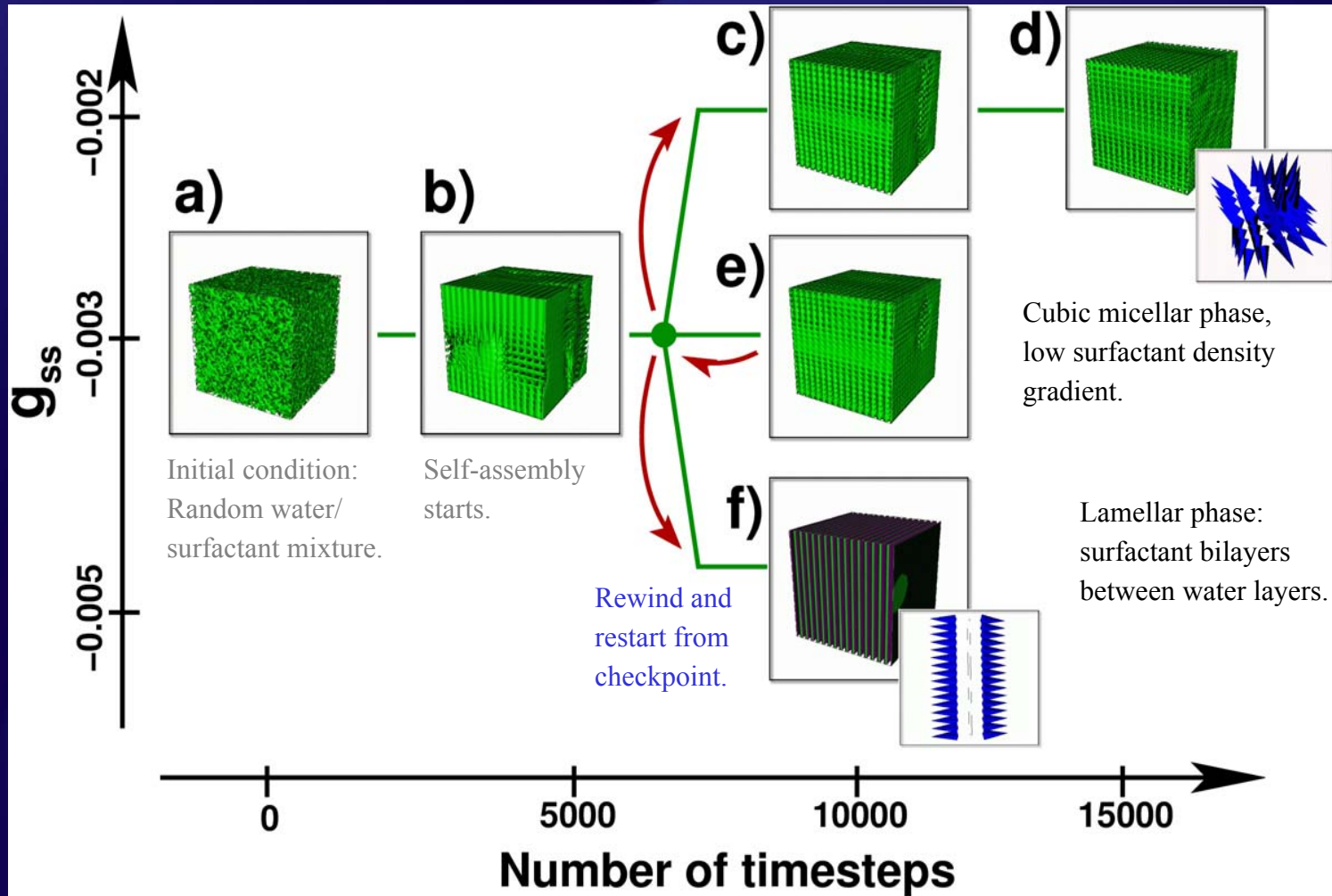


Purdue University

# TeraGyroids Project

- **Amphiphiles: hydrophobic tails and hydrophilic heads, dispersed in solvents or oil/water mixtures, self assemble into complex shapes; gyroids are of particular interest in biology**
- **Shapes from a parameters space:**
  - **Abundance, initial distribution of each component**
  - **strength of the surfactant-surfactant coupling,**
- **Desired structures simulated only in very large systems**
- **Project goal is to study defect pathways and dynamics in gyroid self-assembly**

# Exploring parameter space through computational steering





# Cyberinfrastructure: the future consists of ...

- Computational engines (supercomputers, clusters, workstations – capability and capacity)
- Mass storage (disk drives, tapes, ...) and persistence
- Networking (including optical, wireless, ubiquitous)
- Digital libraries/data bases
- Sensors/actuators
- Software (operating systems, middleware, domain specific tools/platforms for building applications)
- Services (education, training, consulting, user assistance)

***All working together in an integrated fashion.***

# Cyberinfrastructure: Tipping Point

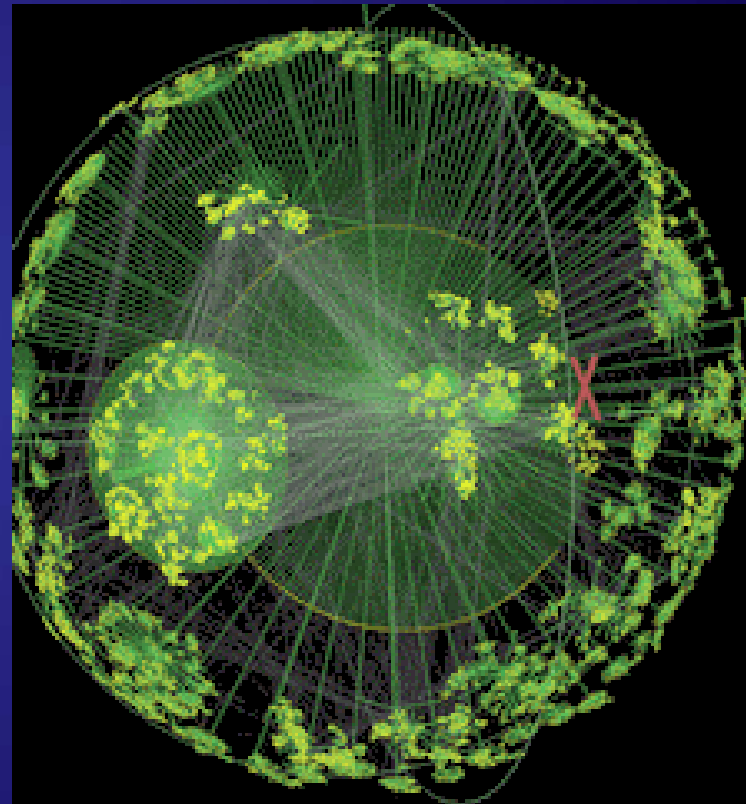
## *Information Flow Reversal*

### Internet Historical Roots

Create & Compute at the core; then broadcast to the periphery

### The Next Wave

Massive data generated at the periphery; novel systems & architectures; revitalized core of high-end computers

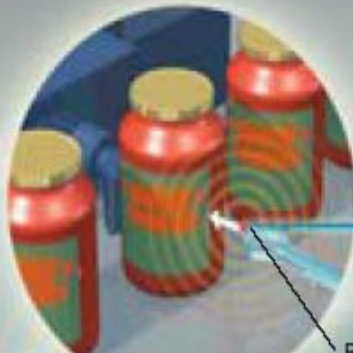


Topological view of the Internet

# Sensor-Nets and the "new" Bar Code

RFID SYSTEMS will let products, such as the hypothetical Mama's tomato sauce, be monitored and reordered automatically, once depleted. Even an individual item will be traced. Here is one scenario.

**1** An RFID tag gets affixed to each can of Mama's tomato sauce as it passes by on a conveyor belt. A reader detects the tag's unique identifier code and stores it in a list ready to be sent to a central database.



RFID reader

RFID tag being affixed

**3** When scanned by the reader at a factory, each box—and each jar in the box—responds sequentially with its identifier code. The plant sends lists of codes, both from the conveyor belt and the packed boxes, to an Internet-based computer system (blue arrows), where they are stored in a database that ties each jar, box and pallet to the originating factory.

Mama's  
Sauces

SEND TO DISTRIBUTION CENTER

**2** A pallet containing boxes loaded with cans of Mama's sauce is prepared for transport.



Scientific American  
Jan. 2004 issue,  
article on RFID  
by R. Want

# Closing Remarks

- Enabling role of CI for S&E Research is the same paradigm for the transformative power of the new wave of the “e” revolution: immediate economic and societal impact
- Massive data generation at the periphery, HEC at the core, and a new architecture linking the core to the periphery - these are the central elements of CI
- A strategy of a balanced and broad CI to serve all of science and engineering; transition from extant CI to the exciting possibilities of future CI