Cyberinfrastructure for Discovery Science and Engineering

Edward Seidel

Incoming Director, Office of Cyberinfrastructure (Sept 1, 2008)

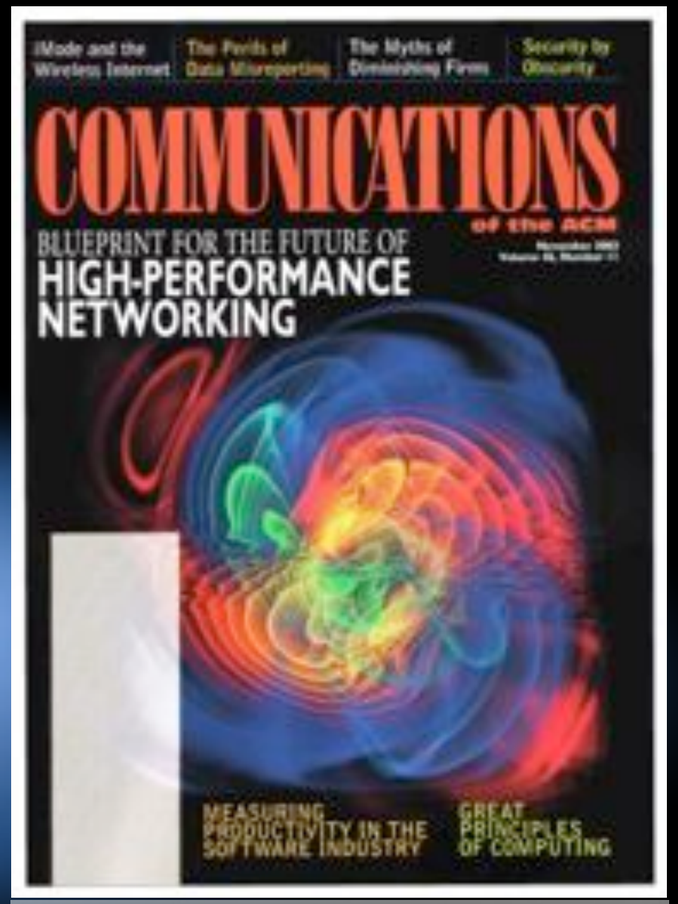
hseidel@nsf.gov

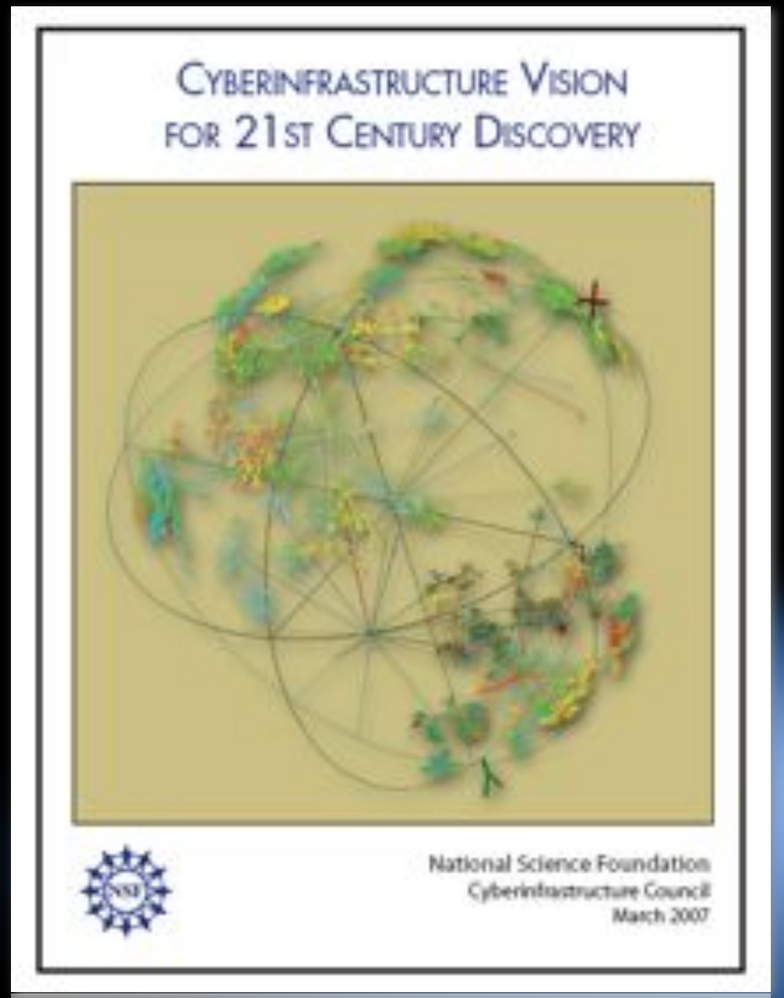
Science is being Transformed

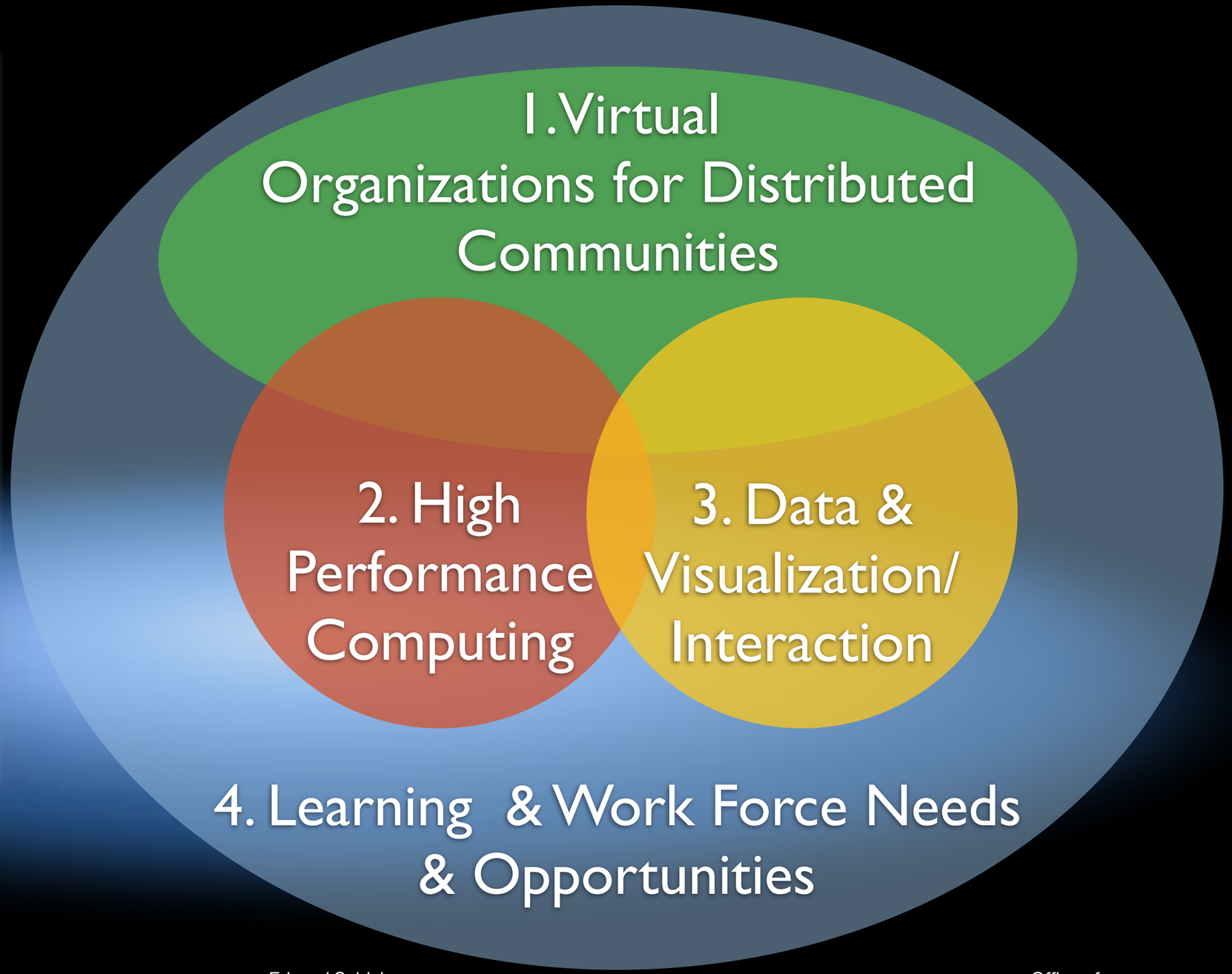
- Complex problems require totally new methodologies
  - Large scale computing beyond current experience/understanding
  - Large scale data, beyond...
  - Large scale collaborations, beyond...
- Good news!
  - Cyberinfrastructure!
  - So different, it somewhat levels playing field for EPSCOR states!

What is Cyberinfrastructure?

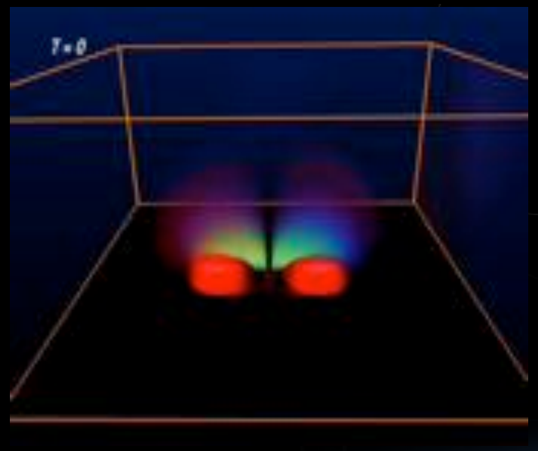
- Comprehensive *environment* for complex problem solving
  - Computers, networks, data archives, sensors, software, algorithms...
  - It is often about building communities and enabling research
- General
  - LONI, SURAgrid, OSG, TeraGrid, Great Lakes Consortium
- Discipline Specific
  - NEON, NEES, SCOOP, UCOMS




CYBERINFRASTRUCTURE VISION
FOR 21ST CENTURY DISCOVERY
National Science Foundation
Cyberinfrastructure Council
March 2007

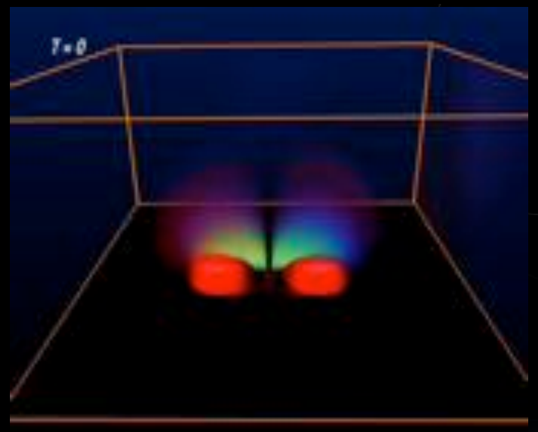

1. Virtual Organizations for Distributed Communities
2. High Performance Computing
3. Data & Visualization/ Interaction
4. Learning & Work Force Needs & Opportunities

Collaborations, CI for Complex Problems

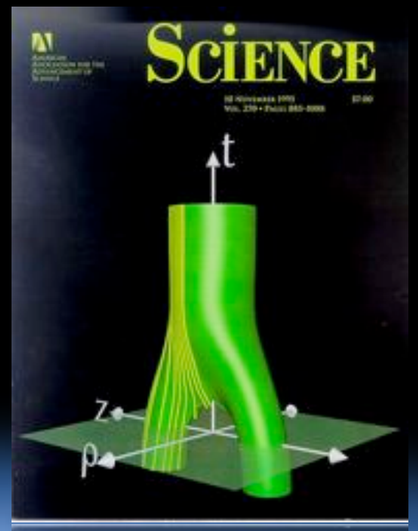


- **NASA Neutron Star Grand Challenge**
  - 5 US Institutions, 3years, $1.4M
  - Solve problem of colliding neutron stars (try…)

Collaborations, CI for Complex Problems

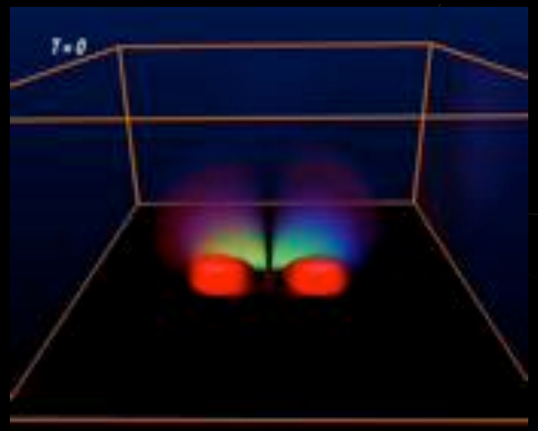


- **NASA Neutron Star Grand Challenge**
  - 5 US Institutions, 3years, $1.4M
  - Solve problem of colliding neutron stars (try…)

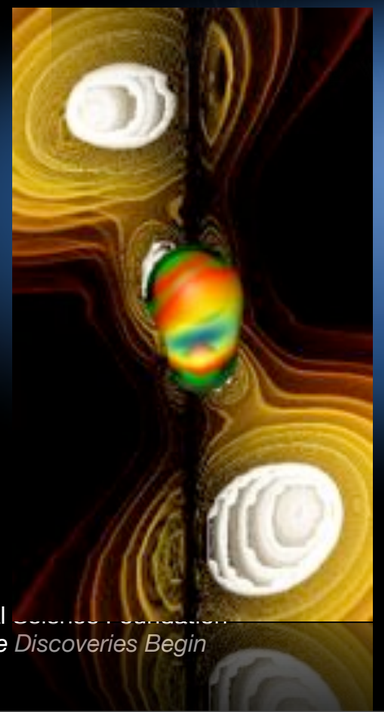


- **NSF Black Hole Grand Challenge**
  - 8 US Institutions, 5 years, $4M
  - Solve problem of colliding black holes (try…)

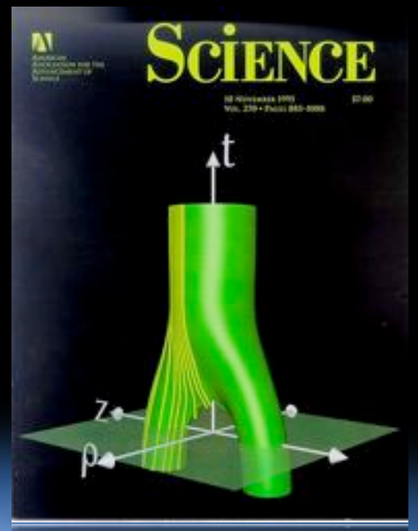
Collaborations, CI for Complex Problems



- **NASA Neutron Star Grand Challenge**
  - 5 US Institutions, 3years, $1.4M
  - Solve problem of colliding neutron stars (try…)



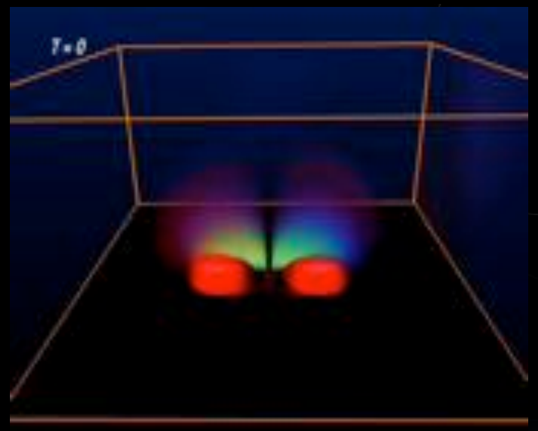
- **EU Network Astrophysics**
  - 10 EU Institutions, 3 years, €1.5M
  - Continue these problems
  - *Entire Community becoming CI-enabled*



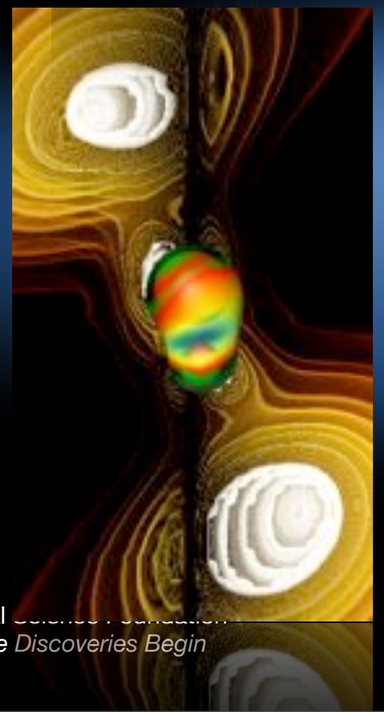
- **NSF Black Hole Grand Challenge**
  - 8 US Institutions, 5 years, $4M
  - Solve problem of colliding black holes (try…)

Edward Seidel
hseidel@nsf.gov

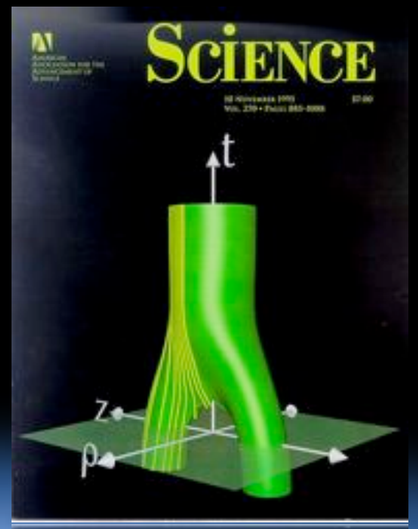
Collaborations, CI for Complex Problems



- **NASA Neutron Star Grand Challenge**
  - 5 US Institutions, 3years, $1.4M
  - Solve problem of colliding neutron stars (try…)



- **EU Network Astrophysics**
  - 10 EU Institutions, 3 years, €1.5M
  - Continue these problems
  - *Entire Community becoming CI-enabled*



- **NSF Black Hole Grand Challenge**
  - 8 US Institutions, 5 years, $4M
  - Solve problem of colliding black holes (try…)

- **Lessons Learned**
  - All disciplines strive to solve complex problems, struggle to build communities & collaboration
  - Require cross disciplinary teams
  - Place demands on software: *abstractions, shared CI*
  - Are socially demanding!

Edward Seidel
hseidel@nsf.gov

Another Complex Problem: Katrina

Much of same expertise, shared infrastructure needed

Louisiana as CI Testbed

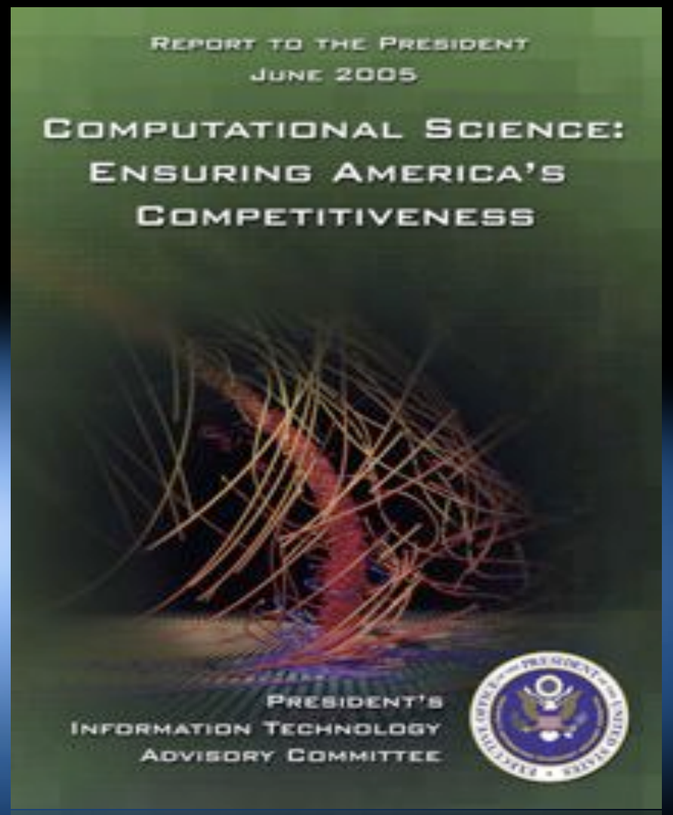
*A Case Study of **Shared** CI*

- 4 Layers: Each reinforces the next
  - 2001-3: Vision 2020
    - Gov. Foster: $25M annual across 5 campuses
    - "IT = Economic development"
    - 2003: CCT, $9M annual at LSU, LITE Center, ULL
  - 2004: LONI and its impact
    - Gov. Blanco: $40M + $10M infrastructure
  - 2007: CyberTools
    - $12M Statewide NSF/EPSCOR CI project
    - Many other multisite projecs
  - 2007: LONI Institute and beyond
    - $15M Statewide project to recrtuit people

Catalyzed many new hires, funding, additional investments

Collaborations for Complex Problems
PITAC Report Summary

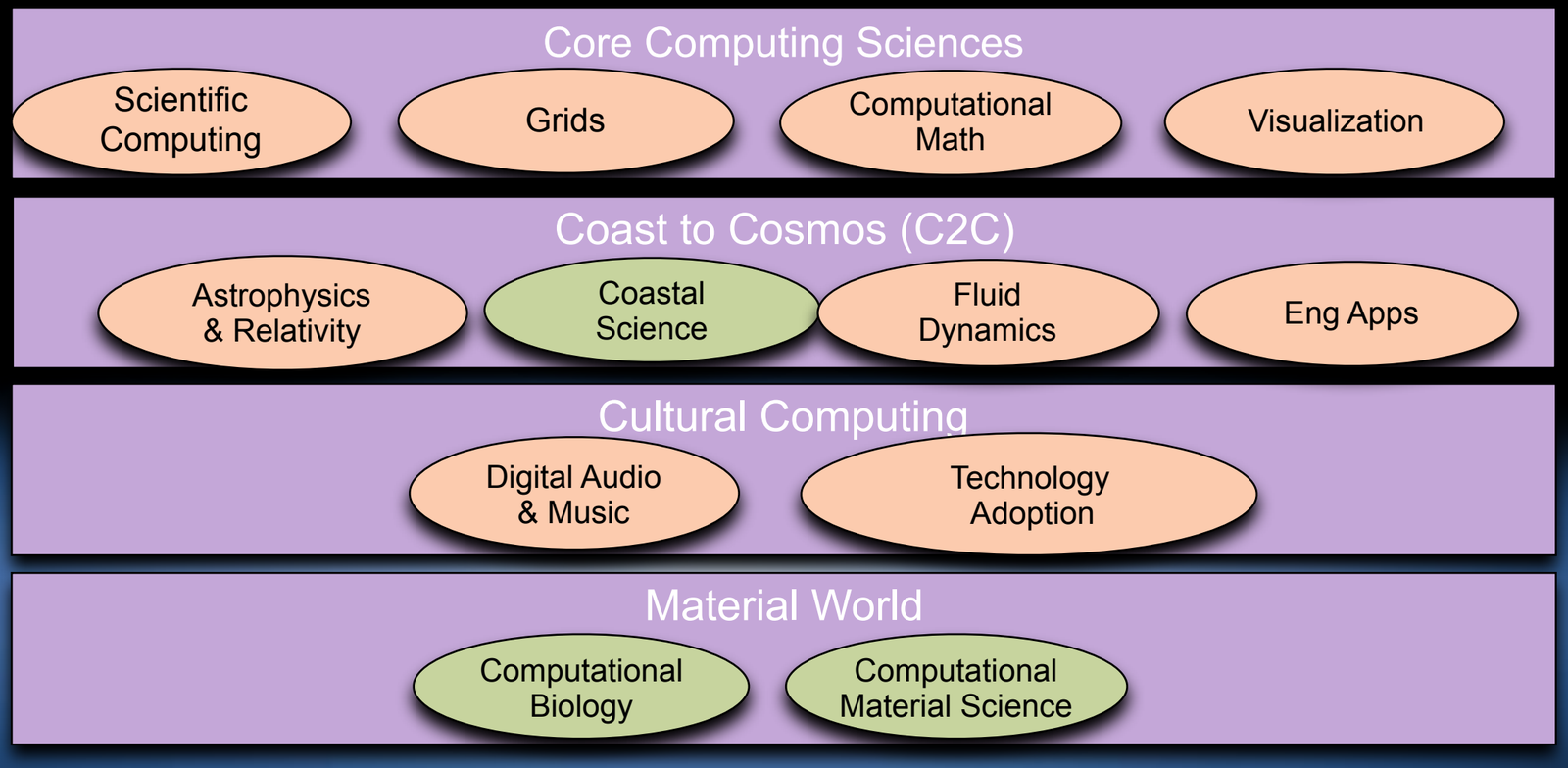
"Universities must significantly change organizational structures: multidisciplinary & collaborative research are needed [for US] to remain competitive in global science"



Layer 1: CCT - Collab for Complex Problems

Building a center through CI-enabled collaboration

This is a model for NSF... Directorates come together through CI: CDI, Grand Challenges, etc

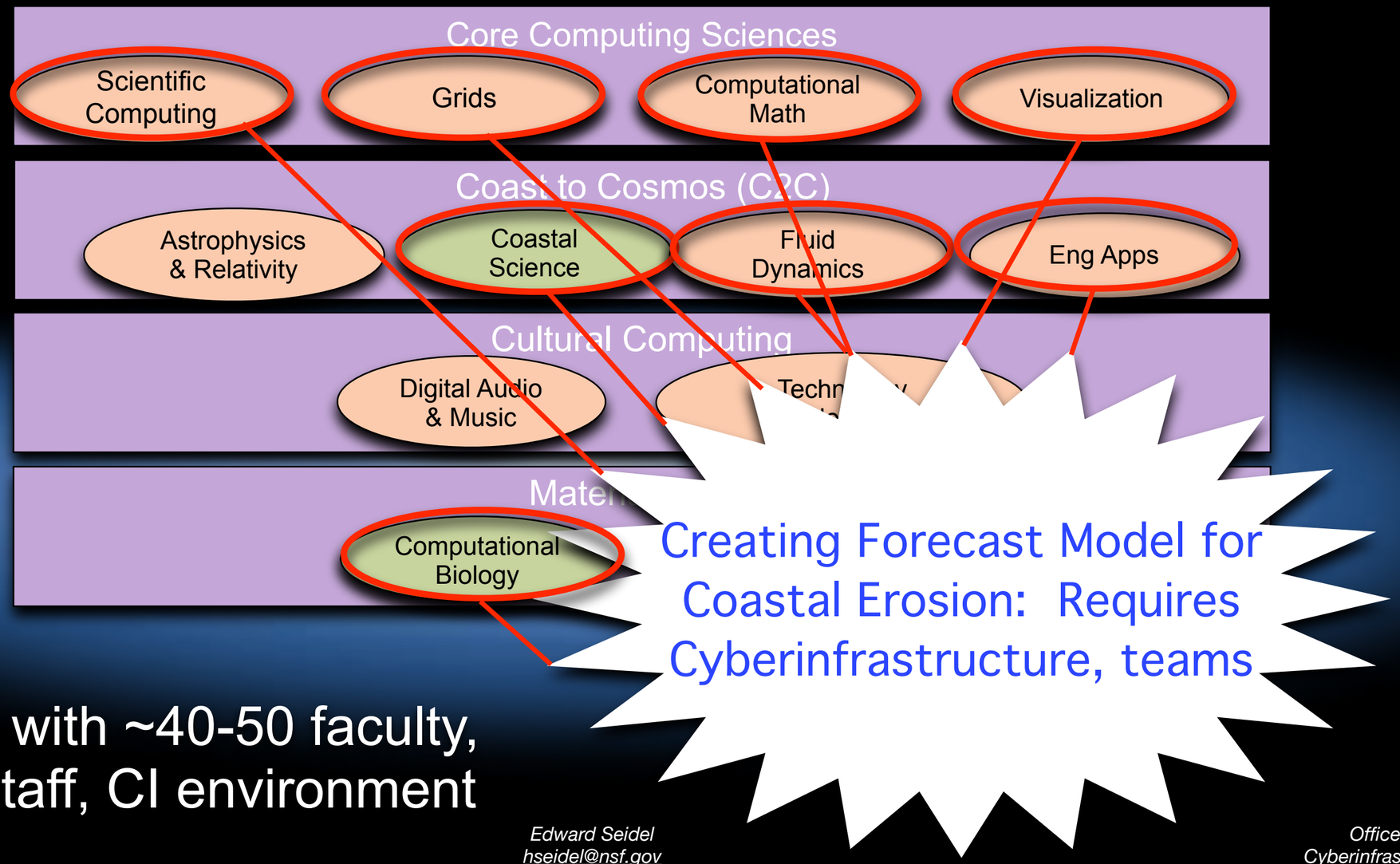


Building center with ~40-50 faculty, ~20 research staff, CI environment

Layer 1: CCT - Collab for Complex Problems

Building a center through CI-enabled collaboration

This is a model for NSF... Directorates come together through CI: CDI, Grand Challenges, etc

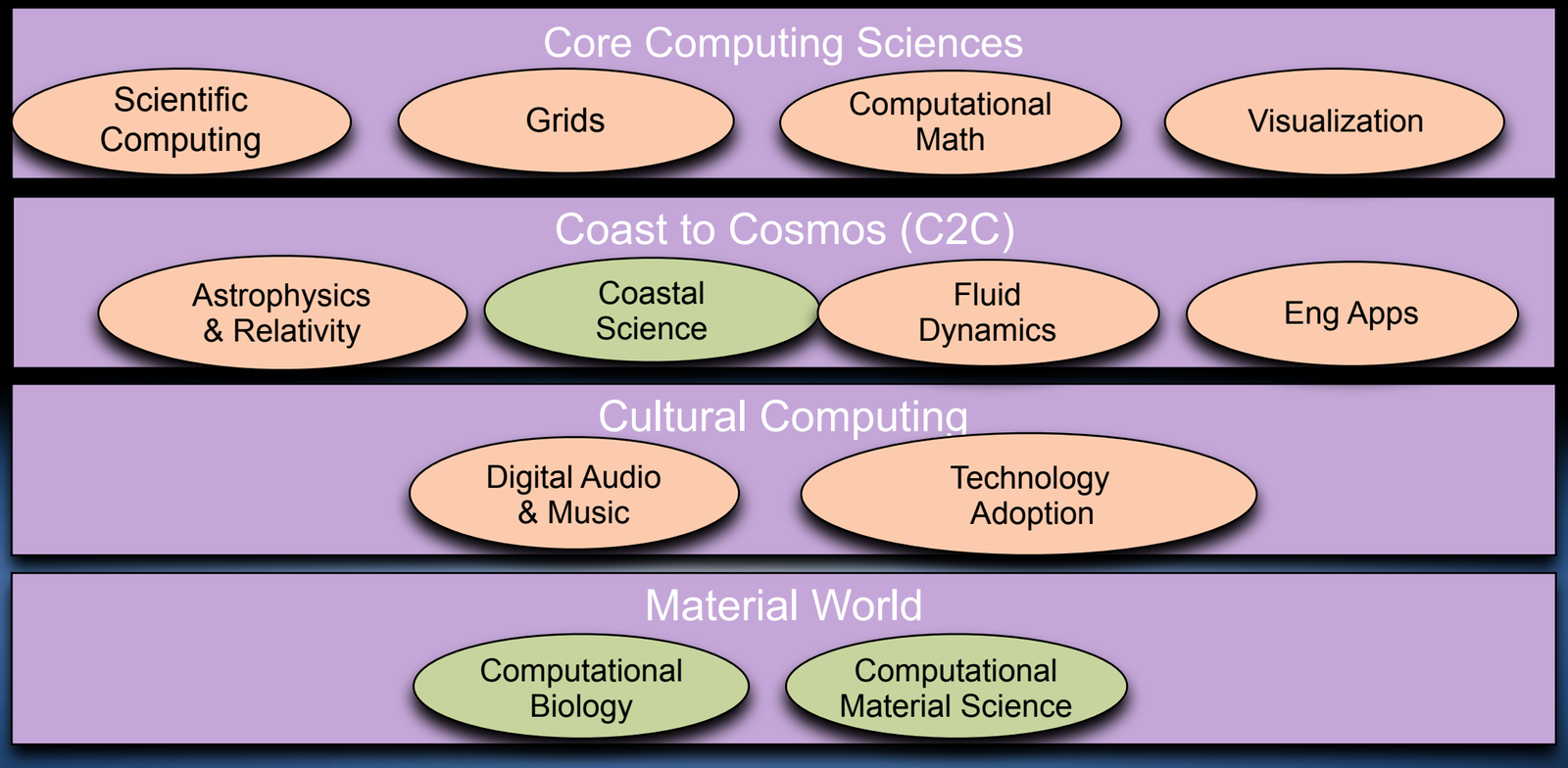


Building center with ~40-50 faculty, ~20 research staff, CI environment

Layer 1: CCT - Collab for Complex Problems

Building a center through CI-enabled collaboration

This is a model for NSF... Directorates come together through CI: CDI, Grand Challenges, etc

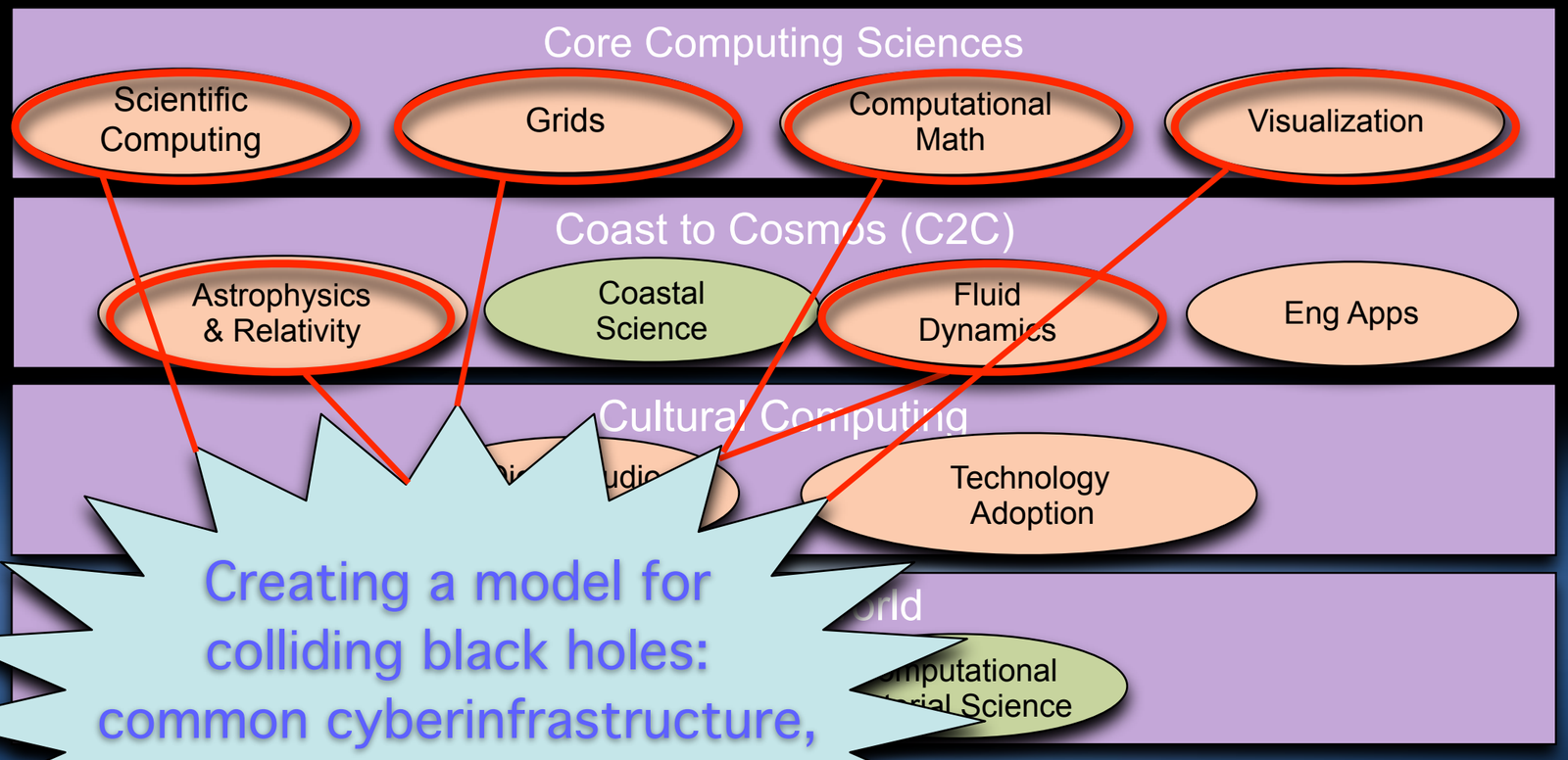


Building center with ~40-50 faculty, ~20 research staff, CI environment

Layer 1: CCT - Collab for Complex Problems

Building a center through CI-enabled collaboration

This is a model for NSF... Directorates come together through CI: CDI, Grand Challenges, etc



Creating a model for colliding black holes: common cyberinfrastructure, teams (many of the same team members!)

Building center with
~20 research staff, CI

Layer 2: LONI

LA Tech

LSU

SUBR

UNO

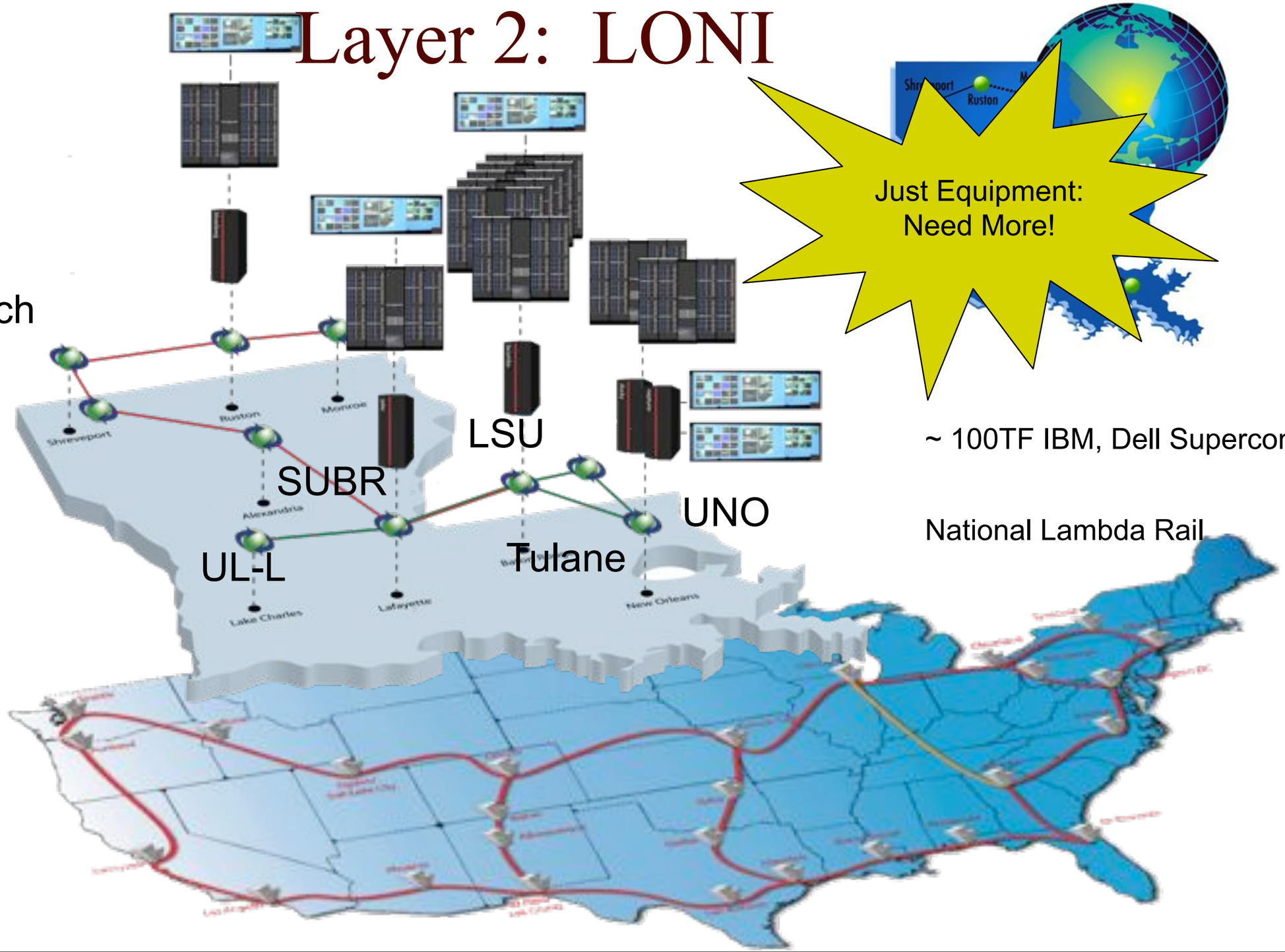
UL-L

Tulane

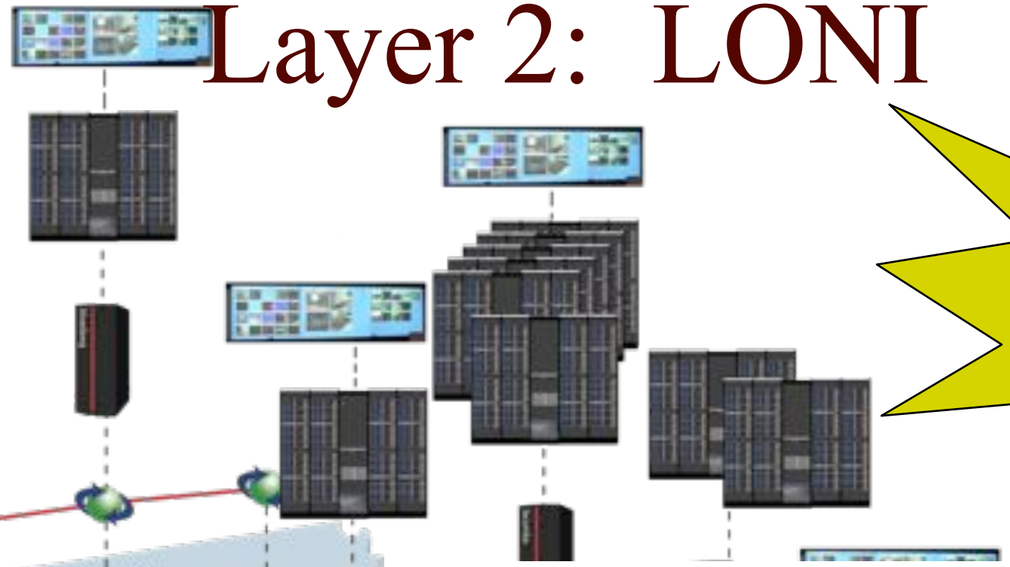
~ 100TF IBM, Dell Supercomputers

National Lambda Rail

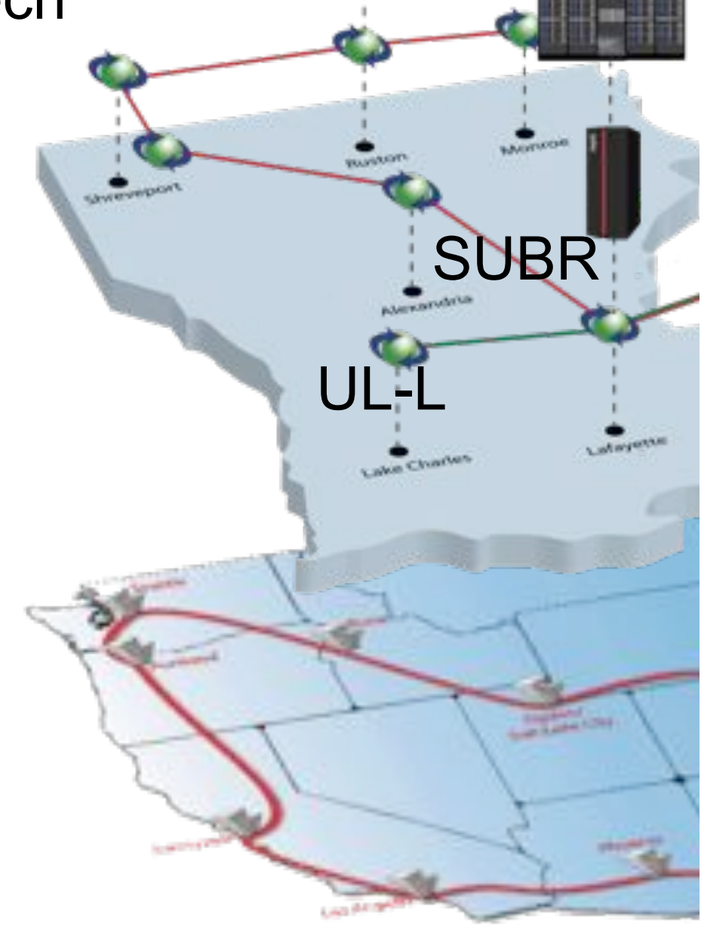
Layer 2: LONI

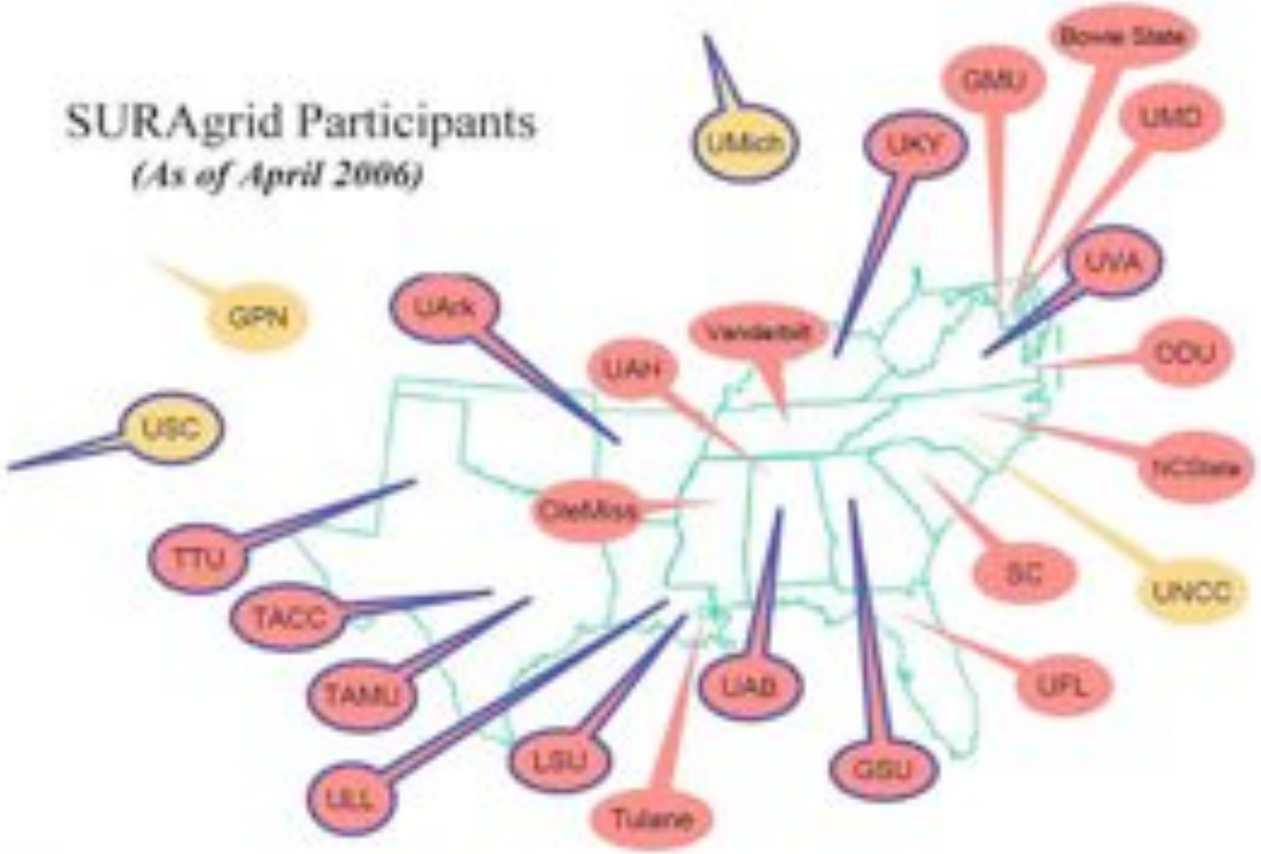

Just Equipment:
Need More!
LA Tech
LSU
SUBR
UNO
UL-L
Tulane
Shreveport
Ruston
Monroe
Alexandria
Lake Charles
Lafayette
New Orleans
~ 100TF IBM, Dell Supercomputers
National Lambda Rail

Layer 2: LONI



LA Tech





iters

LONI Creates a Regional Environment...

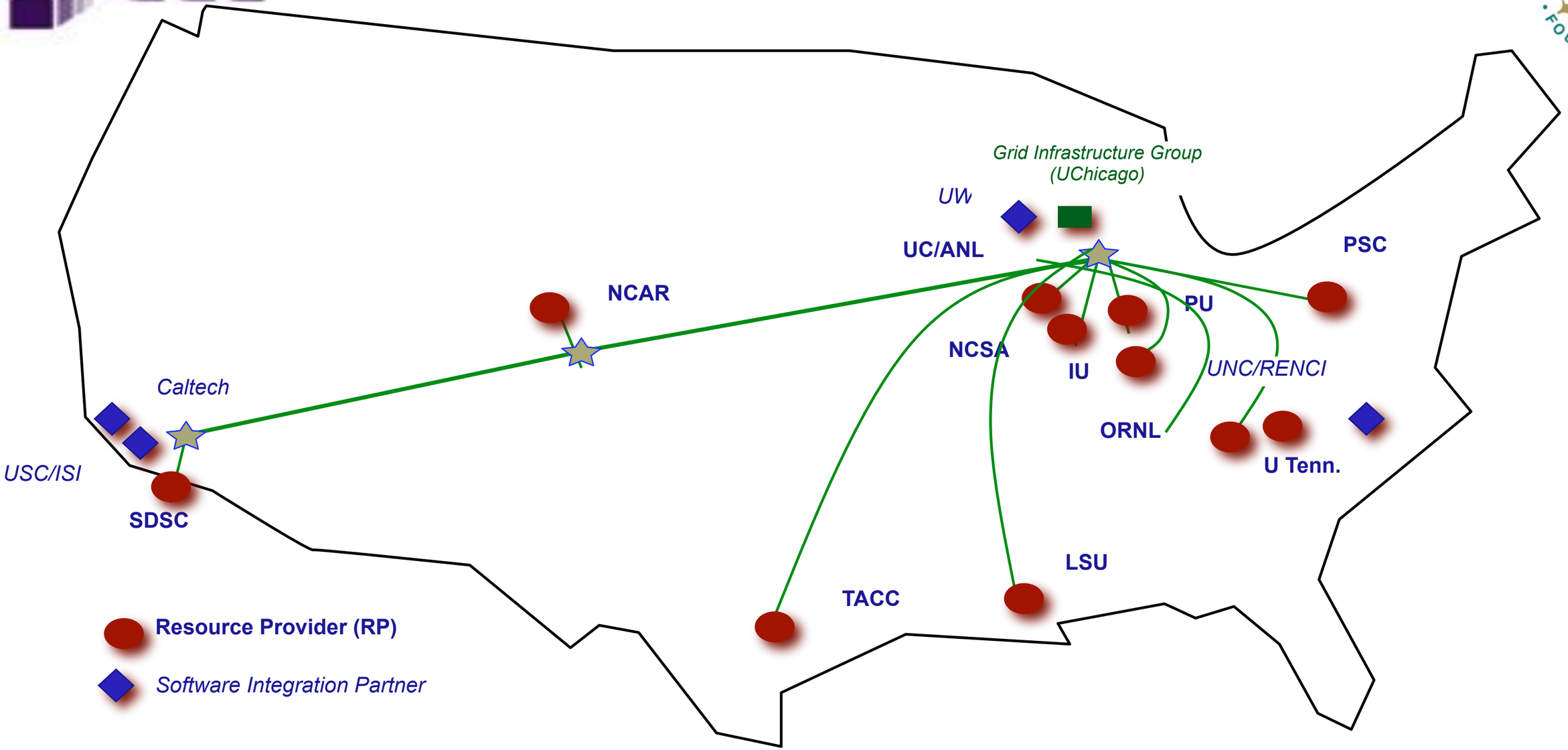
- Suddenly collaborative, competitive
  - Meetings, collaborations between universities: most important!
- Many new multisite projects built on top of LONI
  - NSF ESPCOR RII: CyberTools
    - Common software infrastructure for MD, CFD, Experiment
    - Viz services, data services, computing services, co-scheduling, portals
  - DOE EPSCOR: UCoMS
  - NSF MRI: PetaShare, VizTangibles
  - NIH LBRN, Enlightened, HD Classes, etc...
  - BoR PKSFI: Security CyberSpace Center, LONI Institute
- Prominent Hires
  - Thomas Sterling (LSU), Carolina Cruz-Neira (ULL), others

...Aligned with National CI

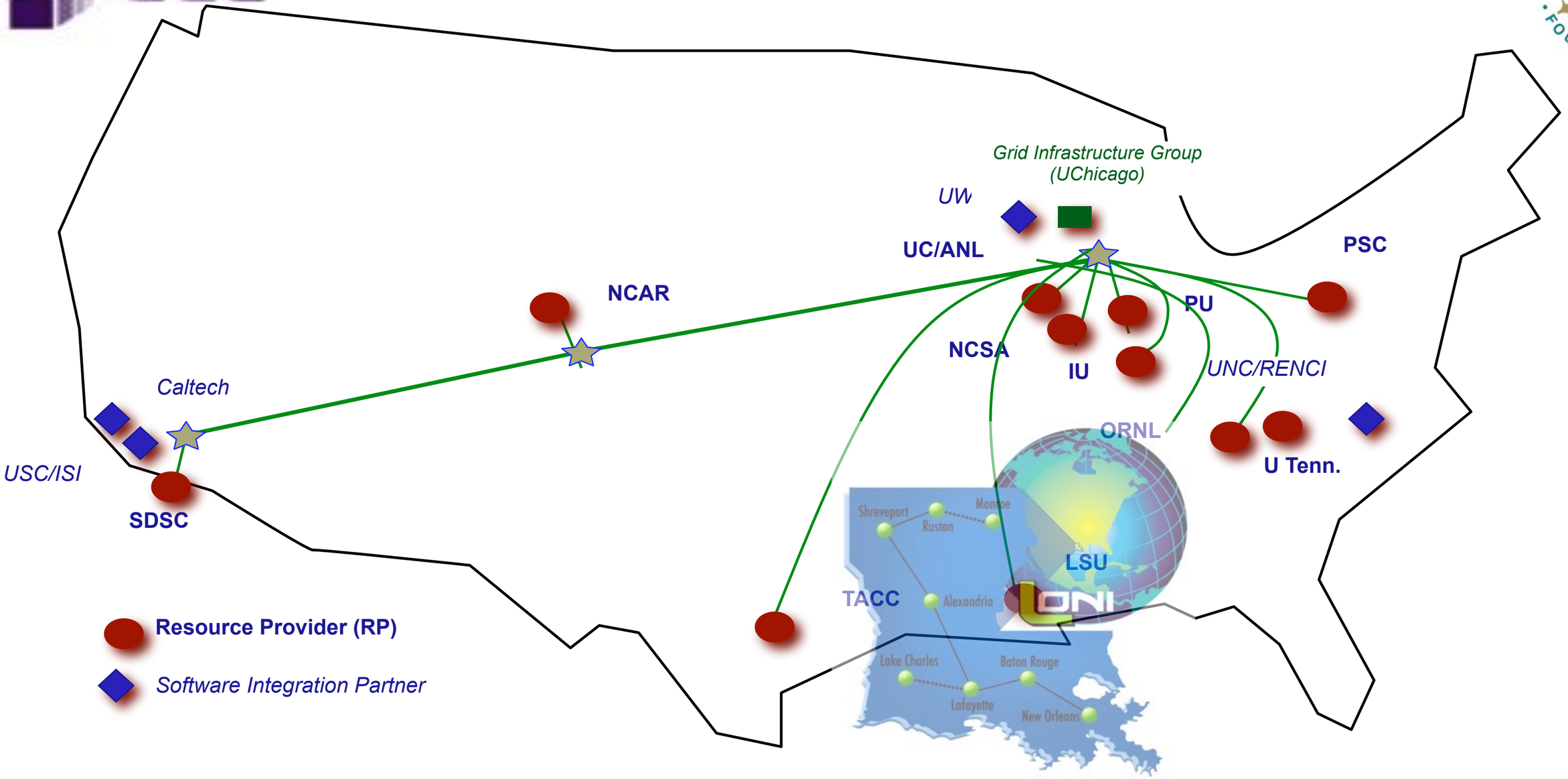
- NSF $2.2M HPCOPS Award
  - LONI joins TeraGrid (Jan08)
  - 7 new HPC staff
  - NCSA subcontracted to support
- NSF $208M Track 1 PetaScale Award
  - LSU partners with NCSA; 10PF-class computer
  - LSU focus on applications, tools, outreach
- State submitted several NSF Track 2c proposals

We are working to ensure that Louisiana researches can take advantage of all this!

LONI and TeraGrid

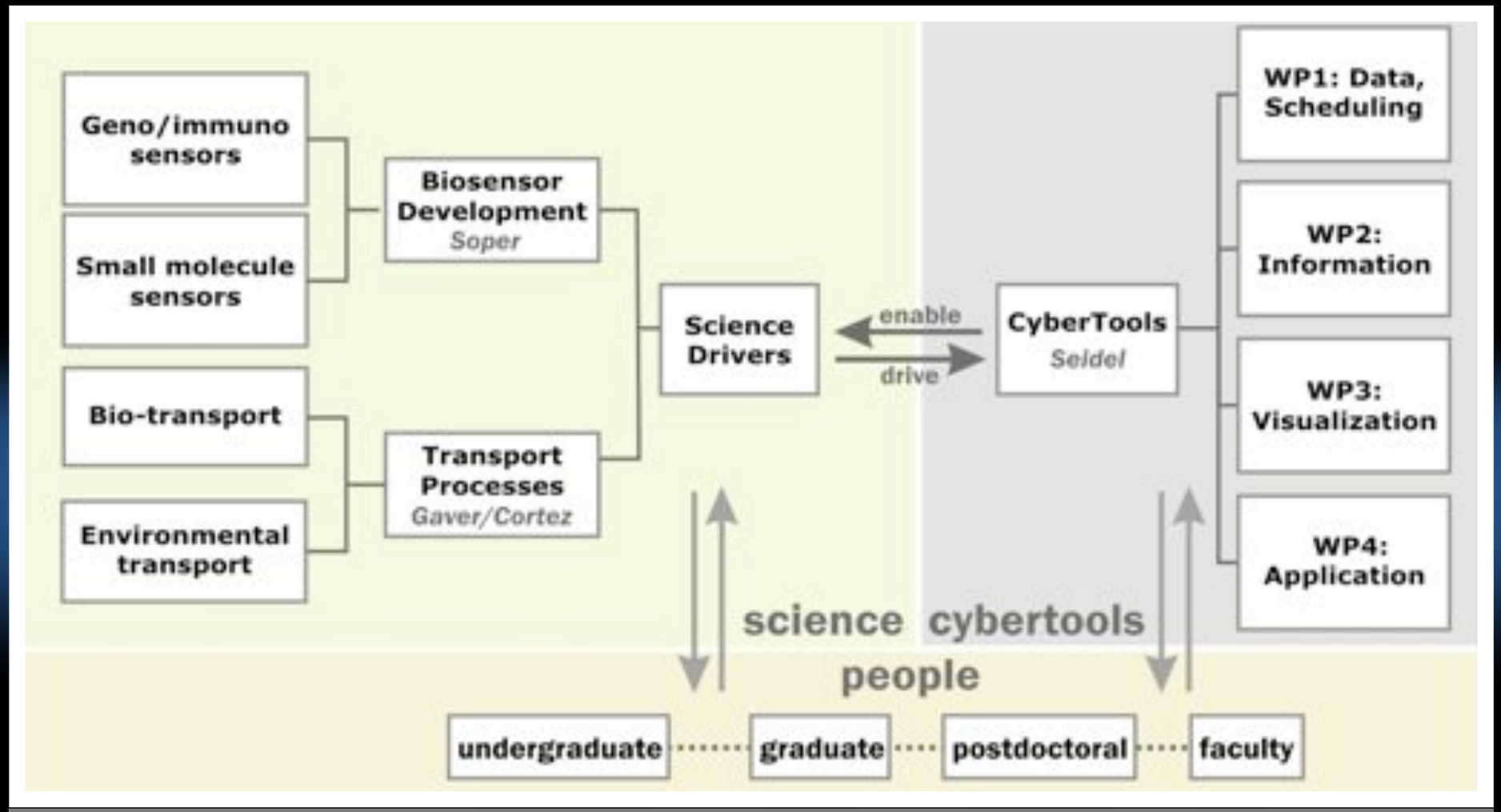


LONI and TeraGrid



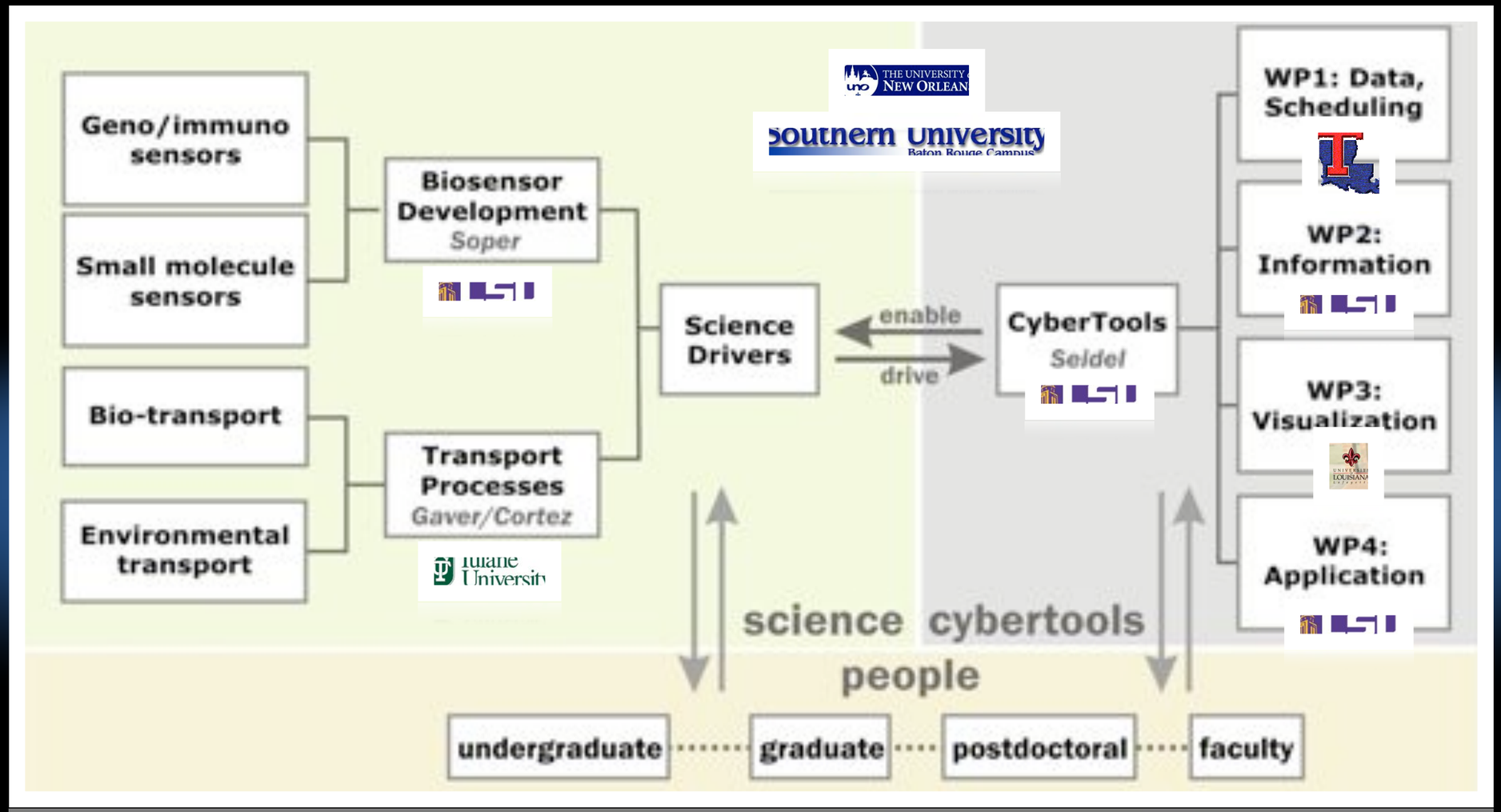
Layer 3: CyberTools

$12M NSF/LA BoR EPSCOR RII Grant



Layer 3: CyberTools

$12M NSF/LA BoR EPSCOR RII Grant

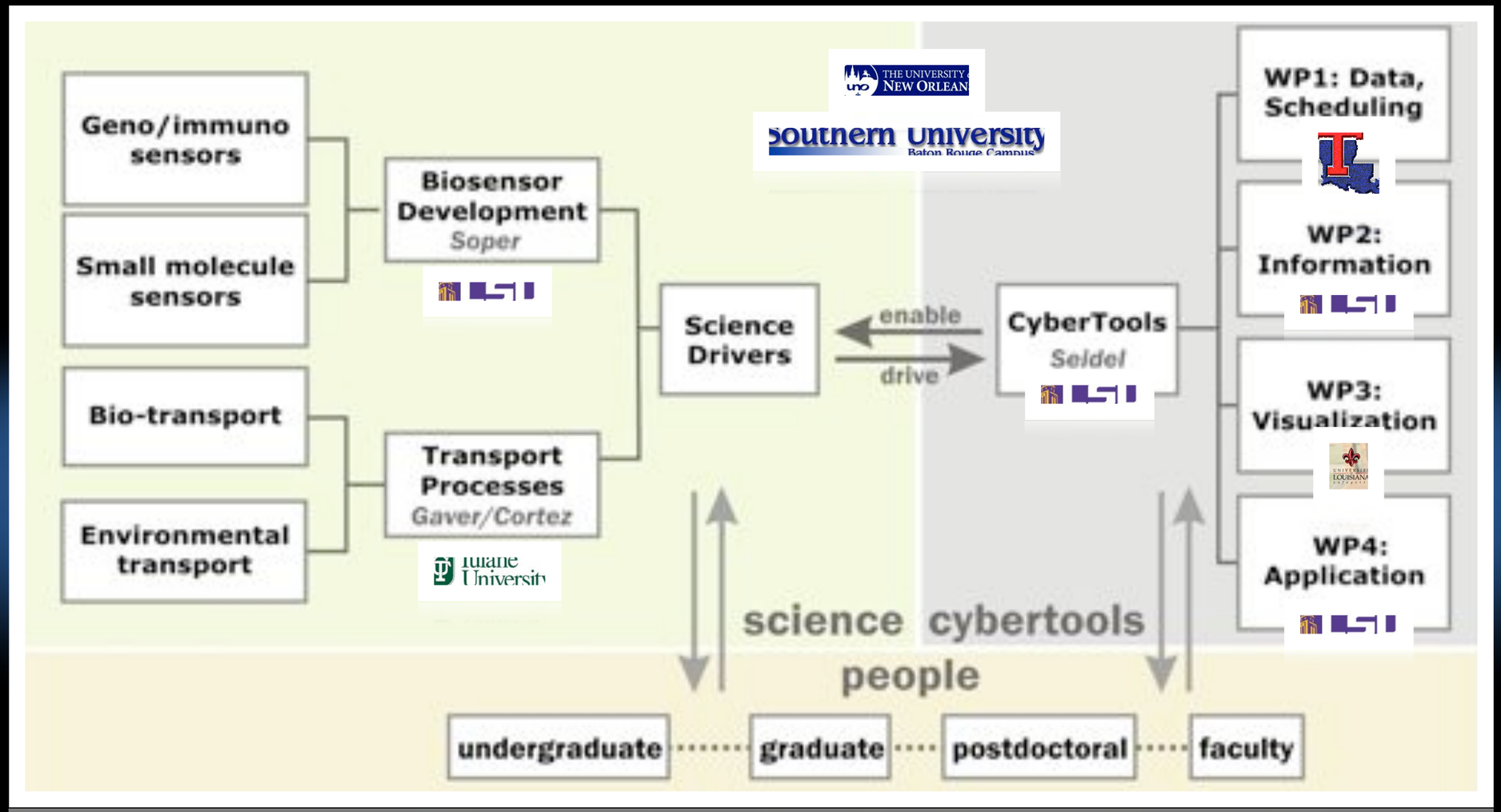


Layer 3: CyberTools

$12M NSF/LA BoR EPSCOR RII Grant

Layer 3: CyberTools

$12M NSF/LA BoR EPSCOR RII Grant



Layer 4: LONI Institute

- $15M, 5-year project
  - $7M BoR, $8M from LA Tech, LSU, SUBR, Tulane, UNO, UL-Lafayette
  - 12 new faculty, 18 grads, 6 CI staff!
- Built on LONI infrastructure, create bold new inter-university superstructure
  - Faculty, staff, students focus on CS, Bio, Materials, but all disciplines impacted; digital art & media!
  - Promote collaborative research at interfaces for innovation
- Draws on, enhance strengths of all universities
  - Strong groups recently created; collectively world-class
  - Much stronger recruiting opportunities for all institutions
- Create University-Industry Research Centers (UIRCs)
  - Economic development!

How does this align with NSF programs in Cyberinfrastructure?

Edward Seidel
hseidel@nsf.gov

Shared Resource Environments

Courtesy of University of Indiana

Shared Resource Environments

Computers

Data services

Visualization services

People

Courtesy of University of Indiana

Shared Resource Environments

Computers

Data services

Visualization services

People

Courtesy of University of Indiana

Modeling and simulation

Data analysis & visualization

User support

Training

Common user environments

Tools for educators

Science Gateways

Shared Resource Environments

Computers

Data services

Visualization services

People

Modeling and simulation

Data analysis & visualization

User support

Training

Common user environments

Tools for educators

Science Gateways

Courtesy of University of Indiana

Shared Resource Environments

Computers

Data services

Visualization services

People

Modeling and simulation

Data analysis & visualization

User support

Training

Common user environments

Tools for educators

Science Gateways

Courtesy of University of Indiana

Shared Resource Environments

Computers

Data services

Visualization services

People

Modeling and simulation

Data analysis & visualization

User support

Training

Common user environments

Tools for educators

Science Gateways

Courtesy of University of Indiana

Shared Resource Environments

Computers

Data services

Visualization services

People

Modeling and simulation

Data analysis & visualization

User support

Training

Common user environments

Tools for educators

Science Gateways

Courtesy of University of Indiana

Shared Resource Environments

Computers

Data services

Visualization services

People

Modeling and simulation

Data analysis & visualization

User support

Training

Common user environments

Tools for educators

Science Gateways

Courtesy of University of Indiana

Shared Resource Environments

Computers

Data services

Visualization services

People

LA parallels TeraGrid locally

Courtesy of University of Indiana

Modeling and simulation

Data analysis & visualization

User support

Training

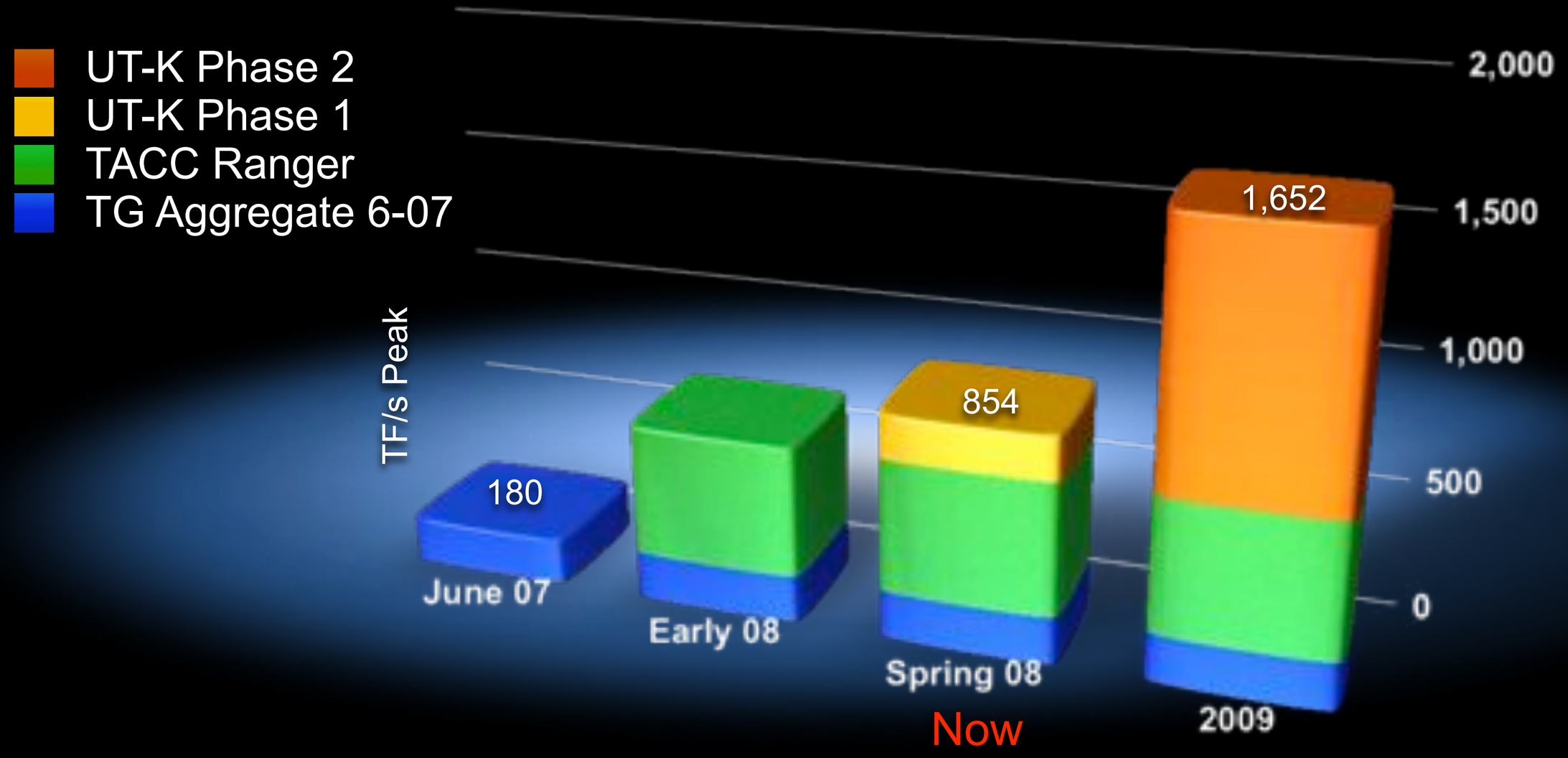
Common user environments

Tools for educators

Science Gateways

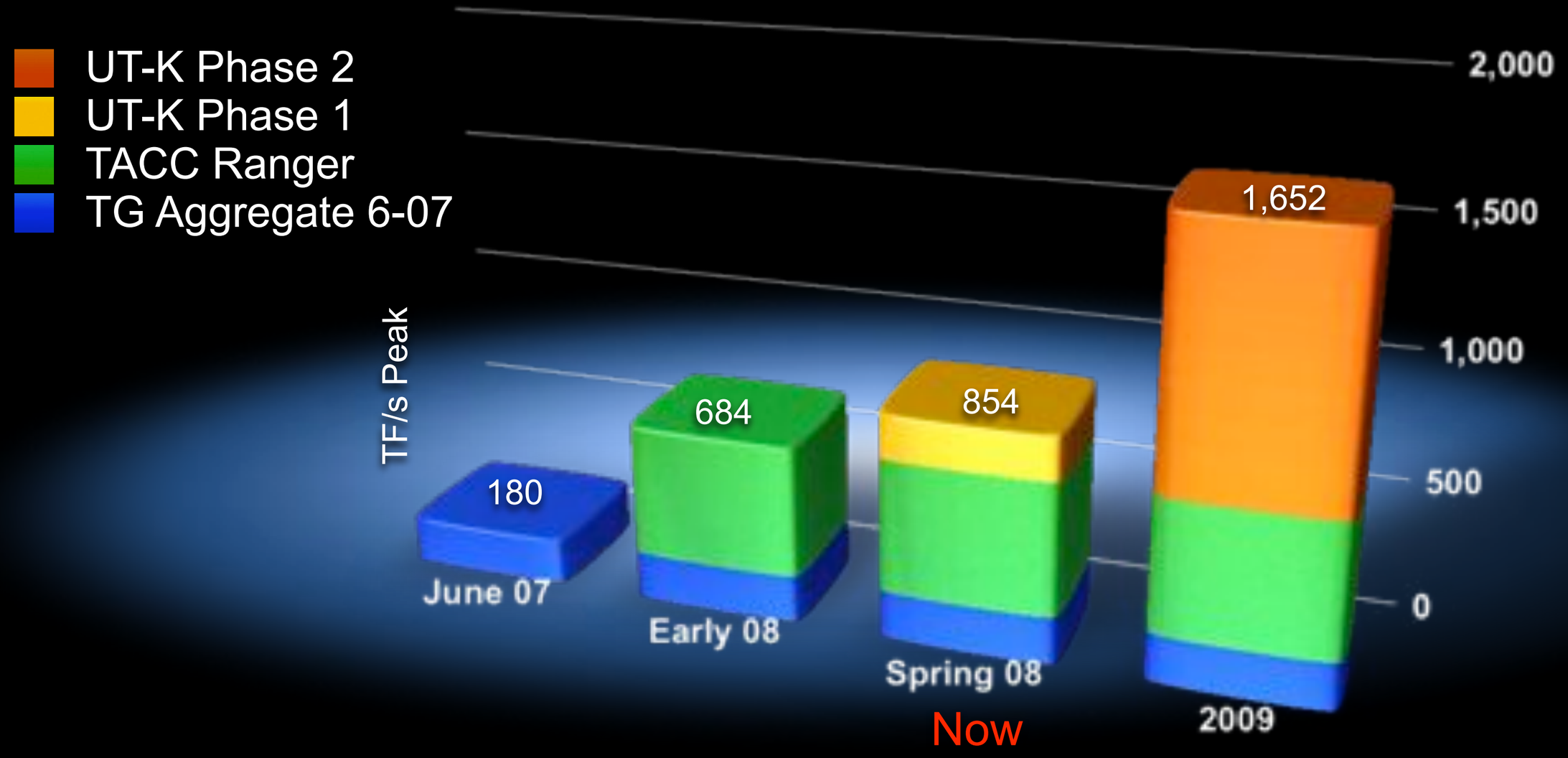
Impact

Greatly expanding capacity of the TeraGrid for digital exploration with reduced oversubscription and queue wait times.



Impact

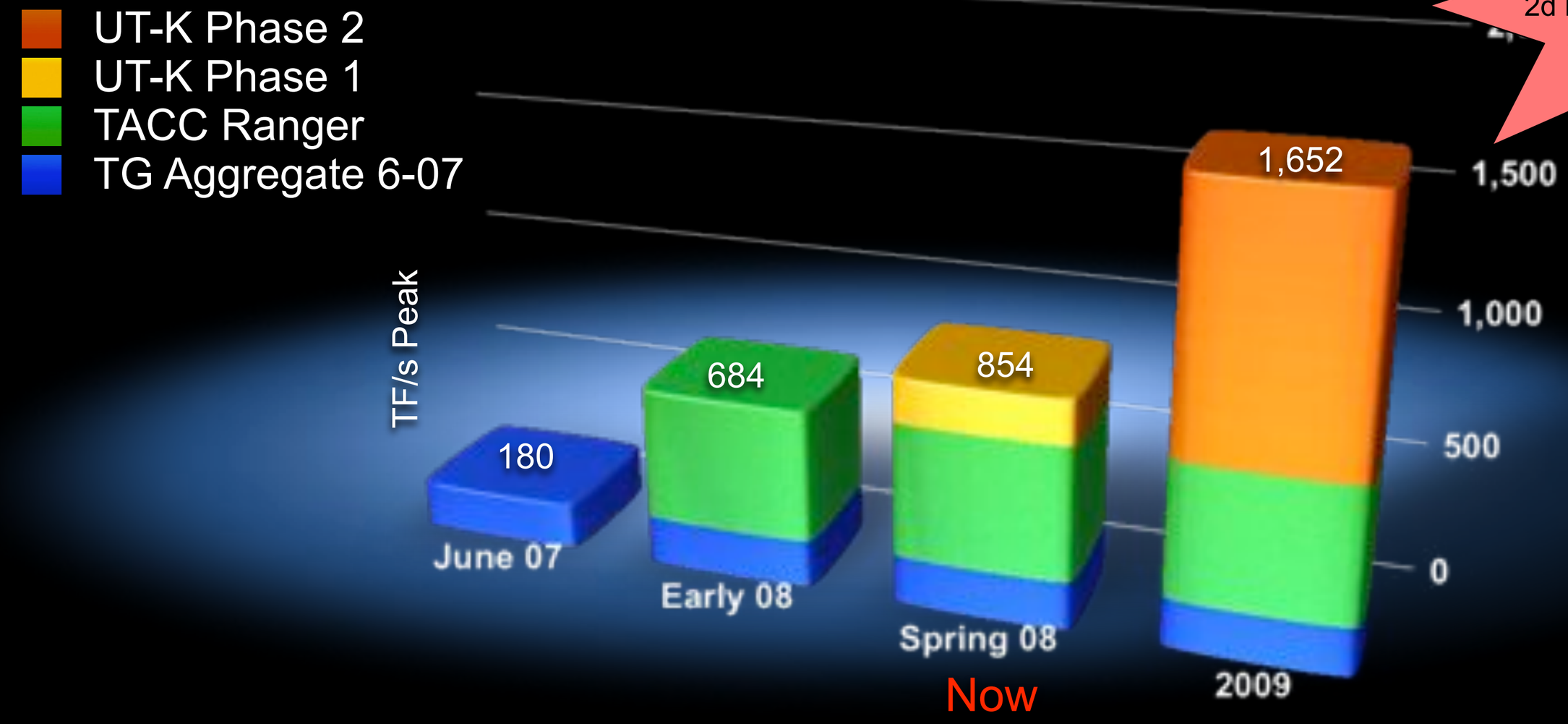
Greatly expanding capacity of the TeraGrid for digital exploration with reduced oversubscription and queue wait times.



Impact

Greatly expanding capacity of the TeraGrid for digital exploration with reduced oversubscription and queue wait times

Track1, Track 2d bring this up by almost 10x!



Track 2d and TeraGrid Phase III (XD)

Solicitations released June 2008

Track 2d (NSF 08-573)

- Aimed at expanding range of TeraGrid portfolio
- Focuses on data-intensive computing and more experimental systems
- Anticipate more than one award

TeraGrid Phase III (eXtreme Digital - XD) (NSF 08-571)

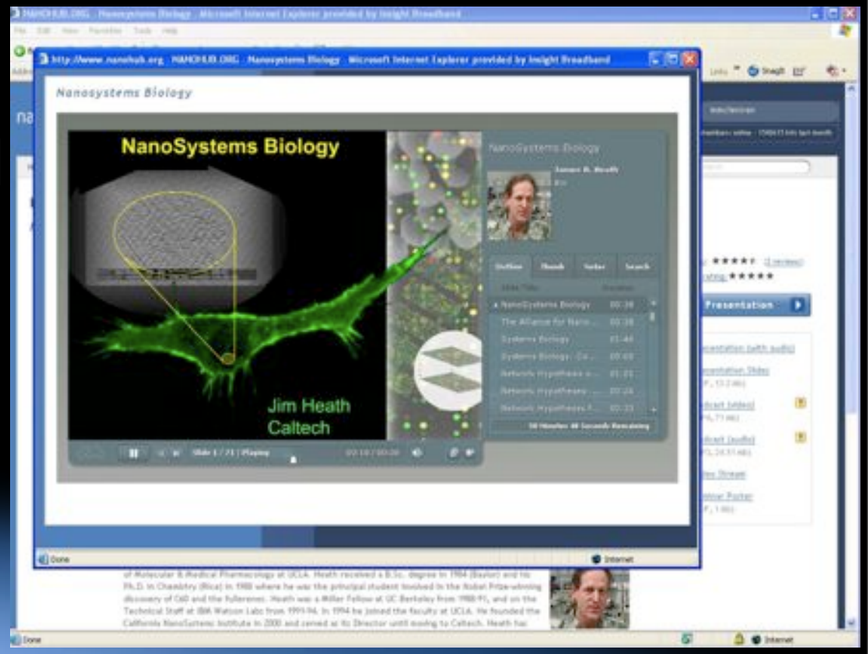
- XD is looking for innovative proposals for a new infrastructure to deliver the next generation of high-end digital services, as national resources
- Goal is to advance science and engineering by providing researchers and educators with the capability to work with extremely large amounts of digitally represented information and making it easy to move between local and national resources
- Anticipate researchers working with much larger range of digital artifacts, including digital text, digitized physical samples, real-time data streams, …

Science Gateways (Virtual Organizations)

- Built to serve communities of practice by bringing together a variety of resources in a customized portal
- Examples include:
  - NanoHub (models, analysis tools, education tools for nano-science and nano-engineering)
  - LEAD (A gateway to support on-demand modeling and analysis of tornados and other strong storms)
  - SCEC Earthworks Project (Earthquake hazard assessment from first principles)
  - CyberTools in Louisiana: (Statewide VO's sharing common CI)
- http://www.teragrid.org/programs/sci_gateways/

Science Gateways (Virtual Organizations)

- Built to serve communities of practice by bringing together a variety of resources in a customized portal
- Examples include:
  - NanoHub (models, analysis tools, education tools for nano-science and nano-engineering)
  - LEAD (A gateway to support on-demand modeling and analysis of tornados and other strong storms)
  - SCEC Earthworks Project (Earthquake hazard assessment from first principles)
  - CyberTools in Louisiana: (Statewide VO's sharing common CI)
- http://www.teragrid.org/programs/sci_gateways/



CI goes far beyond HPC!

NSF Programs in ...

- Software, Applications (need much more in future!)
  - PetaApps, SDCI, CDI
- Networking
  - International Research Network Connections (IRNC)
- Virtual Organizations
  - VOSS
- Data
  - DataNet: bold program for science in data-driven era
- EPSCOR Track I, Track II!

Summary

- Shared CI develops state research & education base
  - Preparing for complex problem solving in 21st century science
  - Perhaps most important aspect is the community building!
- Comprehensive approach is needed
  - Balanced investments in faculty, staff, infrastructure over sustained period of time needed for success
- Build on what is out there, align regional, national approaches
  - Louisiana has layered campus, state, federal investments to build
  - Some project proposals available upon request...
- I want to hear from you! How can we best partner, what can OCI/NSF do in the future?