U.S. Grid Projects and Involvement in EGEE

Ian Foster

Argonne National Laboratory
University of Chicago

http://www.mcs.anl.gov/~foster

EGEE, CERN, February 22, 2003

Overview

- U.S. Grid projects
  - Overview of current state in infrastructure, applications, and middleware
- Next steps
  - NSF "Cyberinfrastructure" report
  - Planned "LHC" ITR proposal & GRIDS-2
  - Building a Grid middleware community
- U.S. involvement in EGEE
  - Integration of infrastructure
  - Collaboration on middleware

U.S. Grid Projects: Current State Infrastructure + Applications

- Infrastructure deployment & operation
  - NSF TeraGrid, iVDGL, DOE Science Grid, NASA IPG, BIRN, various regional Grids
  - Good progress, but still far from critical mass as a national "cyberinfrastructure"
- Applications R&D and deployment
  - (HEP) GriPhyN, iVDGL, PPDG [next slide]
  - (other) Earth System Grid, NEESgrid, NEON, GEON, etc.
  - Substantial engagement of large application communities; much more needed

"Trillium": US Physics Grid Projects

- Particle Physics Data Grid (PPDG)
  - Data Grid tools for HENP experiments
  - DOE funded, $9.5M
- GriPhyN
  - Data Grid research
  - NSF funded, $11.9M
- iVDGL
  - Development of global Grid lab
  - NSF funded, $14.1M

- Data intensive expts.
- Collaborations of physicists & computer scientists
- Infrastructure development & deployment
- Globus + VDT based

=

iVDgL

PPDG

GriPhyN
Data Intensive Science

U.S. Grid Projects: Current State Middleware

- Diverse mix of projects & funding sources
  - No-one wants to fund middleware!
- Unified by two long-lived projects: Globus (since 1995), Condor (since 1990)
  - Persistence, strategic direction, skilled staff
  - Foundation for essentially all Grid projects
- Much recent progress towards creation of professional, distributed support structures
  - With support from NSF Middleware Initiative/GRIDS Center & GriPhyN (VDT)

Impact of NSF NMI/GRIDS Center: Evolution of GT Processes

- Before 2000
  - Email-based problem tracking, aka "req"
- 2000
  - Detailed documentation, release notes (Q1)
  - Legal framework for external contributions (Q1)
- 2001
  - Packaging; module & binary releases (Q4)
  - Substantial regression tests (Q4)
- 2002
  - Bugzilla problem reporting & tracking (Q2)
  - Processes for external contrib (Q2)
  - Distributed testing infrastructure (Q3)
- 2003 (in progress)
  - Distributed support infrastructure: GT "support centers"
  - Standardized Grid testing framework(s)
  - GT "contrib" components
  - Grid Technology Repository

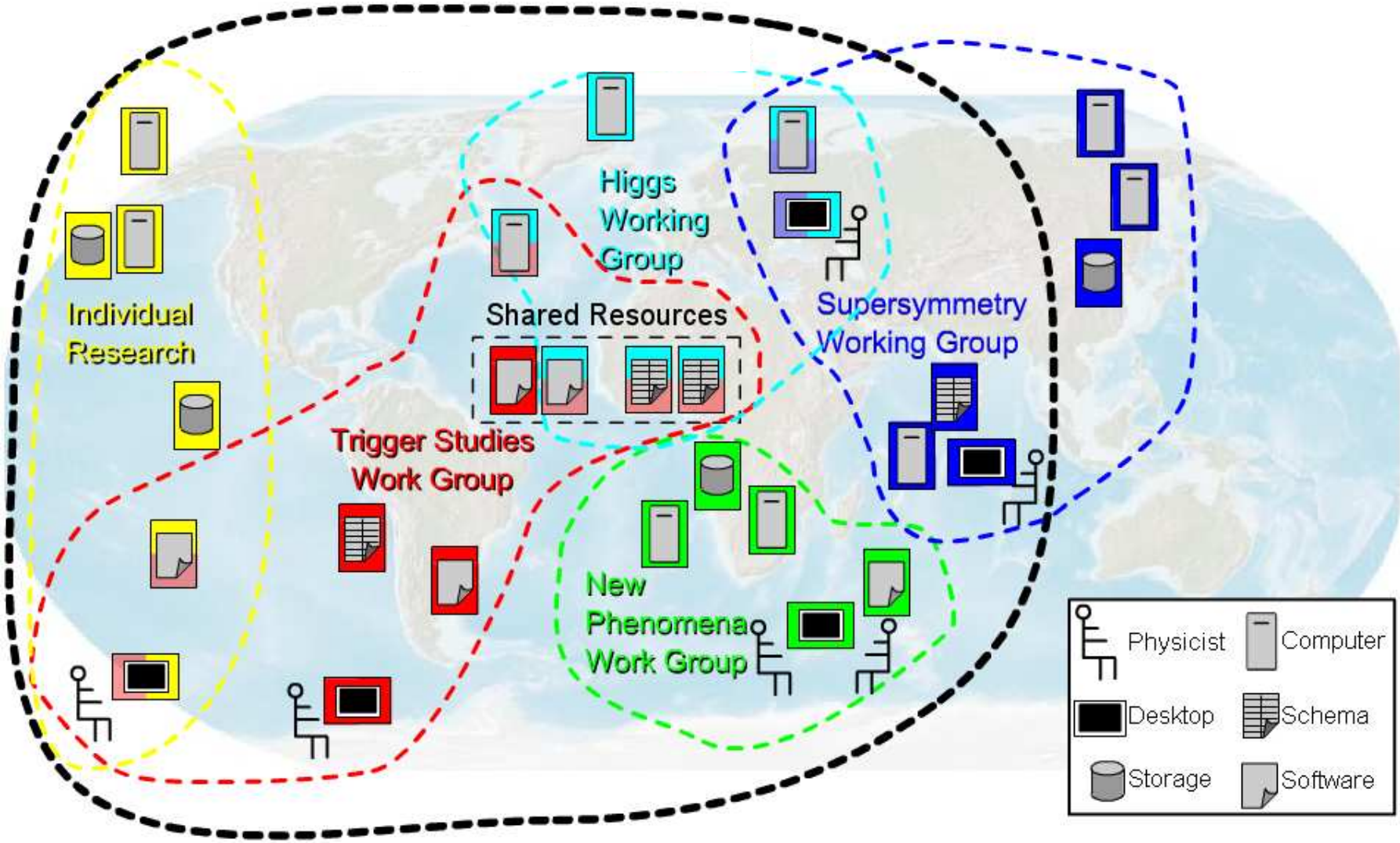
Overview

- U.S. Grid projects
  - Overview of current state in infrastructure, applications, and middleware
- **Next steps**
  - **NSF "Cyberinfrastructure" report**
  - **Planned "LHC" ITR proposal & GRIDS-2**
  - **Building a Grid middleware community**
- U.S. involvement in EGEE
  - Integration of infrastructure
  - Collaboration on middleware

NSF Cyberinfrastructure Report

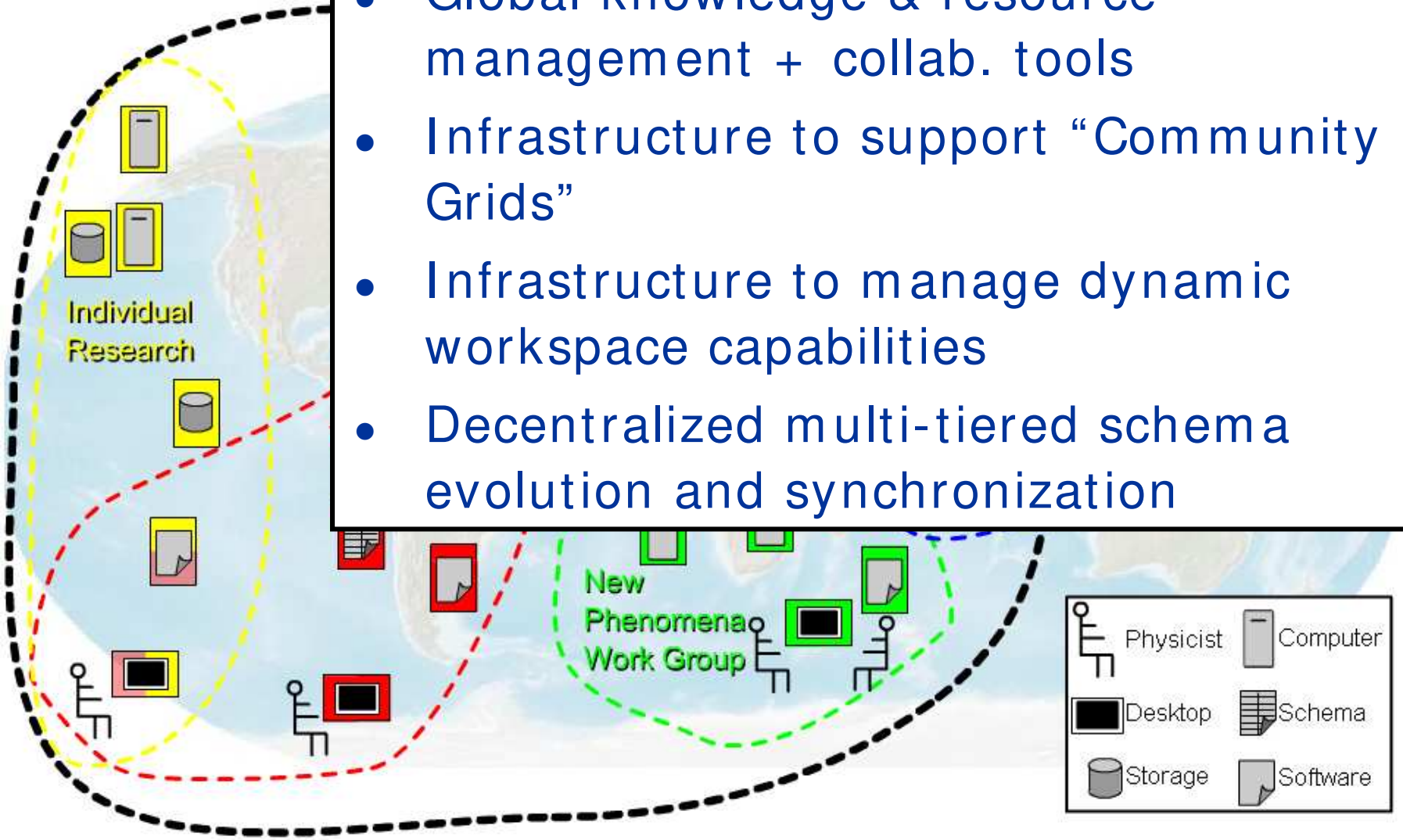
- (summary of report)

NSF ITR: Global Analysis Communities



NSF ITR: Global Analysis Communities

- Global knowledge & resource management + collab. tools
- Infrastructure to support "Community Grids"
- Infrastructure to manage dynamic workspace capabilities
- Decentralized multi-tiered schema evolution and synchronization



GRIDS Center 2 (Proposals to NSF Due March 7th)

- Five principal goals
  - Transition to OGSA standards
  - Expand range of functionality supported
  - Put in place a distributed integration, testing, and support structure
  - Expand the set of communities supported
  - Establish international collaborations

GRIDS Center 2: An Open Grid Technology Community

- Success of Grid concept demands effective community mechanisms for coordinated
  - R&D for core technologies
  - Testing, packaging, documentation
  - Support and training of user communities
- All three must become collaborative activities
  - Based on open standards (GGF)
  - Centered on a common code base
  - Supported by appropriate tools
  - United by appropriate processes & governance

Overview

- U.S. Grid projects
  - Overview of current state in infrastructure, applications, and middleware
- Next steps
  - NSF "Cyberinfrastructure" report
  - Planned "LHC" ITR proposal & GRIDS-2
  - Building a Grid middleware community
- **U.S. involvement in EGEE**
  - **Integration of infrastructure**
  - **Collaboration on middleware**

U.S. Involvement in EGEE (1) Integration of Infrastructure

- EGEE (& U.S. equivalents) to serve science communities with international scope
  - Physics, astronomy, environment, bio, …, …
- We must design in support for international connections from the start
  - Need to learn more about application needs & implications on Grid design and operation
  - But some things clear, e.g., standards
- Strong application community interest in establishing these connections

U.S. Involvement in EGEE (2) Collaboration on Middleware

- With OGSA, EGEE, international science, & evolving distributed support structures, stars are aligned for true international cooperation
  - On software: transform GT into international collaboration (think Linux)
  - On testing & support: integrate EGEE staff and systems into international testing system
- Not obvious, however
  - Many opportunities for not-invented-here, start-from-scratch perspectives!

U.S. Involvement in EGEE (3) Specific Activities

- Explicit collaborative efforts aimed at
  - Integrating U.S. and EGEE resources in support of international science
  - Creation & operation of coordinated international testing and support structure
- Direct engagement of U.S. groups in EGEE work packages
  - Joint development probably appropriate in some cases, e.g., monitoring/operations
  - Establishment and operation of the structures above, e.g., to ensure support