The GRIDS Center: Helping to Build the Cyberinfrastructure

presented by
Tom Garritano
University of Chicago and Argonne National Laboratory

NSF Middleware Initiative (NMI)

- Two Primary Teams
  - GRIDS Center (ISI, NCSA, U. Chicago, UCSD and U. Wisconsin)
  - EDIT (EDUCAUSE, Internet2 and the Southeastern Universities Research Association)
- NMI's main purposes:
  - To design, develop, deploy and support a reusable, expandable set of middleware functions and services that benefit networked applications for research and education in Science and Engineering (S&E)
  - Provide persistent, secure tools for the transparent use and sharing of distributed resources (e.g., computers, data, networks and instruments)
- NMI strives for an architecture and approach to middleware that can be extended to Internet users around the world

GRIDS is Part of NMI

- GRIDS Center
  - GRIDS = Grid Research Integration Development and Support
  - Partnership of leading teams in Grid computing
- Seeks standard components and mechanisms for:
  - Authentication, authorization, policy
  - Resource discovery and directory services
  - Remote access to computers, data, instruments
- Also seeks:
  - Integration with end-user tools (conferencing, data analysis, data sharing, distributed computing, etc.)
  - Integration with campus infrastructures
  - Integration with commercial technologies

GRIDS Center Participants

- Information Sciences Institute (ISI) at USC
  - Carl Kesselman
- The University of Chicago and Argonne Natl Laboratory
  - Ian Foster
- The National Center for Supercomputing Applications (NCSA) at the University of Illinois at Urbana-Champaign
  - Randy Butler
- The San Diego Supercomputer Center (SDSC) at the University of California at San Diego
  - Phil Papadopoulos
- The University of Wisconsin at Madison
  - Miron Livny
- Other Software Contributors (to date: UC Santa Barbara, U. of Michigan)

Goals: Design, Develop, Deploy and Support

- Define an integrated, modular architecture that addresses current & projected middleware requirements for the S&E communities
- Create robust, tested, packaged, documented, and well-supported middleware solutions that are extensible within and beyond S&E
- Work with the middleware research community (e.g., via GGF) to evolve Grid architecture specifications
- Develop procedures for integrating additional software with future NMI releases

GRIDS Outreach to Other Initiatives

- Strong links with other initiatives that are building Grid software, such as:
  - TeraGrid and NEESgrid build on NMI releases
  - GriPhyN/PPDG: Virtual Data Toolkit, including links to EU DataGrid and others
  - UK eScience program: is planning major contributions to OGSA-based Globus Toolkit
  - Companies like IBM, Oracle, Platform Computing: build significant product offerings on components in the GRIDS Center Software Suite
- GRIDS works with NSF and other NMI partners to publicize middleware achievements

The New York Times
nytimes.com

July 15, 2003

Teaching Computers to Work in Unison

By STEVE LOHR

Computers do wondrous things, but computer science itself is largely a discipline of step-by-step progress as a steady stream of innovations in hardware, software and networking pile up. It is an engineering science whose frontiers are pushed ahead by people building new tools rendered in silicon and programming code rather than the breathtaking epiphanies and grand unifying theories of mathematics or physics.

Yet computer science does have its revelatory moments, typically when several advances come together to create a new computing experience. One of those memorable episodes took place in December 1995 at a supercomputing conference in San Diego. For three days, a prototype project, called I-Way, linked more than a dozen big computer centers in the United States to work as if a single machine on computationally daunting simulations, like the collision of neutron stars and the movement of cloud patterns around the globe.

There were glitches and bugs. Only about half of the 60 scientific computer simulations over the I-Way worked. But the participants recall those few days as the first glimpse of what many computer scientists now regard as the next big evolutionary step in the development of the Internet, known as grid computing.

Dr. Ian Foster has helped create the basic software for grid computing. (Photo by Peter Kiar.)

"It was the Woodstock of the grid — everyone not sleeping for three days, running around and engaged in a kind of scientific performance art," said Dr. Larry Smarr, director of the California Institute for Telecommunications and Information Technology, who was the program chairman for the conference.

The idea of lashing computers together to tackle computing chores for users who tap in as needed — almost as if a utility — has been around since the 1960's. But to move the concept of distributed computing utilities, or grids, toward practical reality has taken years of continuous improvement in computer

GRIDS Center Software Suite

- **Globus Toolkit®.** The de facto standard for Grid computing, an open-source "bag of technologies" to simplify collaboration across organizations. Includes tools for authentication, scheduling, file transfer and resource description.
- **Condor-G.** Enhanced version of the core Condor software optimized to work with GT for managing Grid jobs.
- **Network Weather Service (NWS).** Periodically monitors and dynamically forecasts performance of network and computational resources.
- **Grid Packaging Tools (GPT).** XML-based packaging data format defines complex dependencies between components.

GRIDS Center Software Suite (cont.)

- **GSI-OpenSSH.** Modified version adds support for Grid Security Infrastructure (GSI) authentication and single sign-on capability.
- **MyProxy.** Repository lets users retrieve a proxy credential on demand, without managing private key and certificate files across sites and applications.
- **MPICH-G2.** Grid-enabled implementation of the Message Passing Index (MPI) standard, based on the popular MPICH library.
- **GridConfig.** Manages the configuration of GRIDS components, letting users regenerate configuration files in native formats and ensure consistency.
- **KX.509 and KCA.** A tool from EDIT that bridges Kerberos and PKI infrastructure.

GRIDS platforms

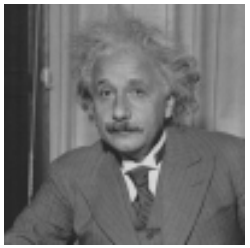
- NMI-R3.1 maintenance release supports:
  - Red Hat Linux 7.2, 7.3, 8.0 and 9.0 on IA32
  - Red Hat Linux 7.2 on IA64
  - SuSE Linux Enterprise Server 8 on IA64
  - Solaris 8.0 on SPARC
- Distributed as binaries
  - Source distribution is available, but is not officially supported

Future GRIDS Components: Criteria for Inclusion

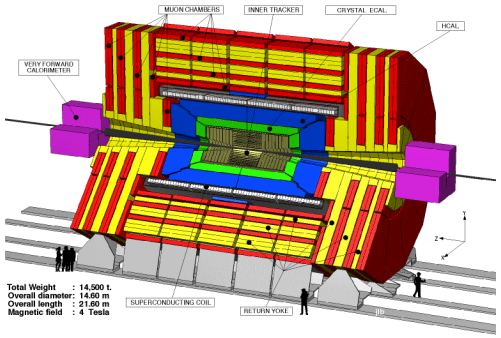
- Production-quality software
- Documentation
- Testing
- Packaging
- Technical Support
- Ability to work with the NMI release process
- Fit with other components
- GRIDS Center Architecture group – PI's

GRIDS Context

- Pre-Internet Collaboration
  - Theorize and/or experiment, usually alone or in small teams; publish paper
  - Travel to instrumentation sites



- Post-Internet Collaboration
  - Build and access large databases based on observation or simulation results
  - Remotely access specialized instruments
  - Exchange information within broadly distributed multidisciplinary teams, preferably in real time



An Emerging "Cyberinfrastructure"

- NSF is viewed favorably for its role in the Internet's and Web's growth
- Information Technology Research initiative (ITR) is in its fourth year
  - Funds research into basic IT and its applications for S&E
  - Fosters teams of computer scientists and disciplinary scientists
- Important report issued in January 2003
  - NSF-commissioned "blue ribbon" panel on Cyberinfrastructure
  - Chaired by Dan Atkins, University of Michigan
- The report is a roadmap for NSF to reinvent itself
  - Not only in terms of IT, although IT is key
  - All NSF S&E directorates will participate and benefit
  - Current Grid-based projects are among the important models
  - Push scientific collaboration and resource sharing onto desktops

Revolutionizing Science and Engineering Through Cyberinfrastructure:

Report of the National Science Foundation
Blue-Ribbon Advisory Panel on
Cyberinfrastructure

January 2003

http://www.cise.nsf.gov/evnt/reports/toc.htm

Daniel E. Atkins, Chair
University of Michigan

Kelvin K. Droegemeier
University of Oklahoma

Stuart I. Feldman
IBM

Hector Garcia-Molina
Stanford University

Michael L. Klein
University of Pennsylvania

David G. Messerschmitt
University of California at Berkeley

Paul Messina
California Institute of Technology

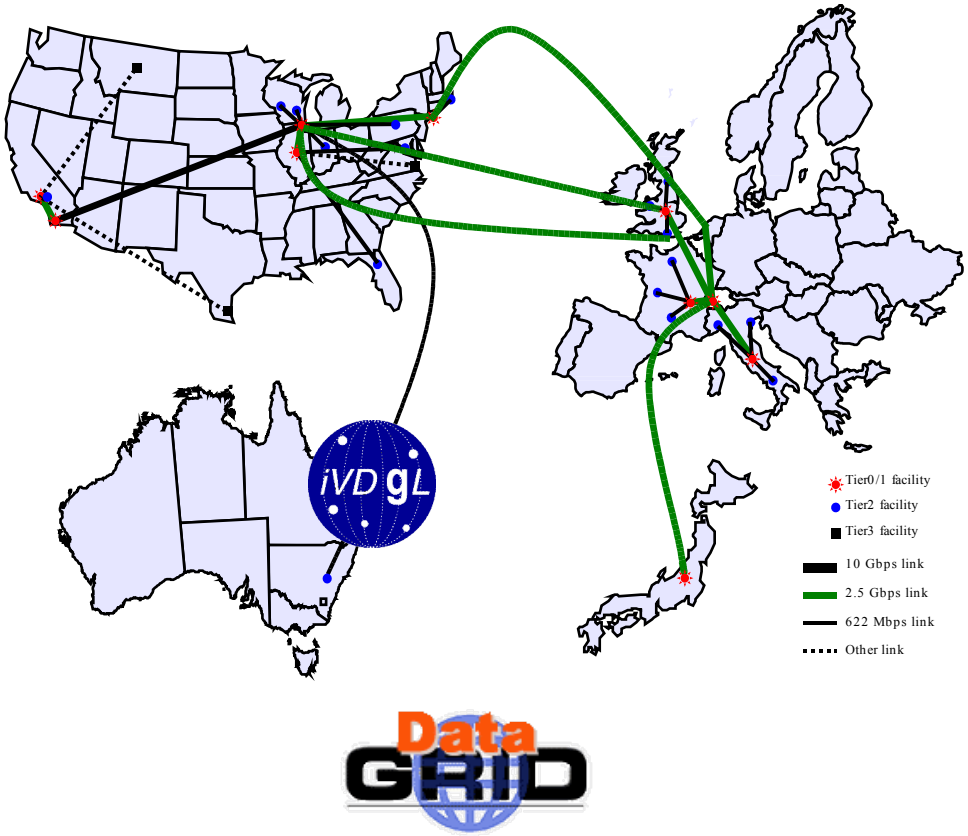
Jeremiah P. Ostriker
Princeton University

Margaret H. Wright
New York University

Model Grid Deployments

Building the National Virtual Collaboratory for Earthquake Engineering Research

NEESgrid



GriPhyN
Data Intensive Science

NVO

LIGO

SDSS

CMS
Compact Muon Solenoid

PPDG

DataGRID

Future GRIDS Plans

- GRIDS will complete its second year in October
  - Original three-year award, through Fall 2004
  - Very successful in terms of establishing processes and meeting twice/year release schedule
  - Strong indications of GRIDS being renewed by NSF, with announcement anticipated in August
- GRIDS Center 2 plans
  - Further develop and refine core NMI releases and processes
  - Move to Open Grid Services Architecture
  - Expand testing capability
  - Increase outreach (e.g., GlobusWorld 2004 in San Francisco, January 20-23 -- see http://www.globusworld.org)
  - Enhance integration with projects like TeraGrid, NEESgrid
  - Help establish a Cyberinfrastructure with tested, hardened Grid software

For more information

The GRIDS Center
http://www.grids-center.org/

NSF Middleware Initiative
http://www.nsf-middleware.org/

NSF Directorate for Computer and Information Science and Engineering (CISE)
http://www.nsf.gov/cise/