The Emerging Grid Experience Base

Dr. Francine Berman

Director, SDSC and NPACI

Professor and High Performance Computing Endowed Chair, Department of Computer Science and Engineering, UC San Diego
Grid Computing

- Google shows 434 refs for “Grid Computing”
- Amazon.com has 69 books on “Grid Computing”
- IBM, SUN, HP, United Devices, Cray, Intel, etc. have efforts in Grid computing
- The Global Grid Forum participants come from over 400 organizations in over 50 countries
- GridToday reaches over 37,000 readers

How come there aren’t more users?
The many faces of Grid Computing

The Grid as a balancing act

The Grid as core infrastructure

The Grid as over-hyped technology

BUILDING GRIDS: HYPE MEETS REALITY

The senior executives from Compaq are part of the new HP3, Platform Computing and renowned analysts from IDC invite you to attend an Exclusive Executive event on Grid Computing.

Insanely Overhyped Technology of the Month - Grid Computing

Distributing executable processing capacity across thousands or millions of Grids isn’t the right approach for most of the problems that computing...
Do Plants Practice Grid Computing?

According to Nature, plants seem to optimize their 'breathing' by conducting simple calculations through a distributed computing scheme. "Plants appear to 'think', according to US researchers, who say that green plants engage in a form of problem-solving computation."  

January 2004
Today’s Presentation

• What is Grid Computing?
• How did we get here?
• Case studies: TeraGrid and NPACI Grid
• Challenges in realizing the Grid vision
• How we know when we’ll get there
Grid Computing

- Grid computing enables a set of resources (compute, data, network, etc.) to be used by an application as an ensemble
The Vision: Grids should be

- **Useful**, able to support/promote *new science*
- **Usable** (accessible, robust, easy-to-use)
- **High-capacity** (rich in resources)
- **High capability** (rich in options)
- **Evolutionary** (able to adapt to new technologies and uses)
- **Persistent** (usable by community in the long term)
- **Stable** (usable by community in the short term)
- **Scalable** (growth must be a part of the design)
- **Integrative** (promoting end-to-end solutions)
The problem is not the vision ... it’s the underestimation of what it will take to get there

• Just building Grid environments makes research questions out of previously solved problems
  • Installation
  • Configuration
  • Accounting
• And brings added complexity to existing problems
  • Performance analysis
  • Debugging
  • Scheduling
  • Security
  • Fault tolerance, etc.
Grid Computing: In the beginning …

- “Science as a team sport”
  - Grand Challenge Problems of the 80’s
- Parallel computation
  - First serious study of program coordination
- Gigabit Testbed program
  - Focus on applications for the local to wide area
- First “modern” Grid: I-Way at SC ‘95
  - First large-scale grid experiment
  - Provided the basis for major grid infrastructure efforts
1995-present: Maturing research and development

- **Important infrastructure and middleware efforts developed**
  - Globus, Legion, Unicore, Condor, NWS, SRB, NetSolve, AppLes, GrADS/VGrADS, etc.

- **Peering provides experience with application scalability in the large**
  - Entropia, Seti@home, United Devices

- **Internationalism**
  - Evolution of the Global Grid Forum
  - International linkage within and between Grid projects (PRAGMA, E-Science, TeraGrid, etc.)

- **Integration of Grid Computing and Web Services**
  - Emerging focus on OGSA, WSRF, etc.

"Grid books" give a comprehensive view of the state of the art
Grid Computing Today – A Work in Progress

NSF Cyberinfrastructure

EUROGRID

NPACI

GRID

NEESgrid

Data GRID

TERAGRID

BIRN

European GRID Forum

e-Science dti

GEON

GLOBAL GF

PRAGMA

SAN DIEGO SUPERCOMPUTER CENTER, UCSD
Many communities focusing on some form of distribution and coordination

- Digital Libraries
- Internet applications
- Network applications
- Enterprise computing
- Gaming
- Collaboratories
- Grid Computing

Seti@home sustains 35 TF on 2M+ laptops

1.7 ZETAflop over last 3 years (10^21, beyond peta and exa …)
Distributed computing is alive and well in the commercial world

- **Walmart Inventory Control**
  - Satellite technology used to track every item
    - Bar code information sent to remote data centers to update inventory database and cash flow estimates
    - Satellite networking used to coordinate vast operations
  - Inventory adjusted in real time to avoid shortages and predict demand
    - *Data management, prediction, real-time, wide-area synchronization*
**Distributed RPGs**

- **Everquest**
  - 45 communal “world servers” (26 high-end PCs per server) supporting 430,000 players
  - Real-time interaction, individualized database management, back channel communication between players
  - Data management adapted to span both client PC and server to mitigate communication delays
  - Game masters interact with players for real-time game management
An Emerging Grid Experience Base

Case Studies:

TERAGRID and NPACI GRID
• > 20 TF distributed at 9 sites (SDSC, NCSA, ANL, PSC, Caltech, Indiana U., Purdue U., Oak Ridge National Lab/U of Tennessee, UT Austin)

• > 1 PB storage distributed at 5 sites

• Fast national network with 40 Gb/sec between hubs

• Linux-based SW environment, uniform administration
Initial TeraGrid Apps providing software shake-out

- ENZO (Astrophysics)
- PPM (Astrophysics)
- GAFEM (Ground-water modeling)
- GridSAT (Computer Science)
- AtlasMaker (Astronomy)
- MEAD (Atmospheric Sciences)
- CMS/GriPhyN (Physics)
- TeraGyroid (Condensed Matter Physics)
- Encyclopedia of Life (Biosciences)
- VTF (Shock Physics)
- BioCoRE (Biomedicine)

NATIONAL PARTNERSHIP FOR ADVANCED COMPUTATIONAL INFRASTRUCTURE
SAN DIEGO SUPERCOMPUTER CENTER, UCSD
**Social Engineering is Critical**

**TeraGrid Management Structure**

- **NSF Review Panels**
  - Currently being formed

- **Institutional Oversight Committee**
  - Frieder Seible, UCSD
  - Richard Herman, UIUC
  - Dan Meiron, CIT (Chair)
  - Robert Zimmer, UC/ANL

- **External Advisory Committee**
  - Are we creating an extensible cyberinfrastructure?

- **Technical Working Group**
  - Are we creating an extensible cyberinfrastructure?

- **Executive Director / Project Manager**
  - Charlie Catlett (UC/ANL)

- **Project Engineer**
  - Pete Beckman (UC/ANL)

- **Chief Architect**
  - Dan Reed (NCSA)

- **Executive Committee**
  - Fran Berman, SDSC (Chair)
  - Mike Levine, Ralph Roskies, PSC
  - Jim Poole, CIT
  - Rick Stevens, UC/ANL
  - Charlie Catlett, ANL, +TEPs

- **User Advisory Committee**
  - Are we effectively supporting good science?

- **Technical Coordination Committee**
  - Project-wide Technical Area Leads
    - **Performance Eval**
      - Brunett (Caltech)
    - **Applications**
      - Williams (Caltech)
    - **Visualization**
      - Papka (ANL)
    - **Networking**
      - Winkler (ANL)
    - **Grid Software**
      - Kesselman (ISI)
      - Butler (NCSA)
    - **Clusters**
      - Pennington (NCSA)
    - **Operations**
      - Sherwin (SDSC)
    - **Data**
      - Baru (SDSC)

- **Site Coordination Committee**
  - Site Leads
    - **ANL**
    - **CIT**
    - **NCSA**
    - **SDSC**
    - **PSC**
    - **Purdue**
    - **IU**
    - **TACC**
    - **ORNL**

- **Policy Oversight**
  - Objectives
  - Architecture
  - Implementation

**SAN DIEGO SUPERCOMPUTER CENTER, UCSD**
Important Ideas

• **TeraGrid “Roaming”**
  - Application can roam TG to assemble resources and execute with a single user interface

• **INCA Test Harness**
  - Use ‘unit test’, ‘version reporter’, and ‘integration tests’ to assure each the quality of each component in the system

• **SAN/WAN innovations**
  - SDSC file system accessible across TeraGrid
  - Remote transfer rates similar to local allowing easy TG-wide data transfer (900 MB/s)

---

**SAN DIEGO SUPERCOMPUTER CENTER, UCSD**

**NATIONAL PARTNERSHIP FOR ADVANCED COMPUTATIONAL INFRASTRUCTURE**
The NPACI Grid: Born Heterogeneous

NPACI alphas and apps incorporate NPACKage and common infra to run on NPACI resources

NPACKage focuses on all-to-all interoperability, packaging, documentation for mature NPACI SW

Community infrastructure deployed on NPACI resources

NPACI resource sites support deployment of NPACKage, NMI, etc.
NPACKage: Focus on interoperability and usability

- NPACKage is an interoperable collection of NPACI SW, providing user-focused middleware for the NPACI Grid
  - Builds on top of NMI

- NPACKage Components
  - The Globus Toolkit™
  - GSI-OpenSSH
  - Network Weather Service
  - DataCutte
  - Ganglia
  - LAPACK for Clusters (LFC)
  - MyProxy
  - GridConfig
  - Condor-G
  - Storage Resource Broker (SRB)
  - Grid Portal Toolkit (GridPort)
  - MPICH-G2
  - APST (AppLeS Parameter Sweep Template)
  - Kx509

- Technology integration
  - All-to-all interoperability

- Packaging and deployment

- Maintenance

- User support
  - Documentation
  - Consulting
  - Help-desk
# NPACI Grid and TeraGrid/ETF: Building an experience base with Grid technologies

<table>
<thead>
<tr>
<th></th>
<th>TeraGrid/ETF</th>
<th>NPACI Grid</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Component hardware</strong></td>
<td>Homogeneous with DTF, heterogeneous with ETF1, ETF2</td>
<td>Heterogeneous from the beginning</td>
</tr>
<tr>
<td><strong>Networks</strong></td>
<td>Custom backplane</td>
<td>Commodity, Internet 2</td>
</tr>
<tr>
<td><strong>Software</strong></td>
<td>Software stack developed to leverage uniform, high capacity resources; now evolving to target heterogeneous resources</td>
<td>Developed from existing mature SW efforts, focus on interoperability, packaging, and usability</td>
</tr>
<tr>
<td><strong>Initial users</strong></td>
<td>TeraGrid Flagship apps</td>
<td>NPACI Alphas, other apps</td>
</tr>
</tbody>
</table>
Challenges in Realizing the Grid Vision
Challenges in Building Grids

- Technical Challenges
  - Scalability of grid infrastructure
  - Fault tolerance and adaptivity – adapting to resource failure and performance variation
  - Next generation programming models (real-time, on-demand, adaptive, etc.)
  - Grid programming environments: libraries, compilers, schedulers, performance tools, etc.

- How should we schedule jobs across different resources?
- How should the system adapt when one of the resources is unacceptably slow or unavailable?
- What went wrong? How can I debug my Grid program?
Challenges in Building Grids

- **Logistical Challenges**
  - Who maintains the SW?
  - Who fixes the bugs?
  - Who documents the code?
  - Who writes the manual?
  - Who answers the phone?

- **Legal Challenges**
  - How do we deal with varying institutional IP policies, open source policies, licensing policies, etc.

- **What do I need to do to “join the Grid”?**
- **Who do I call when I’m having trouble?**
- **Will my institution/company allow me to contribute to open source software?**
Challenges in Building Grids

• Ideological Challenges
  • Commercial vs. academic community – should someone make a profit?
  • Policy: What is shared/private, free/charged for, centralized/distributed, etc.
  • Who makes decisions? Who enforces decisions? Do we need a “Grid U.N.”?

• Political challenges
  • How to share resources across national boundaries?
  • How to ensure stable, persistent, long-term, adequate funding
  • How to achieve scientific goals in the context of political constraints

• Can I get an account in Egypt?

• What if I don’t want to run your application on my resources?

• What am I willing to pay for?
Next Generation Challenges
Policy – The Grid as a social system

• The Grid as a social system: Aggregate user behavior must not destabilize the system
  • Grid users seek individual application performance – what “rules of the road” do we need to put into place to ensure stability?
Policy – The Grid as an economic system

• The Grid as an economic system:
  Users want “end-to-end” performance – computer cycles, network bandwidth, data storage must be balanced
  • “Multiple currencies, multiple users” problem

Social, economic, political policies for the Grid are uncharted territory
Advanced User Environments

- Every 14 year old gamer regularly uses more sophisticated user environments than most scientists

- Today’s RPG games provide robust, interactive, dynamic, distributed environments with 10’s of thousands of users

- Gaming technologies demonstrate sophisticated adaptation to available resources (home PCs, cable modems)
The problem is not the vision, it’s the underestimation of what it takes to get there

- What should our aggregate Grid/Cyberinfrastructure/integrated technology platforms look like in 10 years?
  - What kind of research and development should we be encouraging now?
    - Research must continually inform the discussion
    - Platforms for testing and development are needed to provide experience with mature and tested prototypes

- A long-term strategy is critical.
How we’ll know when we get there …

• Our Cyberinfrastructure should be as easy to use as our cars
  • All components should work together to provide end-to-end performance
  • Things should be basically where you expect them to be – you can get in almost any car and drive it with a minimum of effort
  • There should be an instruction manual that’s user friendly
  • It should be relatively easy to find someone who can fix it when it breaks
  • It’s more about where you’re going than whether your car works
Thank You