The Neutron Science TeraGrid Gateway: TeraGrid Cyberinfrastructure at ORNL

Presented by

John W. Cobb, Ph.D.
Computer Science and Mathematics
Office of Cyberinfrastructure
National Science Foundation
Outline

• Cyberinfrastructure (CI)
• TeraGrid
• Spallation Neutron Source
• Neutron Science TeraGrid Gateway
• Other collaborations
TeraGrid Partners Network

Also see TeraGrid SC10 exhibit (1613) and other TeraGrid Resource Provider exhibits. See TeraGrid pylon in this booth for map and directions.

- Dedicated links
- Resource to resource
- Bandwidth unit: 10 Gbps
- Substrate for high-performance, high-quality data movement and communications

Resource Provider (RP)
Software Integration Partner
TeraGrid resources

- 11 distributed centers
- >2 petaflops compute
- Multi-petabyte disk
- Multi-petabyte archive
- Different resource types
  - Capability
  - Capacity
  - High throughput
  - Visualization
  - Specialized use
- Common user environments
- Tuned data transfer over dedicated 10+ gbs network
Science gateways

• TeraGrid size: How can we manage it?
  – Today: >4000 user accounts across 20 resources
  – Target: 10–100× increase in user base; increase in number of resources

• TeraGrid science gateways
  1. Web portal with users in front and TeraGrid services in back
  2. Grid-bridging gateways: Extending the reach of a community grid (devoted to a single area of science) so it can use TeraGrid resources
  3. Application programs running on users’ machines (e.g., workstations and desktops) that access TeraGrid (and other) services

• The Neutron Science TeraGrid Gateway NSTG is one of 33 existing gateways

Charting new territory for wide cyberinfrastructure services

• Federated user management
• Callback end-user identification with automated service auditing across federated identity space “gateway community accounts”
• Grid interoperability
• Participation in attribute-based authorization
### Spallation Neutron Source

“The next generation of materials research”

<table>
<thead>
<tr>
<th>Large world-class user facility</th>
<th>Neutrons as a probe of matter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction complete</td>
<td>1998–2006, TPC $1.4 billion</td>
</tr>
<tr>
<td>Accelerator power-up</td>
<td>1 MW beam power record set in September 2009</td>
</tr>
</tbody>
</table>

#### Instruments
- User Program
  - SNS: 12
  - HFIR: 9
- Additional under construction and development

#### User program
- 983 beam-time proposals submitted for 5,306 user days in 2010 (2+ times oversubscription)
- Next proposal deadline: March 2, 2011
HFIR instrument suite

Fixed-incident-energy triple-axis spectrometer - HB-1A
Low-energy excitations, magnetism, structural transitions
Jared Zarecki - 865.574.6951
zareckij@ornl.gov

Polarized triple-axis spectrometer - HB-1
Polarized neutron studies of magnetic materials, low-energy excitations, structural transitions
Mervin Webber - 865.574.6006
mwebber@ornl.gov

Neutron powder diffractometer - HB-2A
Structural studies, magnetic structures, texture and phase analysis
Ovidiu Curtei - 865.574.5941
curteoi@ornl.gov

US/Japan WAND - HB-2C
Diffuse-scattering studies of single crystals and time-resolved phase transitions
Jon Fernández-Ruiz - 865.970.8999
fernandezdr@ornl.gov

Future development - HB-2D
Neutron residual stress mapping facility - HB-2D
Strain and phase mapping in engineering materials
Candis Hubbard - 865.574.4372
hubbardc@ornl.gov

THIRD-AXIS SPECTROMETER - HB-3
Medium- and high-resolution inelastic scattering at thermal energies
Mark Lumley - 865.574.8000
lumleym@ornl.gov

Four-circle diffractometer - HB-3A
Small unit-cell crystal structural studies, particularly H-bonding
Byron Chakravartika 865.574.5205
chakravartika@ornl.gov

Development beamline - CG-1
Imaging, optics, SERGIS, sample alignment
Lee proporcion - 865.574.5204
propionel@ornl.gov

General-purpose SANS - CG-2
Polymer blends, flux lattices in high-Tc materials, soft materials processing and structure
Ken Kittrell - 865.574.4526
kittrellk@ornl.gov

Bio-SANS - CG-3
Proteins and complexes, pharmaceuticals, biomaterials
Walker Urban - 865.574.5275
urbanw@ornl.gov

Future development - CG-4A
High-resolution inelastic scattering at cold neutron energies
Harry Whin - 865.574.5202
whin@bnl.gov

Image-plate single-crystal diffractometer (IMAGINE) - CG-4D (2010+)
Chemical, organic, metal-organic, protein single crystals
Flora Mieleur - 865.574.3807
mieleurf@ornl.gov

LEGEND
Installed, commissioning, or operating
In design or construction
Under consideration

* Scheduled commissioning date.

Managed by UT-Battelle
for the U.S. Department of Energy
### SNS data access via portal (Reduction, analysis, and simulation as well)

Turnkey solutions are being requested by users and user groups

<table>
<thead>
<tr>
<th>Problem</th>
<th>SNS approach</th>
</tr>
</thead>
</table>
| - Providing advanced tools in an easy fashion while retaining customizability | - Neutron Science User Portal  
  - Totally in Java |

Function deployment order
- First: Data acquisition and storage/stewardship
- Then: Reduction, analysis, simulation, proposals for one-stop access

Acknowledgment: S. Miller and SNS Advisory SW group; J. Kohl, S. Vazhkudai, and ORNL/CSM division
Neutron science portal view

Currently hosts data from five facilities:

SNS
HFIR
IPNS
LANSCE
LENS

Acknowledgement: From SNS ASG
Simulation of inelastic scattering off a $^4$He sample in backscattering spectrometer (BASIS)

- Using McStas
- 1010 simulation neutrons
- Performed on NSTG cluster

Acknowledgment: V. E. Lynch, M. Hagen
NeXus
A community-driven standards process

- Data format for neutron, X-ray, and muon science
- Based on HDF5 (http://www.hdfgroup.org)
- Self-describing
- Compression
  - Data: 95% compression
  - Metadata remain uncompressed (more easily searchable)
- Use increasing as new facilities come on line
  - SNS
  - OPAL at ANSTO
  - ISIS second target station
  - J-SNS
- McStas (neutron Monte Carlo ray-trace simulation tool) can read and write NeXus format

http://www.nexusformat.org/
# Neutron Science TeraGrid Gateway

One of 11 TeraGrid partner resource providers

<table>
<thead>
<tr>
<th>Focus areas</th>
<th>Resources provided</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutron science</td>
<td>Outreach to a specific science community (neutron science)</td>
</tr>
<tr>
<td>Connecting facilities with cyberinfrastructure</td>
<td>- Expertise</td>
</tr>
<tr>
<td>Bridging cyberinfrastructures</td>
<td>- Technology transfer</td>
</tr>
<tr>
<td>Data movement within and across TeraGrid</td>
<td>- Education, in a broad sense</td>
</tr>
<tr>
<td></td>
<td>- User outreach</td>
</tr>
<tr>
<td></td>
<td>A science gateway interface for neutron science</td>
</tr>
<tr>
<td></td>
<td>Exploration of large science facility integration with national-scale cyberinfrastructure</td>
</tr>
<tr>
<td></td>
<td>Combine TeraGrid computational resources with available neutron scattering datasets for analysis and simulation comparison</td>
</tr>
</tbody>
</table>
NSTG operational components

- Wide-area network connections to TeraGrid
- Local connections, at maximum bandwidth, to local facilities (SNS and HFIR)
- Modest, local cluster with complete TeraGrid environment for small, local TeraGrid jobs
- Long-term, high-performance storage system archival storage via ORNL’s LCF/NCCS
- System operations
- Application consulting
Neutron science simulations

Next step: Provide analysis and relevant simulations in support of neutron science

Instrument design optimization
Moving simulations to TeraGrid allows larger runs and faster turnaround. Assisted in effort to redesign instrument to lower cost. Simulations provided confidence that resolution would not be sacrificed.

Porting and deployment of simulation tools to TeraGrid cyber-infrastructure: McStas
Many codes are currently I/O limited. Next step is to modify I/O strategy in codes to take advantage of I/O architectures.

Improved and faster data analysis: Reflectometry
Faster and better $\chi^2$ minimization
Portal-initiated simulations ... under the covers!

Interactive user inputs

Portal applet (simulation GUI)

Job information service

Globus front-end

Globus back-end

Back-end web server

Simulation servlet

“runas”, params

Authorization as user

SNS-user

ssh

Results

Private key

TeraGrid computational resources as “Jimmy Neutron” community account

GridFTP

Community certificate

User “Jimmy Neutron”

grid-proxy-init

globus-job-submit

globus-job-status

globus-url-copy

Portal applet (simulation GUI)

Job information service

Globus front-end

Globus back-end

Back-end web server

Simulation servlet

“runas”, params

Authorization as user

SNS-user

ssh

Results

Private key

TeraGrid computational resources as “Jimmy Neutron” community account

GridFTP

Community certificate

User “Jimmy Neutron”

grid-proxy-init

globus-job-submit

globus-job-status

globus-url-copy

Portal applet (simulation GUI)

Job information service

Globus front-end

Globus back-end

Back-end web server

Simulation servlet

“runas”, params

Authorization as user

SNS-user

ssh

Results

Private key

TeraGrid computational resources as “Jimmy Neutron” community account

GridFTP

Community certificate

User “Jimmy Neutron”

grid-proxy-init

globus-job-submit

globus-job-status

globus-url-copy
NSTG was and is “postured as a pathfinder”

- Proposed a Science Gateway as part of ORNL RP in 2003, before creation of TeraGrid Science Gateways Program
- Included a focus on cyber-enabled science from the outset: enabling neutron science
- One of the first intergrid production operations
- Early (but not first) concentration on intra-TeraGrid high-performance wide-area data transfer
- Concentration on high-performance transfer between TeraGrid and external endpoints
- Early adopter and pathfinder postures fit with NSTG RP role
McStas

- Widely used Monte Carlo neutron ray tracing neutron instrument simulation package for instrument design and experiment planning
- Developed at Risø National Laboratory (Denmark) and Institut Laue-Langevin (France)
- Simulations from moderator through sample to detectors agree with experimental results
- Use cases
  - Instrument design and construction
  - Experiment interpretation
  - Experiment planning
  - Experiment proposal evaluation
- E. Farhi has completed a preliminary study to computationally compute scattering kernels using VASPE
- McStas scalable: 1024+ cores
REDDnet
Research and Education Data Depot Network

- National-scale logistical storage network
- Distributed (but acts local!)
- Funding from NSF and Library of Congress
- >700 TB (500 disk, 200 tape)
- Multiple application domains
  - Satellite imagery (AmericaView)
  - High energy physics: LHC
  - Terascale Supernova Initiative
  - Structural biology
  - Vanderbilt TV news archive
  - National Geospatial Data Archive
- Data everywhere
  (under the cover tools to manage movement, replication, ...)

See Vanderbilt research booth for SC'09 demo, bandwidth challenge and more details
http://www.reddnet.org/

ORNL REDDnet nodes in NSTG and LCF
Orbiter

- Thick client portal implementation
- Multitiered portal architecture
- DOE (BES) SBIR funded (phase I and phase II) to Mark Green, TechX Corp.
- Includes SNS and HFIR data catalog
- Also, data prefix caching and search ("metadata rich" for NeXus files)
- Includes grid execution enabling templates
- Together with REDDnet can form a nomad, or disconnected use model, portal + data store for neutron data—a "gateway on the go"
## Acknowledgments and thanks

### Staff and Effort

- **NSTG**: M. Chen¹, J. Cobb¹, D. Giles¹, V. Hazlewood⁷, S. Hicks¹, G. Hinkel¹, D. Hudson¹, V. Lynch¹, P. Newman¹, J. Nichols¹, D. Pack¹, G. Pike¹, J. Rome¹, J. Travieso⁹, W. Wing¹

- **SNS**: J. Bilheux¹, M. Green⁶, G. Granroth¹, M. Hagen¹, J. Kohl¹, D. Mikkelson⁸, R. Mikkelson⁸, S. Miller¹, P. Peterson¹, S. Ren¹, M. Reuter¹, J. Schwidder¹, B. Smith⁷, T. Swain⁷, J. Trater¹, S. Vazhkudai¹

- **McStas group**: E. Farhi³, K. Lefmann⁵, P. Willendrup⁵

- **OSG**: R. Pordes², M. Livny²

- **ESG**: D. Bernholdt¹, D. Middleton⁴

- **TeraGrid partners and others**

### Support

- National Science Foundation
- U.S. Department of Energy
- ORNL Laboratory Director’s research and development fund

### Organization

¹ORNL, ²Fermilab, ³Institut Laue-Langevin, ⁴NCAR, ⁵RISO, ⁶TechX Corp., ⁷University of Tennessee, ⁸U Wisconsin–Stout, ⁹Austin Peay State University, ¹⁰U Wisconsin–Madison
Contact

John W. Cobb, Ph.D.

Computer Science and Mathematics
(865) 576-5439
cobbjw@ornl.gov