Science on the TeraGrid

Daniel S. Katz[1,2,3], Scott Callaghan[4], Robert Harkness[5], Shantenu Jha[2,6,7],
Krzysztof Kurowski[8], Steven Manos[9]*, Sudhakar Pamidighantam[10],
Marlon Pierce[11], Beth Plale[11,12], Carol Song[13], John Towns[10]

[1]*Computation Institute, University of Chicago and Argonne National Laboratory, USA*
[2]*Center for Computation & Technology, Louisiana State University, USA*
[3]*Department of Electrical and Computer Engineering, Louisiana State University, USA*
[4]*University of Southern California, USA*
[5]*San Diego Supercomputer Center, University of California San Diego, USA*
[6]*Department of Computer Science, Louisiana State University, USA*
[7]*e-Science Institute, University of Edinburgh, UK*
[8]*Poznan Supercomputing and Networking Center, Poland*
[9]*Information Technology Services, University of Melbourne, Australia*
[10]*National Center for Supercomputer Applications, University of Illinois, USA,*
[11]*Pervasive Technology Institute, Indiana University Bloomington, USA*
[12]*School of Informatics and Computing, Indiana University Bloomington, USA*
[13]*Rosen Center for Advanced Computing, Purdue University, USA*

Abstract: The TeraGrid is an advanced, integrated, nationally-distributed, open, user-driven, US cyberinfrastructure that enables and supports leading edge scientific discovery and promotes science and technology education. It comprises supercomputing resources, storage systems, visualization resources, data collections, software, and science gateways, integrated by software systems and high bandwidth networks, coordinated through common policies and operations, and supported by technology experts. This paper discusses the TeraGrid itself, examples of the science that is occurring on the TeraGrid today, and applications that are being developed to perform science in the future.

Key words: computational science applications, high performance computing, grid computing, production grid infrastructure

I. INTRODUCTION

TeraGrid is an an advanced, integrated, nationally-distributed, open, userdriven, US computational science cyber-infrastructure that enables and supports leading edge scientific discovery and promotes science and technology education, operated in a partnership comprising the Grid Infrastructure Group (GIG) and eleven Resource Provider (RP) institutions. The TeraGrid includes supercomputing resources, storage systems, visualization resources, data collections, soft-ware, and science gateways, integrated by software systems and high bandwidth networks, coordinated through common policies and operations, and supported by technology experts. The TeraGrid has been supported by funding from the National Science Foundation's (NSF) Office of Cyberinfrastructure (OCI) and significant matching funds from the participating institutions since it began as the Distributed Terascale Facility in 2001.

The TeraGrid began in 2001 when the U.S. National Science Foundation (NSF) made an award to four centers to establish a Distributed Terascale Facility (DTF). The DTF became known to users as the TeraGrid, a multi-year

---

effort to build and deploy the world's largest, fastest, most comprehensive, distributed infrastructure for general scientific research. The initial TeraGrid was homogeneous and very "griddy", with users foreseen to be running on multiple systems, both because their codes could run "anywhere", and because in some cases, multiple systems would be needed to support the large runs that were desired. The software that made up the TeraGrid was a set of identical packages on all systems. The TeraGrid has since expanded in capability and number of resource providers. This introduced heterogeneity and thus added complexity to the grid ideals of the initial DTF, as the common software no longer could be completely identical. This led to the concept of common interfaces, with potentially different software underneath the interfaces. Additionally, the users of national center supercomputers were merged into TeraGrid, which led to TeraGrid increasing its focus on supporting these users and their traditional parallel/batch usage modes.

The TeraGrid is freely available for US researchers, with allocations determined by a peer-review mechanism. The TeraGrid resources are generally at least 50% dedicated to TeraGrid. Roughly similar to the TeraGrid in terms of a continental-scale set of resources with a focus on HPC users is DEISA/PRACE in Europe, which is a set of national resources with some fractions provided to a set of European users, again determined through a review process. An alternative type of infrastructure is one that is more focused on high throughput computing (HTC), initially in high-energy physics, but increasingly in other domains as well. Filling this role in the US is Open Science Grid (OSG), and in Europe, Enabling Grids for E-sciencE (EGEE), which has recently been transformed into the European Grid Infrastructure (EGI).

This paper discusses the TeraGrid itself (Section II), examples of the interesting current uses of the TeraGrid cyberinfrastructure for science (Section III), and examples of applications that are being developed to perform science on the TeraGrid in the future (Section IV)[1].

II. THE TeraGrid

TeraGrids mission has the potential to advance the nations scientific discovery, ensure global research leadership, and address important societal issues. This is done through a three-pronged strategy, known as deep, wide, and open:

- *Deep: ensure profound impact for the most experienced users, through pro vision of the most powerful computational resources and advanced computational expertise.*

  The TeraGrid's deep goal is to enable transformational scientifoc discovery through leadership in the use of high-performance computing (HPC) for high-end computational research. The TeraGrid is designed to enable high-end science utilizing powerful supercomputing systems and high-end resources for the data analysis, visualization, management, storage, and transfer capabilities required by large-scale simulation and analysis. All of this requires an increasingly diverse set of leadership-class resources and services, and deep intellectual expertise in the application of advanced computing technologies.

- *Wide: enable scientific discovery by broader and more diverse communities of researchers and educators who can leverage TeraGrid's high-end resources, portals and science gateways.*

  The TeraGrid's wide goal is to increase the overall impact of TeraGrid's advanced computational resources to larger and more diverse research and education communities. This is done through user interfaces and portals, domain specific gateways, and enhanced support to facilitate scientific discovery by people who are not high performance computing experts. The complexity of using TeraGrid's high-end resources will continue to grow as systems increase in scale and evolve with new technologies. TeraGrid broadens its user base by providing simpler, but powerful, interfaces to resources, such as establishing common user environments and developing and hosting Science Gateways and portals. TeraGrid also provides focused outreach and collaboration with science domain research groups, and con- ducts educational and outreach activities that help inspire and educate the next generation of America's leading-edge scientists.

- *Open: facilitate simple integration with the broader cyberinfrastructure through the use of open interfaces, partnerships with other grids, and collaborations with other science research groups that deliver and support open cyberinfrastructure facilities.*

  TeraGrid's open goal is twofold: to enable the extensibility and evolution of the TeraGrid by using open standards and interfaces; and to ensure that the TeraGrid is interoperable with other open, standards-based cyberinfrastructure facilities. While TeraGrid

---

[1] The examples in this paper were chosen by the authors from the much larger set of science being performed and planned on TeraGrid for having particularly interesting interactions with cyberinfrastructure. It's likely that another set of authors writing a similar paper would choose other examples

only provides (integrated) high-end resources, it must enable its high-end cyberinfrastructure to be more accessible from, and even federated or integrated with, cyberinfrastructure of all scales. That includes not just other grids, but also campus cyberinfrastructures and even individual researcher's labs or systems. The TeraGrid leads the community forward by providing an open infrastructure that enables, simplifies, and encourages scaling out to its leadership- class resources by establishing models in which computational resources can be integrated for current and new modalities of science. This openness includes standards and interfaces, but goes further to include appropriate policies, support, training, and community building.

II.1. Organization

The coordination and management of the TeraGrid partners and resources requires organizational and collaboration mechanisms that are different from a classic organizational structure for single organizations. The existing structure and practice has evolved from many years of collaborative arrangements between the centers, some predating the TeraGrid. As the TeraGrid moves forward, the interrelationships continue to evolve in the context of a persistent collaborative environment.

The TeraGrid team comprises eleven RPs:
- Indiana University (IU),
- the Louisiana Optical Network Initiative (LONI),
- the National Center for Atmospheric Research (NCAR), the National Center for Supercomputing Applications (NCSA),
- the National Institute for Computational Science (NICS), Oak Ridge National Laboratory (ORNL),
- Pittsburgh Supercomputer Center (PSC),
- Purdue University,
- San Diego Supercomputer Center (SDSC),
- Texas Advanced Computing Center (TACC),
- University of Chicago/Argonne National Laboratory (UC/ANL)

and the GIG (at UC/ANL), which in turn issues subawards to the RPs and to additional Software Integration partners. The GIG provides coordination, operations, software integration, management and planning. GIG area directors (ADs) direct project activities involving staff from multiple partner sites, coordinating and maintaining TeraGrid central services. TeraGrid policy and governance rests with the TeraGrid Forum (TG Forum), comprising the eleven RP principal investigators and the GIG principal investigator.

TeraGrid management and planning is coordinated via a series of regular meetings, including weekly TeraGrid AD teleconferences, biweekly TG Forum teleconferences, biweekly project-wide Round Table meetings (held via Access Grid), and quarterly face-to-face internal project meetings. Management of GIG activities is led by the GIG Project Director and accomplished primarily by the GIG Area Directors. Coordination of project staff in terms of detailed technical analysis and planning, particularly across RP and GIG efforts, is done through two types of technical groups: working groups and Requirement Analysis Teams (RATs). Working groups are persistent coordination teams and in general have participants from all RP sites; RATs are short-term (6-10 weeks) focused planning teams that are typically small, with experts from a subset of both RP sites and GIG partner sites. Both types of groups make recommendations to the TeraGrid Forum or, as appropriate, to the GIG management team.

To support the great diversity of research activities and their wide range in resources needs, the TeraGrid user support and operations teams leverage the expertise from the eleven RP sites. For example, operations staff at each center talk to each other about what generally is and isn't working on their systems, and about their experiences with the TeraGrid-supplied software and services. In addition, users benefit from our coordinated external relations, education, outreach, and training activities.

II.2. Facilities

The TeraGrid's integrated resource portfolio is constantly evolving. During 2009, it included 22 high-performance computational (HPC) systems, several massive storage systems, remote visualization resources, and a dedicated inter- connection network. This infrastructure is integrated at several levels: policy and planning, operational and user support, and software and services. Policy and planning integration facilitates coordinated management and evolution of the TeraGrid environment and allows us to present – as much as possible – a single cyberinfrastructure to the user community. Operational and user support integration allows the user community to interact with one or more of the many distinct resources and HPC centers that comprise TeraGrid, through a common service, training, and support organization – masking the complexity of a distributed organization. In addition, this allows the national user community to request allocations through a single national review process and use the resources of the distributed facility with a single allocation. Software and services integration creates a user environment with standard service interfaces, lowering barriers to porting applications,

enabling users to readily exploit the many TeraGrid resources to optimize their workload, and catalyzing a new generation of scientific discovery through distributed computing modalities. Science Gateways, web-based portals that present TeraGrid to specific communities, represent a new paradigm in high-performance computing with the potential to broaden usage. Gateways benefit significantly from the integration of software and services across the TeraGrid.

TeraGrid's diverse set of HPC resources provides a rich computational science environment. As of the end of 2009, TeraGrid RPs operate 22 highly reliable HPC resources and 6 storage resources that are available via a central allocations and accounting process for the national academic community. The storage resources include highly scalable parallel systems, novel architectures, and condor-based resources. These resources are used singly or in multi-site workflows, and accessed via command line interface or science gateway. The storage systems include high speed disk, globally-accessible file systems (accessible from 7 RP sites as of 2010), and several distributed massive data storage systems. More detailed information is available on compute resources at http: //www.teragrid.org/web/user-support/compute_resources and on storage resources at http://www.teragrid.org/web/user-support/data_resources.

In 2009, TeraGrid compute resources delivered 27.4 billion NUs[2] to users representing a 255% increase over the NUs delivered in 2008.

II.3. Users and usage

The national and global user community that relies on TeraGrid has grown tremendously during the past five years, from fewer than 1 000 active users in October 2005 to more than 4 800 active users, working on almost 1 600 projects, at the end of 2009. The number of unique users submitting jobs continues to increase about 20% each year. More than 2 300 new individuals became TeraGrid users in 2009.

The group of users can be analyzed in a number of different ways. One is by their domain of science, as seen in Fig. 1. Additionally, they can be studied by the modalities in which they use TeraGrid, such as traditional batch processing on a single HPC system, metascheduling a workflow across multiple systems, and using a science gateway or portal.
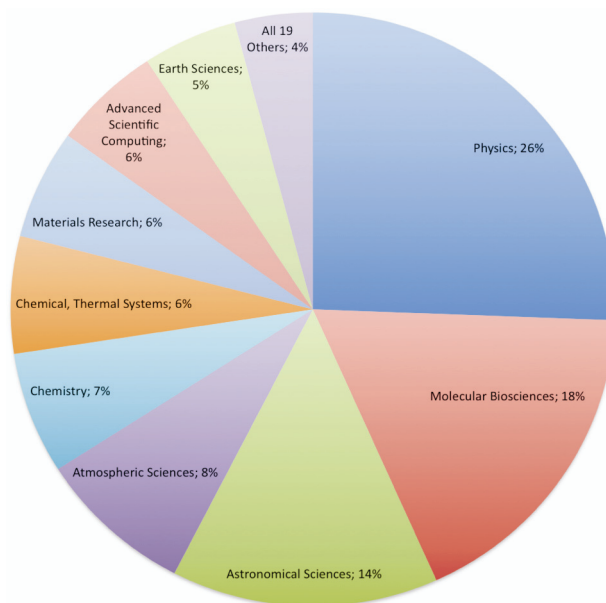
[2] NUs are normalized units, Cray X-MP-equivalent SUs. SUs are core-hours. Roughly, a core-hour on a modern (2010) large cluster can be thought of as 30-40 NUs.

Fig. 1. TeraGrid usage in 2009 by area of science

III. CURRENT SCIENCE EXAMPLES

There are a number of interesting and innovative science applications currently running on TeraGrid today. This section includes a few of those that have interesting cyberinfrastructure challenges.

Enzo, as described in Section III.1, is a good example of a large code developed by one group and used by many other groups. It runs on a variety of systems, and at some times, has been run with multiple TeraGrid resources each doing different parts of a science run (data initialization, processing, and data analysis and visualization).

LEAD II (Section III.2) runs computationally intense and large-data workflows that must be responsive to weather. It crawls data sources every five minutes to make last minute binding to the freshest weather data possible.

The SCEC example in Section III.3 is interesting because it combines two workflows, one involving a small number of medium-scale MPI jobs, and the other with a large number of sequential jobs.

GridChem (Section III.4) is a science gateway developed for computational chemists. As such, it hides the complexity of the TeraGrid behind a simple user interface. Unlike most TeraGrid science gateways, the client interface is a Java application, rather than a web portal.

GENIUS, described in Section III.5, combines real-time steering and visualization, advance reservations and a Java client interface so researchers can launch and interact with brain blood-flow simulations between the TeraGrid and the Centre for Computational Science in London, UK.

III.1. Enzo

Enzo is a community astrophysics code developed by many contributors over many years. Enzo is maintained by the Laboratory for Computational Astrophysics at the University of California, San Diego. On the TeraGrid, Enzo uses advanced used support for code development, and large-scale computational resources for simulation, archival storage, and high-capacity networking for application execution.

At the LCA, Enzo is used to simulate various aspects of large-scale structure formation in the Universe. Figure 2 shows an example visualization from Enzo. These cosmology simulations are bound by the aggregate memory of the computing systems, and by I/O capability and long-term archival storage performance and capacity. For example, a full-scale run is presently in progress on the NICS Cray XT5, Kraken, that requires the memory and processing power of almost 94 000 cores. Each re-start/checkpoint file is 30 TB in size. It is expected that the entire simulation will require tens of millions of core hours to complete and will output over 1 PB to NICS Lustre. Enzo is also used to model supersonic turbulence in the magnetized interstellar medium, where these processes are key to understanding star formation and fundamental MHD processes in space plasmas.

Enzo calculations aid the theoretical understanding that is essential to support multi-billion dollar space- and ground-based observatories such as the James Webb Space Telescope (JWST), the Large Synoptic Survey Telescope (LSST), and many others.

The present calculations are at a scale where only NICS has the necessary resources to support them, although the Teragrid 10 Gbit/s network is essential for transmitting data snapshots to SDSC/UCSD and other Teragrid systems for some types of detailed analysis. Previously, however, LCA used multiple TeraGrid resources, and components of the entire simulation were performed on the appropriate systems. Large shared memory systems are no longer required for initialization, but are increasingly important for some aspects of the final analysis. Additional facilities, such as the visualization walls at SDSC and ORNL, are very valuable for analysis and for outreach activities.

Enzo is constantly expanding, with contributions from many researchers across the USA. Several new capabilities, including multiple MHD and RHD models are being added to the new Enzo 2.0 community code, which will be released in summer 2010. These new physics capabilities call for ever-expanding computational resources at the largest scales available.
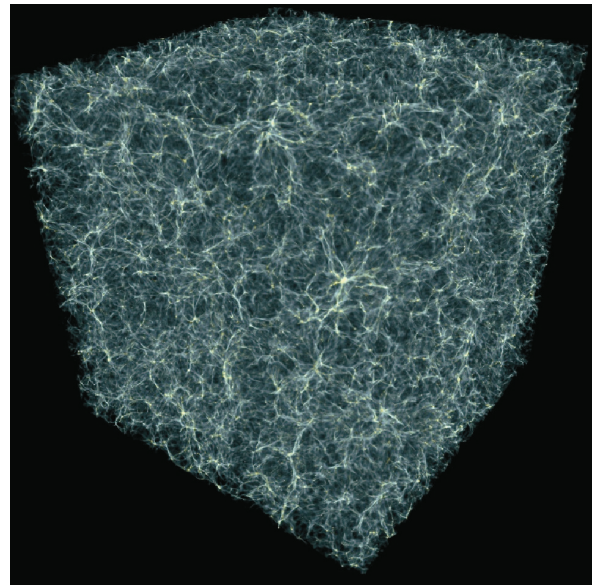

Fig. 2. This image was created at NCSA using the results from an simulation which ran in parts on NCSA Cobalt and SDSC DataStar. The model represents a volume of the Universe about 800 million light years on a side and used a 512^3 top level mesh with 7 levels of adaptive mesh refinement, plus 134 million dark matter particles

III.2. Weather prediction and Linked Environments for Atmospheric Discovery II (LEAD II)

The Linked Environments for Atmospheric Discovery II (LEAD II) [1, 2] project is a multi-disciplinary collaboration that applies advanced cyberinfrastructure to bridge the gap between environmental scientists, weather data and tools, and large-scale computational resources needed to interpret the weather. LEAD II evolved from a National Science Foundation (NSF) funded grant (2003-2009) that involved nine institutions and more than 100 scientists, students and technical staff. The LEAD II gateway has been a major consumer of computational resources on the TeraGrid through ongoing use and a number of notable initiatives. NOAA Hazardous Weather Testbed (HWT) researchers have used the LEAD II gateway to run on-demand forecasts to research severe storms during the high activity spring period. LEAD has also played an important role in promoting interest and training in science, technology, engineering, and math (STEM) disciplines. Each year, for several years, more than 1 000 students participate in a playoff style team challenge to predict weather conditions via LEAD through the Wx-Challenge Weather Forecast Challenge sponsored annually by the University of Oklahoma. Finally, LEAD II participated in the NSF funded Vortex 2 program, a six-week national nomadic field effort to understand tornadoes. LEAD II provides

hourly morning forecasting that incorporates the latest observational and model data including WSR-88D Doppler radar data which has shown a marked improvement in the forecasts, and automatic delivery and posting of images to mobile devices, including smart phones, so they can be instantly utilized in the field.

As an example of computational needs, in support of Vortex2, LEAD runs a 15-hour WRF 3.1.1 forecast using 512 cores on the TeraGrid resources (Big Red at IU) to run the WRF application, 64 cores on Big Red to run the Oklahoma University ADAS tools that assimilate data, and a 16-core Windows HPC Server box to generate 2D visualizations of eight variables including helicity, vorticity, and precipitation, using the NCAR Command Language. This end-to-end pipeline must be accomplished within a strict one-hour timeline, which places rigorous demands on the data subsystem to retrieve and index large amounts of realtime weather data frequently so the data is available to assimilate into the model the moment the model is ready to run. An output from a forecast of this type is shown in Fig. 3.
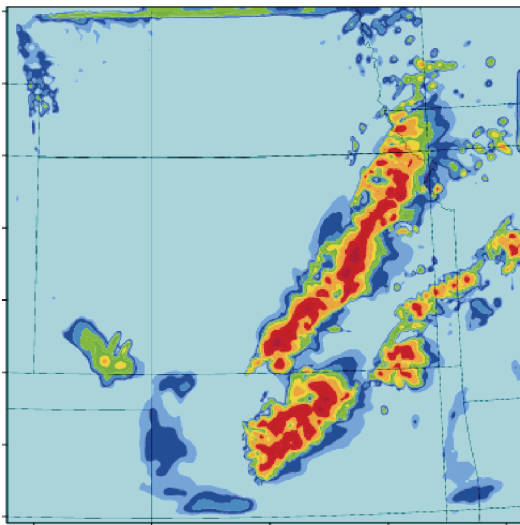

Fig. 3. Model-generated radar results of a storm that caused severe tornado damage in Norman, Oklahoma. The forecast is generated using ADAS 5.10, WRF 3.1.1, and NCL to generate the visualizations, and runs on the LEAD II gateway

LEAD II pioneered XBaya workflow composer [3], XMC Cat metadata catalog [4], and the GFAC service wrapper [5]. The workflows are computationally intense and take several hours to run the "typical workflow" that has forecast duration of greater than 36 hours, and then ingest and produce gigabyte-sized data files. The system is architected as a Service Oriented Architecture (SOA) and is written in Java. Application tasks, that is, nodes in the workflow, are added through a web tool provided as part of

GFAC that "wraps" the application in a web service wrapper so that it can be invoked through the SOA. Communication in the architecture is through asynchronous and synchronous web service calls, and through an eventing system that allows simultaneous notification of events in the system. During execution of a workflow, data is passed from node to node primarily through files, but key contextual information is accumulated along the workflow path in a notification and then converted to a file that the application service accepts.

The workflows are largely dataflow-driven. We build upon the conventional workflow systems used in scientific computing and extend them with support for data streams. Our goal is to extend the existing support for workflow composition (e.g., XBaya) in a way that preserves the workflow programming model for the user. Users are familiar with directed acyclic graph (DAG) execution. Stream processing is an eventing paradigm that, when put side by side in a graphical interface during the task of workflow composition, tends to confuse the scientists. LEAD II integratea continuous events processing into the familiar DAG model while preserving the familiar DAG control flow [6].

3.3. SCEC

The Southern California Earthquake Center (SCEC) uses the TeraGrid's large-scale resources to advance their program of earthquake system science research. SCEC scientists have developed a new technique to use full 3D waveform modeling to perform probabilistic seismic hazard analysis (PSHA) more accurately than with traditional empirically-based attenuation relationships. SCEC's PSHA computational platform, CyberShake [7, 8], enables seismologists to estimate the probability that ground motions will exceed an intensity measure, such as peak ground acceleration, over a given time period. This seismic hazard information is reported in a hazard curve (Fig. 4). These kinds of seismic hazard estimates are useful for civic planners, building engineers, and insurance agencies and can inuence billions of dollars of construction.

To manage the CyberShake calculations, SCEC utilizes distributed, grid-based workflows using NSF-supported HPC middleware including Pegasus, Condor, and Globus to manage job submission and execution and provide error recovery. Calculating a CyberShake PSHA hazard curve for a site of interest requires executing two workflows. The first contains several large parallel jobs, while the second is made up of a large number of loosely coupled, short-running jobs. The current CyberShake platform executes about 840 000 jobs, using 17 000 CPU-hours to generate

a single hazard curve. However, PSHA curves are often calculated for hundreds of sites and synthesized into a hazard map, representing the hazard for an entire region. To meet SCEC's goal of producing the world's first physics-based PSHA map for Southern California, curves needed to be generated for over 200 sites in the Los Angeles area.
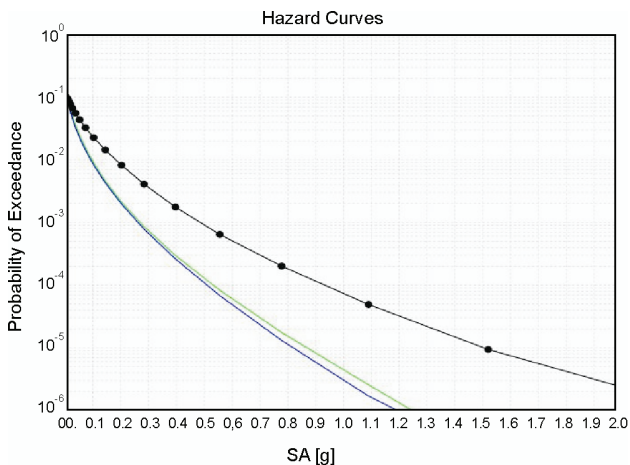

Fig. 4. Three hazard curves for Whitter Narrows, near Los Angeles. The black line is the physics-based CyberShake simulation results. The blue and green lines are two common empirically-based attenuation relationships
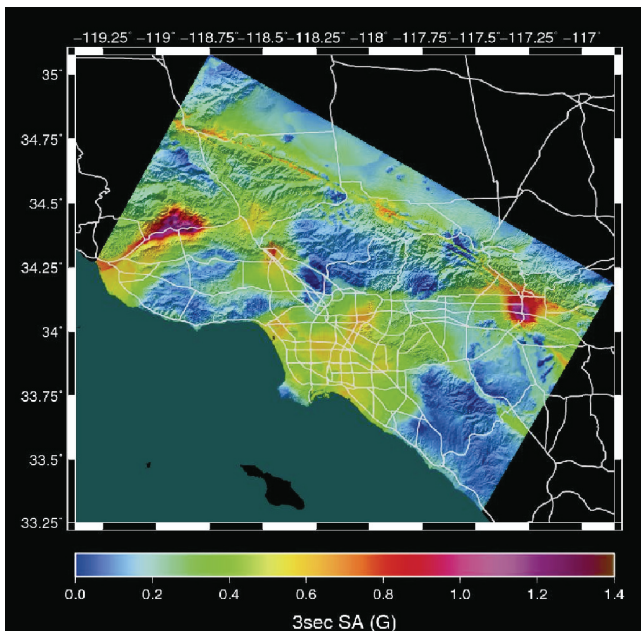

Fig. 5. CyberShake hazard map; 2% probability of exceedance in 50 years

Executing CyberShake at this scale requires sustaining high throughput for the large number of short-running jobs.

SCEC uses TeraGrid systems with large core counts, such as Ranger and Kraken, to take advantage of the inherent parallelism. Grid computing tools are used to transfer data in, and results out, from the TeraGrid execution platform, and to support remote job submission. SCEC's largest CyberShake run in 2009 used Ranger for execution. Over a two-month period, CyberShake ran on an average of 4 500 cores on Ranger, using 5 million CPU-hours and executing 190 million code invocations. 223 hazard curves were calculated and combined into a hazard map (Fig. 5). Facilities such as those provided by the TeraGrid are key to running both high performance and sustained high throughput applications, enabling SCEC to produce transformative science.

III.4. GridChem

The "Computational Chemistry Grid" (CCG), a.k.a. GridChem [9, 10], is a virtual organization that provides access to high performance computing resources for computational chemistry with distributed support and services, intuitive interfaces and measurable quality of service. GridChem is a Science Gateway; it allows users access to resources for their work while hiding the details of both the resources and the access mechanisms from them, allowing them to focus on the science. The CCG client, GridChem, is a Java desktop application that provides an interface to integrate the hardware, software and middleware resources necessary to solve quantum chemistry problems using grid technologies. Currently available applications via the GridChem science gateway include Gaussian, Gamess, NWChem, Molpro, ADF, DMOL3, CASSTEP, QMCpack, and Amber, and other programs will be available in the future. In general, these are applications with which the user is familiar; having them accessible on the TeraGrid from GridChem allows the user to solve larger problems than they could by using the application locally, but with less effort to actually run the application. All applications are site-licensed to the HPC providers; a GridChem science gateway user does not have to deal with the licensing issue of the software and systems. One of the important recent developments of the GridChem science gateway is the ongoing incorporation of the Xbaya workflow composer [3] into the gateway. This will be beneficial for complicated scientific calculations that are effectively executed as a workflow, rather than as a set of single steps that are manually run by the user.

III.5. GENIUS

Cardiovascular disease is the cause of a large number of deaths in the developed world. Cerebral blood flow behavior plays a crucial role in the understanding, diagno-

sis and treatment of the disease, where problems are often due to anomalous blood flow behavior in the neighborhood of bifurcations and aneurysms within the brain. The GENIUS (Grid Enabled Neurosurgical Imaging Using Simulation) project uses blood flow modeling on the TeraGrid along with patient-specific medical imaging data. These are fundamentally used to help understand blood flow patterns in the presence of neurovascular pathologies. More powerfully however, it provides a tool which researchers and clinicians can use in realtime to help plan courses of surgical treatment for the embolization of arterio-venous malformations and aneurysms, amongst other neuropathology. To achieve the science aims of GENIUS, and more broadly, to enable patient-specific medical simulation to work in a real-world setting, not only is the correctness of the results important, the timeliness of results is imperative to the success of simulation methodologies [11]. Various facets of high-performance computing have been used to achieve these aims that run beyond the batch-submission scenario typical in high-performance computing today.

At the core of GENIUS is HemeLB [12], a lattice-Boltzmann (LB) solver that has been specifically designed to address fluid flow in sparse geometries such as the neuro-vasculature. The vasculature is decomposed into a number of computational domains for parallel execution using MPI. Computation is overlapped with communication between different processor domains, resulting in an essentially linear scaling up to 1024 processors and beyond on machines such as Ranger on the TeraGrid [13].

Rather than relying on post-processing visualization, an in-situ volume-rendering approach is used within HemeLB where the ray-traced domains correspond to the LB computational domains. Sub-images are composited to form a complete frame, and transmitted over the network to a lightweight client, resulting in immediate real-time visualization of the bloodflow simulation as the MPI simulation is running. The lightweight client allows a researcher or clinician to steer HemeLB in realtime, where physical parameters of the vascular system along with various visualization properties can be adjusted in real-time (rotation, zoom, color mapping, opacity etc.). From the computational lab at University College London to the Ranger cluster, 25 frames per second was readily achievable [14]. The scalability of the algorithm results in real-time rendering capabilities, even for high-resolution datasets. An image from such a rendering is shown in Figure 6.

To conduct investigations into whole brain blood flow simulation, for which a single supercomputer does not typically have enough CPU resources, HemeLB has been combined with MPIg, the successor to MPICH-G2 [15],

a version of MPI that allows simulations to be distributed across multiple resources, using the Globus middleware for intermachine communications. Distributing MPI in this manner also allows conglomeration of various resources for the ability to rapidly run important medical simulations, rather than waiting for the availability of a single large resource.

For scenarios that require co-allocation, such as distributed MPI, as well the realtime interactive nature of HemeLB, advance reservations play a crucial role in achieving the scientific aims [16]. Two systems exist to allow users to co-reserve time on the TeraGrid. GUR (Grid Universal Remote) [17] is one such system, developed at SDSC. The other, HARC (Highly Available Robust Co- scheduler), is a widely deployed open-source system that allows users to reserve multiple distributed resources in a single step [18], and is used extensively by the GENIUS project.
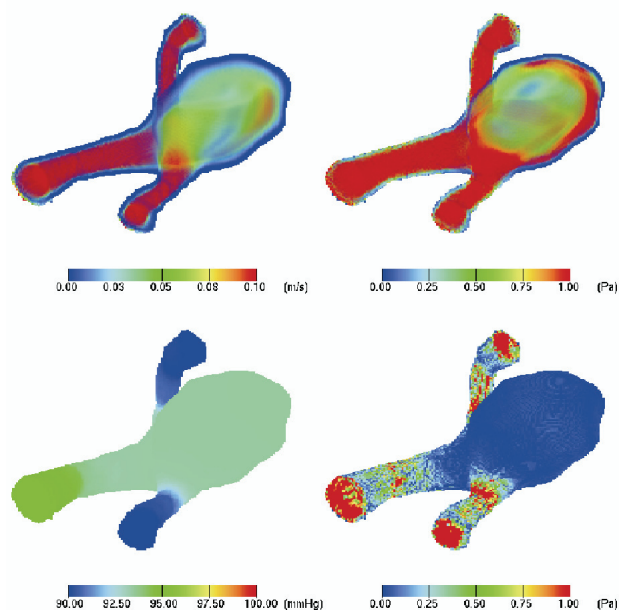

Fig. 6. An example snapshot of the interactive visualization simulation results obtained with the HemeLB in situ rendering system. For this patient-specific intracranial aneurysm, the volume rendering of the blood flow velocity and stress flow fields half way through a pulsatile cycle are depicted on the top-left and top-right images. The wall pressure and wall stresses are shown in the bottom-left and bottom-right images

Bringing together the various requirements – job launching, file transfer, computational steering, real-time visualization and advance reservations – is the Application Hosting Environment (AHE) [19], providing researchers with a single and simple middleware tool to access patient-specific computational haemodynamics on the TeraGrid.

IV. FUTURE SCIENCE EXAMPLES

This section contains science projects that are examples of new ways of using cyberinfrastructure.

PolarGrid, as described in Section IV.1, is investigating the use of Google-led OpenSocial gadget components and online social networking services (Google's GData-based services, Facebook, and Twitter) for building science gateways and scientific community building. Another innovative use of gadgets is occurring at Purdue, where a team is building gadgets to automatically provide satellite data, as shown in Section IV.2.

Obtaining scale-out across multiple resources has been very challenging currently, and in Section IV.3, we discuss several applications that are attempted to use SAGA to scale-out on the TeraGrid.

Finally, beyond scaling out across a single grid, a number of interesting examples of applications and infrastructure that can interoperate across different grids are being developed, with some examples discussed in Section IV.4.

IV.1. The PolarGrid Gateway: Combining TeraGrid and Web 2.0

PolarGrid is an NSF-funded MRI project that provides computing support for the Center for the Remote Sensing of Ice Sheets (CReSIS[3]). One of CReSIS's primary activities [20] is to use Synthetic Aperture Radar (SAR) to obtain information on the depths of the Greenland and Antarctic ice sheets and to map out the rock beds underlying the glaciers. Detailed knowledge of these rock beds is needed to develop improved models of glacial movement and melting. Terabytes of raw SAR data from each expedition must be processed with signal image processing software to produce data products for distribution to the CReSIS community. PolarGrid provides both in-the-field computing clusters for initial image processing (useful for finding problems with radar equipment, for example) and larger clusters on the TeraGrid for full-scale image processing needed to make initial ("level 0") community data products. This initial processing pipeline is fixed and has little need for user-driven job submission typical of science gateways. However, the "level 0" imagery data products are quite noisy and need to be additionally processed by human-driven filtering steps before the rock bed signals can be extracted. These filtering applications form the basis of the PolarGrid science gateway. This is described in greater detail in [21]. A summary description is provided here.

---

[3] https://www.cresis.ku.edu/

Science gateways [22, 10] are typically composed of user interface components managed by a container, with computing resource access managed by Web services invoked by these component clients. Each component is dedicated to a particular, self-contained functionality, such as composing and launching a computational job. The component's container manages layouts of the components into dashboards, user accounts, group memberships, and similar features. Web Services running separately from the container's server are responsible for managing jobs, files, and information about resources.

The PolarGrid gateway follows this general model. However, in addition to providing typical gateway capabilities, the PolarGrid portal also is used to evaluate the applicability of several recent trends in Web development to science gateways. Specifically, PolarGrid uses Open Social-compatible Google Gadgets as its component model. Data filter services are implemented using Matlab and then wrapped as REST services using the OGCE's GFAC tool. Finally, PolarGrid offloads identity management and social networking tools to third party services. For example, it supports OpenID logins: after creating an account, a user can associate this account with any recognized OpenID provider. This association is under the user's control and not set by the portal provider. For common gateway services such as calendars, event updates, image repository management, and discussion forums, PolarGrid uses Google Calendar, Twitter, Google Picasa, and Google FriendConnect, respectively. PolarGrid's motive in making these choices is to make a collaborative framework that is sustainable beyond the life of this specific gateway instance.

Integration with these third party services is done through both the GFAC-based submission services and the gadget user interfaces. For example, an invocation of a filter run will invoke Twitter as a service to post metadata about the run. The images made by the filter are uploaded to a Picasa image repository. The user interfaces to both of these services are off-the-shelf Google Gadgets that have been imported into the gadget container. The basic processing flow is as follows:

1. A user visits his/her gadget home page, which is served by gadget layout manager
2. The gadget layout manager constructs the user's custom gadget layout in their browser and makes use of a gadget renderer (Shindig in this case) to render each gadget XML to HTML/JavaScript. Then the generated HTML/JavaScript code is displayed in the user's browser.
3. Different gadgets may interact with different backend RESTful services to generate output. A JSON

response is sent back to the gadget to display the results.

4. Gadgets and RESTful services also query social data using OpenSocial API's by sending requests to Shindig server.

Although the PolarGrid gadgets can be deployed in iGoogle, PolarGrid primarily uses a gadget container from the OGCE project developed by the authors of [21]. This is displayed in Fig. 7. The OGCE gadget container is a downloadable piece of software built on top of Apache Shindig that can be used to host both locally deployed gadgets as well as remote gadgets registered with iGoogle. The gadget container supports tabbed layouts for organizing gadgets, skins and themes, OpenID logins, and OAuth secured gadgets. The gadget container also supports HTTPS connections between the gadgets and the container as well as between the user's browser and portal server.

IV.2. Realtime Satellite Data Gadgets

Purdue University, one of the eleven TeraGrid resource providers, hosts a number of data collections in support of Earth and environmental research and education. These data collections, accessible through the Purdue data management system[4], are linked with processing and visualization capabilities provided by TeraGrid resources. Among the observational data, Purdue makes available the remote sensing satellite data from the Purdue Terrestrial Observa-

tory's[5] two satellite ground stations, which are capable of continuous ingestion of remote sensing data from the NOAA GOES-12, NASA MODIS Terra/Aqua, NOAA AVHRR, and Fengyun-1D MVISR sensor systems. Up to 30 GB of raw data are received and processed each day. A large number of satellite data products can be generated from the raw satellite observations. For instance, data from the MODIS sensor on the Terra and Aqua satellites can generate 14 different products. Each data product may contain up to 50 different variables. These data products describe features of the land, oceans, and atmosphere and are used to study processes and changing, at both the local and global scales, in many study areas, including atmospheric science and climatology, hydrology, agriculture, ecology, transportation and logistics, aviation and air traffic safety, and disaster and emergency management.

To help potential users easily discover and access the available satellite data products, the Purdue team developed viewer gadgets as a new way to deliver satellite data and information to a broad user community. The satellite data viewer gadgets deliver real time satellite data continuously in near real-time to the user via a compact and yet rich featured, interactive web interface. A user may select from a list of products (e.g., cloud cover, water vapor, etc.) and a time period (e.g., 12 or 24 hours). Behind the scene, the gadgets connect to a data generation and processing system where a video consisting of a sequence of images for the

Fig. 7. PolarGrid Science Gateway is composed of Google Gadgets running in the OGCE gadget container. Gadgets displayed in the figure include (top left to bottom right) the Filter Launching gadget, a Picasa gadget showing results of a previous filter run, a Twitter gadget (showing filter job submission events), an Atom feed showing blog updates associated with the job run, a calendar, and a discussion gadget based on FriendConnect

requested time period is generated and streamed to the user. These gadgets changed the way data is typically provided to the users: the data is pushed in near real-time to user's web page, instead of waiting for the user to come to the data provider. The viewer gadgets are linked back to the Purdue satellite data science gateway PRESTIGE[6] where users may learn more about the available data products, view product images, and subscribe to the products of their interests [23].

Purdue released two Flash-based satellite data viewer gadgets to the Google gadget directory in August 2009,

Fig. 8. GOES-12 viewer gadget

leveraging Google's infrastructure for scientific outreach. The GOES-12 satellite data viewer (Fig. 8) streams an animation of three types of products (cloud cover, cloud pressure and water vapor) to user's gadget page (e.g., iGoogle). As the real-time GOES-12 data is received at PTO at 15-minute intervals, the updated video is automatically delivered to the user. The MODIS satellite data viewer gadget (Fig. 9) delivers seven days worth of data products in a video, including the fire detection and sea

surface temperature. In these gadgets, users can pause, zoom, and seek an area of interest. During the seven months since the two gadgets were released to the Google gadget directory, they have been visited more than 18 936 times by users from 63 countries according to Google Analytics.

Fig. 9. MODIS viewer gadget

IV.3. EnKF Reservoir Simulations

Ensemble Kalman Filters (EnKF) are recursive filters that can be used to handle large, noisy data; the data are sent through the Kalman filter to obtain the true state of the data. EnKF are the kernel for a wide range of science and engineering applications, ranging from CO_2 sequestration to simple reservoir engineering problems (history matching). The Kalman filter stage is where the data assimilation is performed and compared to the real measured value from the oil-field (or historical data from well production logs) with different models. Models are modified so that they are "closer" to the real produced value. After assimilation of all the historical data, a set of models are generated that are representative of what the actual oil reservoir looks like. This approach can be extended to study CO_2 sequestration. The history of a depleted/abandoned oil field is matched to

---

[6] http://www.purdue.teragrid.org/prestige

get its properties and inject CO_2 based on the newly determined good reservoir description. Using the EnKF formulation, a scientific problem is solved using multiple models (ensemble members).

The physical models are represented as ensembles with members that vary in size from large MPI-style jobs to long-running single processor tasks. Varying parameters sometimes also lead to varying systems of equations and entirely new scenarios, increasing both computational and memory requirements. Because all models in a stage must complete before the next stage can begin, dynamic load-balancing can be used to ensure that all models complete as close to each other as possible. In the general case, the number of jobs required varies between stages. The resource requirement for these applications is dynamic and unpredictable, and it is difficult to initially define a static scheduling strategy that will be effective throughout the execution of a complete application.
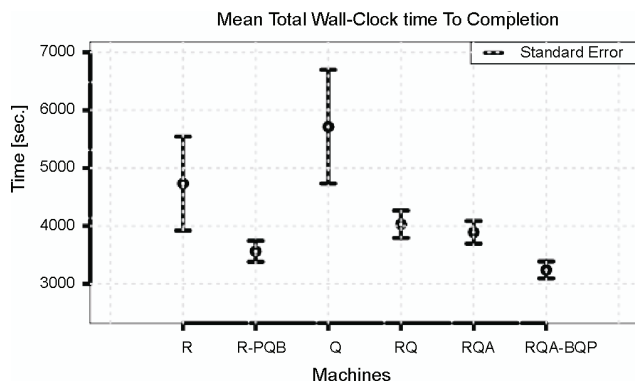

Fig. 10. Time to completion for the Ensemble Kalman Filter application. From left to right: Ranger; Ranger when using BQP (batch queue prediction, which predicts how long a job will wait in a queue on a system); QueenBee; Ranger and QueenBee; Ranger, QueenBee and Abe; Ranger, QueenBee and Abe when using BQP [28]

EnKF-based applications must be able to scale-up (scale to use larger fractions of a single machine) as well as scale-out (scale to use larger numbers of different machines), because the relative importance of these two types of scaling changes with the specific problem under investigation. Ensuring optimal time-to-solution requires that scaling-up be balanced with scaling-out. To decrease the time-to-solution for increasingly larger and more realistic physical models, a broad range of resource types – HPC grids (TeraGrid), HTC grids (OSG), and immediately available resources (clouds) – can be used to effectively match resources to application requirements.

The SAGA [24] programming system provides a high level API that forms a simple, standard and uniform interface for the most commonly required distributed functionality. SAGA can be used to program distributed applications [25, 26] or tool-kits to manage distributed applications [27], as well as implement abstractions that support commonly occurring programming, access and usage patterns. Figure 10 shows how a SAGA-based solution has successfully used multiple TeraGrid resources – Ranger, Abe, and QueenBee – to improve time-to-solution for an Ensemble Kalman Filter application.

IV.4. Combining TeraGrid with other Resources

The Polish Ministry of Science has invested in supporting international collaboration between Poznan Supercomputing and Networking Center (PSNC) and TeraGrid partners. A goal of this project is the integration of PSNC computing environment and PL-GRID applications (the Polish national grid infrastructure) with highly efficient computing resources at TeraGrid over fiber optic high-speed networks provided by PIONIER [29], Geant3 and Internet2 infrastructures in Poland, Europe, and the United States respectively. Consequently, new virtual testbeds and virtual collaborative environments can be created to allow researchers to perform tests of advanced grid middleware tools and applications in the above mentioned areas.

One area in which this is being applied is in UCoMS, the Ubiquitous Computing and Monitoring System (UCoMS) for discovery and management of energy resources. UCoMS addresses key research issues to arrive at appropriate technical solutions in the areas of HPC computing, sensor network systems, and distributed monitoring software. The technical solutions are expected to effectively facilitate drilling and operational data logging and processing, on-platform information distribution and displaying, infrastructure monitoring/intrusion detection, and management of complex surface facilities and pipelines. UCoMS is supported in part by the U.S. Department of Energy (DOE) and by the Louisiana Board of Regents, and it has been recently extended by Polish Ministry of Science. This is leading scientists at PSNC to closely cooperate with research teams at Louisiana State University (LSU), Center for Computation and Technology (CCT), and University of Louisiana at Lafayette to deal with challenging problems, including real-time high resolution video processing, large scale computing fluid dynamics, parallel visualization and data analysis, remote computer tomography and structural analysis, and advanced web-based science gateways, as shown in an example in Figure 11. Moreover, further development and
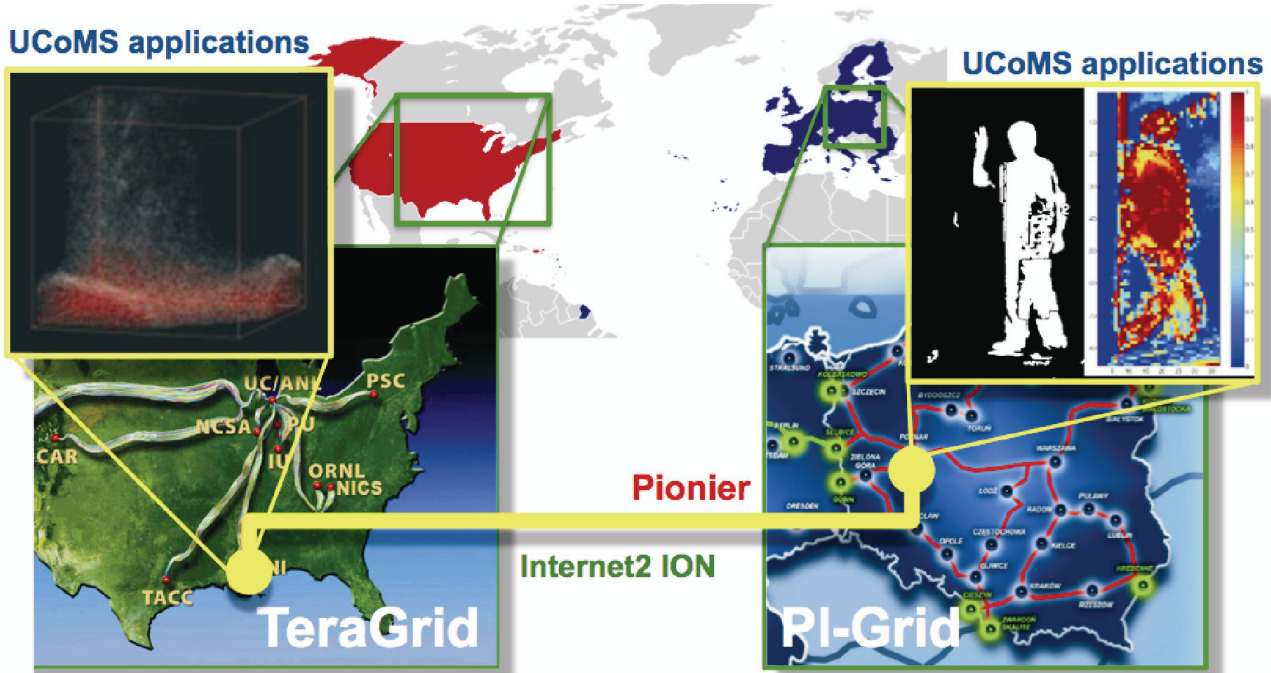
Fig. 11. An experimental cross-continental testbed established between TeraGrid and PSNC resources in the UCoMS project

deployment of grid services, tools and applications created at PSNC in EU-funded projects, such as GridLab [30], QosCosGrid [31] and MAPPER, are planned as joint activities with some developers using TeraGrid.

There are also other ongoing efforts to provide both infrastructure-level and application-level interoperability between different grids (e.g., TeraGrid, DEISA, EGEE/-EGI, and NAREGI). The argument for doing this is that federated grids can have the potential to facilitate scientific advances that would not otherwise be possible. These advances can arise from the advantages that federated resources provide, such as aggregated higher throughput, lower time-to-solution, exploitation of opportunistic scheduling, and larger peak-resource utilization at otherwise unreachable scales. Application-level interoperability can have different interpretations, but in this case it is defined as the ability to utilize multiple machines without refactoring for the specifics of resources.

One scientific project in the area of application-level interoperability is *Rapid and accurate calculation of Free-Energy for Bio-molecular System*[7]. It aims to address a significant problem in the treatment of HIV: the development of drug resistant mutant strains. In order to tailor treatment to individual patients, clinicians routinely genotype the virus infecting a patient. Knowledge-based decision

support software built on existing clinical records and associated viral genotypic data is then used to aid drug selection. This project seeks to enhance the existing decision support rules by using predictive molecular dynamics simulations of the available drugs with their enzymatic targets (the HIV-1 protease and reverse transcriptase enzymes). Specifically, the project aims to use several Replica-based and Replica-Exchange simulations of HIV-1 and multiple TeraGrid, LONI, and DEISA resources to rapidly compute free-energies of binding with high-levels precision.

There are also ongoing studies to use the federated capacity of the TeraGrid with EGEE/EGI in the field of lattice QCD and with the Japanese NAREGI in medical imaging and particle physics. All three of these use the multi-platform support that SAGA provides, either directly by simple SAGA-based job submission (TeraGrid-NAREGI), by SAGA-based abstractions (cross-platform Pilot-Jobs on TeraGrid-DEISA), or by existing tools that have been modified to use SAGA (GANGA-DIANE for TeraGrid-EGEE/EGI) interoperability.

V. CONCLUSIONS

The TeraGrid is both an operational infrastructure and a research project. As an infrastructure, it's a facility that

---

[7] http://www.teragridforum.org/mediawiki/index.php?title=LONI-TeraGrid-DEISA_Interoperabilty_Project

some users use directly, while others build tools that are then used by larger groups. As a research project, TeraGrid is investigating how best to enable a diverse range of science projects. This paper has discussed the range of users and uses of TeraGrid, ranging from large- scale applications that use one or more TeraGrid systems for modeling and simulation to portals and gateways that enable communities to use TeraGrid without having to understand the details to applications that require multiple 22 sets of infrastructure to be federated.

Many traditional applications, such as Enzo, are starting to have more complicated needs, since running a code on a single computer is often no longer the only bottleneck. Such applications are really parts of a relatively simple workflow that involves data generation, simulation, data analysis and archiving. Many of these components are also supported on TeraGrid, though not always on the same system or even at the same one on which the simulation itself ran.

By providing services such as large-data transfers, advance reservations, different types of infrastructure (storage, compute, visualization), grid-computing/standardized access interfaces, etc. the TeraGrid can give researchers the flexibility to enrich their science by developing interesting workflows (new ways of doing research). By using standard interfaces, researchers, communities and entire research disciplines can start building their tools/workflows/- applications on top of this nationally-distributed infrastructure.

Science gateways represent an important mechanism for broadening participation. Gateways themselves must track larger developments in the Web community. In particular, important Web trends over the last five years include Representational State Transfer (REST) services, RSS/Atom feeds as a machine-to-machine messaging format, social networks, reusable Web gadget components, and an emphasis on richer interactivity through significant improvements in JavaScript and related client side technologies. Some projects, such as PolarGrid, represent an effort to investigate building science gateways that integrate gateway-specific services with general-purpose third party collaboration services. This leads to a greater integration of a user's various roles in an online world.

When both researchers and resources are distributed with independent infrastructures, federation is needed at either the infrastructure or applications level. The ongoing work between TeraGrid and other infrastructures such as PL-Grid and DEISA is a start towards understanding the issues that must be solved in order to allow scientists to make progress.

References

[1] K.K. Droegemeier, D. Gannon, D. Reed, B. Plale, J. Alameda, T. Baltzer, K. Brewster, R. Clark, B. Domenico, S. Graves, E. Joseph, D. Murray, R. Ramachandran, M. Ramamurthy, L. Ramakrishnan, J.A. Rushing, D. Weber, R. Wilhelmson, A. Wilson, M. Xue, S. Yalda, *Service-oriented environments for dynamically interacting with mesoscale weather.* Computing in Science and Engg. 7 (6), 12-29 (2005).

[2] B. Plale, D. Gannon, Y. Huang, G. Kandaswamy, S. Lee Pallickara, A. Slominski, *Cooperating services for data-driven computational experimentation.* Computing in Science and Engineering 7, 34-43 (2005).

[3] S. Shirasuna, *A Dynamic Scientific Workflow System for the Web Services Architecture.* PhD thesis, Indiana University, September 2007.

[4] S. Jensen, B. Plale, S*chema-independent and schema-friendly scientific metadata management.* In ESCIENCE '08: Proceedings of the 2008 Fourth IEEE International Conference on e-Science, pages 428-429, Washington, DC, USA, 2008. IEEE Computer Society.

[5] G. Kandaswamy, L. Fang, Y. Huang, S. Shirasuna, S. Marru, D. Gannon, *Building web services for scientific grid applications.* IBM J. Res. Dev. 50 (2/3) 249-260 (2006).

[6] Ch. Herath, B. Plale, *Streamow – programming model for data streaming in scientific workflows.* In Proceedings of the 10th IEEE/ACM Int'l Symposium on Cluster, Cloud, and Grid Computing (CCGrid 2010) 2010.

[7] S. Callaghan, E. Deelman, D. Gunter, G. Juve, P. Maechling, Ch. Brooks, K. Vahi, K. Milner, R. Graves, E. Field, D. Okaya, T. Jordan, *Scaling up workflow-based applications.* J. Comput. System Sci., 2010 (Special issue on scientific workflows, in press).

[8] R. Graves, T. Jordan, S. Callaghan, E. Deelman, E. Field, G. Juve, C. Kesselman, P. Maechling, G. Mehta, K. Milner, D. Okaya, P. Small, K. Vahi, *Cybershake: A physics-based seismic hazard model for southern California.* Pure Applied Geophys., 2010 (accepted for publication).

[9] R. Dooley, K. Milfeld, Ch. Guiang, S. Pamidighantam, G. Allen, *From proposal to production: Lessons learned developing the computational chemistry grid cyberinfrastructure.* Journal of Grid Computing, 4 (2), 195-208 (2006).

[10] N. Wilkins-Diehr, D. Gannon, G. Klimeck, S. Oster, S. Pamidighantam, *TeraGrid science gateways and their impact on science.* Computer 41 (11), 32-41 (2008).

[11] S.K. Sadiq, M.D. Mazzeo, S.J Zasada, S. Manos, I. Stoica, C. V Gale, Simon J Watson, P. Kellam, S. Brew, P.V Co-

veney, *Patient-specific simulation as a basis for clinical decision-making*. Philosophical Transactions of the Royal Society A, 366 (1878), 3199-3219 (2008).

[12] M.D Mazzeo, P.V Coveney. *Hemelb: A high performance parallel lattice-Boltzmann code for large scale fluid flow in complex geometries.* Computer Physics Communications 178 (12), 894-914 (2008).

[13] R.S Saksena, B. Boghosian, L. Fazendeiro, O.A. Kenway, S. Manos, M.D. Mazzeo, S.K. Sadiq, J.L Suter, D. Wright, P. V Coveney, *Real science at the petascale*. Philos T R Soc A 367 (1897) 2557-2571 (2009).

[14] M.D Mazzeo, S Manos, P.V Coveney, *In situ ray tracing and computational steering for interactive blood flow simulation.* Computer Physics Communications 181, 355-370 (2010).

[15] N. Karonis, B. Toonen, I. Foster, *A grid-enabled implementation of the message passing interface*. Journal of Parallel and Distributed Computing (JPDC) 63 (5), 551-563 (2003).

[16] S. Manos, S. Zasada, P.V. Coveney. *Life or Death Decision-making: The Medical Case for Large-scale, On-demand Grid Computing.* CTWatch Quarterly Journal 4 (2), 35-45 (2008).

[17] K. Yoshimoto, P. Kovatch, P. Andrews, *Co-scheduling with usersettable reservations*. In D.G. Feitelson, E Frachtenberg, L Rudolph, U. Schwiegelshohn, editors, Job Scheduling Strategies for Parallel Processing 146-156, Springer Verlag, Lect. Notes Comput. Sci. 3834 (2005)

[18] J. MacLaren, M. Mc Keown, S. Pickles. *Co-Allocation, Fault Tolerance and Grid Computing.* In Proceedings of the UK e-Science All Hands Meeting 2006, 155-162 (2006).

[19] P.V. Coveney, R.S. Saksena, S.J. Zasada, M. McKeown, S. Pickles, *The application hosting environment: Lightweight middleware for grid-based computationa*l science. Computer Physics Communications 176 (6), 406-418 (2007).

[20] S. Gogineni, D. Braaten, Ch. Allen, J. Paden, T. Akins, P. Kanagaratnam, K. Jezek, G. Prescott, G. Jayaraman, V. Ramasami, C. Lewis, D. Dunson, *Polar radar for ice sheet measurements (PRISM).* Remote Sensing of Environment 111 (2-3), 204-211 (2007) (Remote Sensing of the Cryosphere Special Issue).

[21] Z. Guo, R. Singh, M. Pierce, *Building the PolarGrid portal using web 2.0 and OpenSocial*. In GCE '09: Proceedings of the 5th Grid Computing Environments Workshop 1-8, New York, NY, USA, ACM (2009).

[22] J. Alameda, M. Christie, G. Fox, J. Futrelle, D. Gannon, M. Hategan, G. Kandaswamy, G. von Laszewski, M.A. Nacar, M. Pierce, E. Roberts, Ch. Severance, M. Thomas, *The open grid computing environments collaboration: portlets and services for science gateways: Research articles.* Concurr. Comput.: Pract. Exper. 19 (6), 921-942 (2007).

[23] R. Kalyanam, L. Zhao, T. Park, L. Biehl, C.X. Song, *Enabling useroriented data access in a satellite data portal.* In Proceedings of the 3rd International Workshop on Grid Computing Environment (2007).

[24] T. Goodale et al., *A Simple API for Grid Applications (SAGA)*. http://www.ogf.org/documents/GFD.90.pdf.

[25] Shantenu Jha, Hartmut Kaiser, Yaakoub El Khamra, and OleWeidner. De- sign and implementation of network performance aware applications using saga and cactus. e-Science and Grid Computing, International Conference on, 0:143{150, 2007.

[26] S. Jha, Y. El Khamra, H. Kaiser, O. Weidner, A. Merzky, *Developing adaptive scientific applications with hard to predict runtime resource requirements*. In Proceedings of TeraGrid 2008 Conference (2008).

[27] A. Luckow, S. Jha, A. Merzky, B. Schnor, J. Kim, *Reliable Replica Exchange Molecular Dynamics Simulation in the Grid using SAGA CPR and Migol.* In Proceedings of UK e-Science 2008 All Hands Meeting, Edinburgh, UK (2008).

[28] Y. El-Khamra, S. Jha, *Developing autonomic distributed scientific applications: a case study from history matching using ensemble kalman-filters*. In Proceedings of the 6th International Conference on Autonomic Computing (ICAC); Industry session on Grids meets Autonomic Computing 19-28, New York, NY, USA (2009). ACM.

[29] A. Binczewski, N. Meyer, J. Nabrzyski, S. Starzak, M. Stroiński, J. Węglarz, *First experiences with the polish optical internet.* Comput. Netw. 37 (6):747-759 (2001).

[30] M. Kosiedowski, K. Kurowski, C. Mazurek, J. Nabrzyski, J. Pukacki, *Workflow applications in gridlab and progress projects: Research articles*. Concurr. Comput.: Pract. Exper. 18 (10), 1141-1154 (2006).

[31] K. Kurowski, W. Back, W. Dubitzky, L. Gulyás, G. Kampis, M. Mamonski, G. Szemes, M. Swain, *Complex system simulations with qoscosgrid.* In ICCS '09: Proceedings of the 9th International Conference on Computational Science, Berlin, Heidelberg, Springer-Verlag, 387-396 (2009),.

DANIEL S. KATZ received his B.S., M.S., and Ph.D degrees in Electrical Engineering from Northwestern University, Evanston, Illinois, in 1988, 1990, and 1994, respectively. His research interests include: numerical methods, algorithms, and programming applied to supercomputing, parallel computing, cluster computing, distributed computing, and embedded computing; and fault-tolerant computing. He is the TeraGrid GIG Director of Science, and a Senior Fellow in the Computation Institute (CI) at the University of Chicago and Argonne National Laboratory. He is also an affiliate faculty member at the Center for Computation and Technology (CCT), Louisiana State University (LSU), where he was previously Director for Cyberinfrastructure Development from 2006 to 2009, and is an Adjunct Associate Professor in the Department of Electrical and Computer Engineering at LSU.

SCOTT A. CALLAGHAN graduated with his B.S. and M.S. degrees in Computer Science in 2004 and 2007 from the University of Southern California (USC) in Los Angeles. He is a research programmer at the Southern California Earthquake Center (SCEC). His research interests include high performance computing, high throughput computing, scientific workflows, and seismic hazard analysis.

ROBERT HARKNESS obtained a B.Sc. degree in Astronomy and Physics from University College London and a D.Phil in Astrophysics from the University of Oxford. His research interests include computational astrophysics and cosmology, radiative transfer in supernova and classical nova explosions, massively parallel computer architectures, and programming techniques for supercomputing. He is one of the developers of the ENZO astrophysics code and specializes in adapting ENZO for extremely large scale applications in cosmology. He is currently a research programmer/analyst at SDSC and works with the UCSD Laboratory for Computational Astrophysics.

SHANTENU JHA is the Director for Cyberinfrastructure Development at the CCT, and a Research Professor in Computer Science at Louisiana State University (LSU). He is also a theme-leader at the e-Science Institute, Edinburgh and a Visiting Researcher at UC-London. His research interests lie at the triple point of Computer Science, Cyberinfrastructure Development and Computational Science. He is a co-PI for $2.6M NSF award that enables LSU/LONI participation as a TeraGrid Resource Provider. Shantenu leads the SAGA project and is currently working on co-writing a book on "Abstractions for Distributed Applications and Systems: A Computational Science Perspective" to be published by Wiley in 2010.

KRZYSZTOF KUROWSKI holds the PhD degree in Computer Science and he is leading now Applications Department at Poznan Supercomputing and Networking Center, Poland. He was involved in many EU-funded R&D projects in the areas of Information Technology and Grids over the last few years, including GridLab, inteliGrid, HPC-Europa, or QosCosGrid. He was a research visitor at University of Queensland, University of Wisconsin, University of Southern California, and CCT Louisiana University. His research activities are focused on the modeling of advanced applications, scheduling and resource management in HPC and networked environments. Results of his research efforts have been successfully presented at many international conferences and workshops.

STEVEN MANOS received his B. Sc (Hons) and PhD (Physics) degrees from the University of Sydney, Australia in 2002 and 2006 respectively. He is currently the Manager of Research Services at the University of Melbourne, Australia, where his role is to develop and deliver research infrastructure services throughout the University in the areas of high-performance computing, collaboration tools, large-scale data storage, data management and visualisation. Prior to this role he was a post-doctoral research fellow at University College London, London, UK, where he was involved in infrastructure developments in the EU-funded Virtual Physiological Human project and other initiatives involving computational models of physiological processes for patient-specific predictive medical applications. His interests include distributed computing, clinical applications of high-performance computing, the interface of art and scientific visualisation, and user education and training in eResearch.

DR. SUDHAKAR PAMIDIGHANTAM has been a consulting and research scientist for high performance computing and applications at NCSA for the last 14 years. He received his Ph. D. from University of Alabama at Birmingham after spending couple of preparatory year at IISc Bangalore, after Graduating with an M.Sc. from University of Hyderabad, India. Dr. Pamidighantam provides scientific consulting services for chemistry and computational biology communities at NCSA/TeraGrid and deployed Chemviz (chemviz.ncsa.uiuc.edu) the chemistry educational portal with integration of NCSA Condor resources. Currently he is leading the middleware integration for ParamChem virtual organization (www.paramchem.org) and continues to extend the services in Gridchem supported by US National Science Foundation. His interests are in multiscale molecular and materials modeling and simulation and cyberinfrastructure for science and engineering.

MARLON E. PIERCE is the Assistant Director for the Science Gateways Group in Research Technologies Applications at Indiana University. Pierce received his Ph.D. Florida State University (Physics) in 1998 in computational condensed matter physics. His current research and development work focuses on computational sciences with an emphasis on Grid computing and computational Web portals. Prior to forming the Science Gateway Group, Pierce served as assistant director for the Community Grids Laboratory at Indiana University's Pervasive Technologies Institute. Pierce supervises the research activities of software engineering staff and Ph.D. students, and serves as principal investigator on multiple federally-funded research projects. Pierce leads research efforts in the following areas: the application of service-oriented architectures and real-time streaming techniques to geographical information systems and sensor networks; the development of open source science Web portal software for accessing Grid computing and data resources; and Grid-based distributed computing applications in computational chemistry and material science, chemical informatics, and geophysics.

BETH PLALE is Director of the Data to Insight Center of Pervasive Technology Institute and Associate Professor of Computer Science in the School of Informatics and Computing at Indiana University Bloomington. Professor Plale did Postdoctoral study in the Center for Experimental Research and Computer Systems at Georgia Institute of Technology and has a Ph.D. in computer science from State University of New York Binghamton. Plale is an experimental computer scientist researching data related issues around cyber-infrastructure in an interdisciplinary research setting. Her research interests are in data provenance, metadata catalogs, automated digital curation, workflow systems in e-Science, and complex events processing. Plale is a recipient of the DOE Early Career award and is an ACM Senior Member and IEEE Member.

X. CAROL SONG received her Ph.D degrees in Computer Science from the University of Illinois at Urbana-Champaign, Illinois. Her research interests include scientific distributed computing, data interoperability, and advanced cyberinfastructure for scientific applications. Dr. Song is the Principal Investigator of the Purdue TeraGrid Resource Partner project, and a Senior Research Scientist at the Rosen Center for Advanced Computing, Purdue University, where she leads the group of Scientific Solutions.

JOHN TOWNS is Director of the Persistent Infrastructure Directorate at the National Center for Supercomputing Applications (NCSA) at the University of Illinois. He is also the Chair of the TeraGrid Forum, the leadership body for the TeraGrid Project. His background is in computational astrophysics utilizing a variety of computational architectures with a focus on application performance analysis. At NCSA, he provides leadership and direction in the support of an array of computational science and engineering research projects making use of advanced computing resources. He earned M.S. degrees in Physics and Astronomy from the University of Illinois and a B.S in Physics from the University of Missouri Rolla.