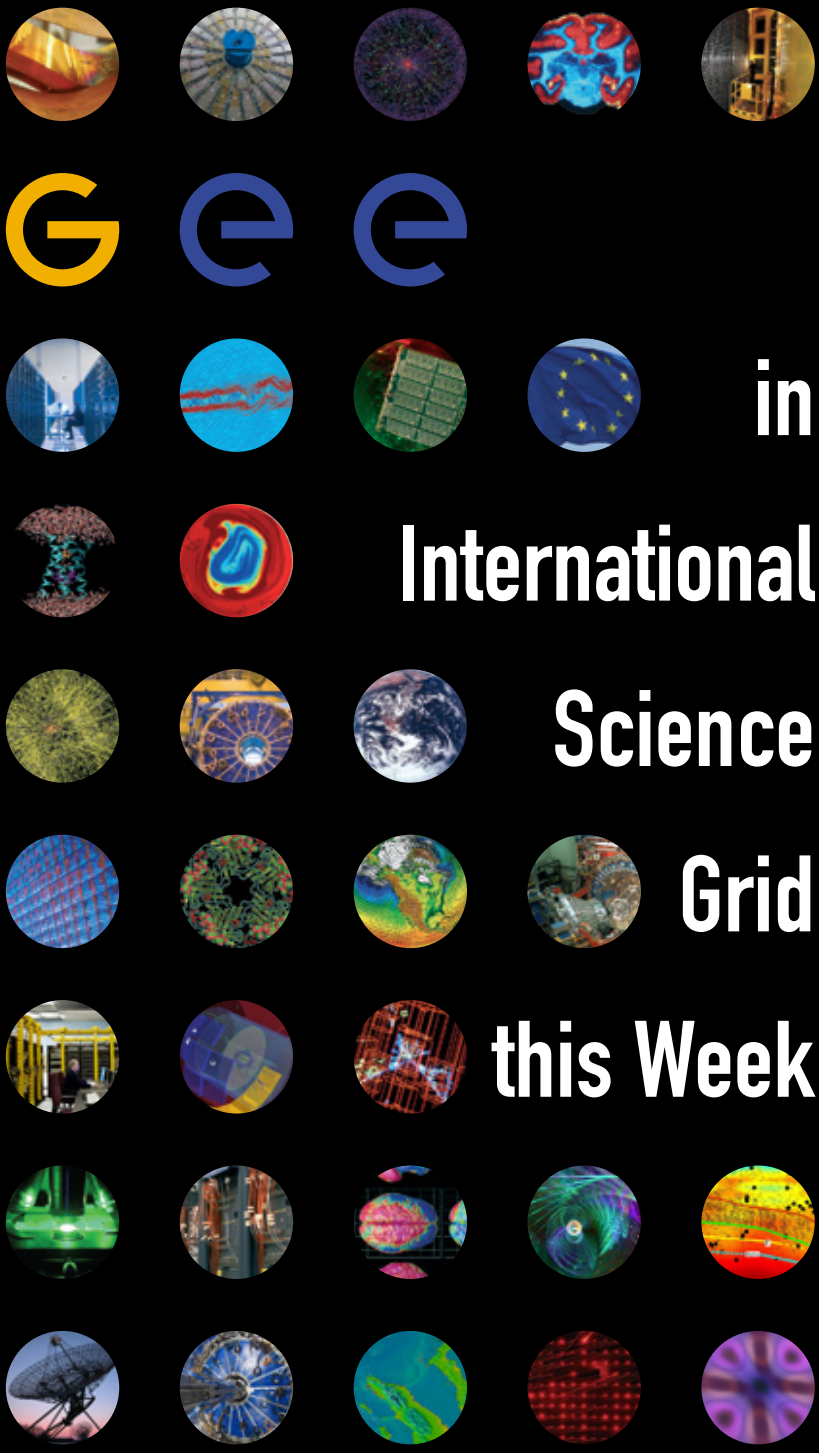


EGEE in International Science Grid this Week

Enabling Grids for E-sciencE

iSGTW

Table of contents

About this book

This book contains a selection of articles from *International Science Grid this Week* (iSGTW) that feature the Enabling Grids for E-sciencE (EGEE) project. These articles were published between May 2008 and April 2009. All the featured articles are available online at www.isgtw.org and within the *iSGTW* archives.

Acknowledgements
This publication was edited and designed by Danielle Venton, with thanks to Daniel Drollette, Anna Cook, Catherine Gater, Anne Heavey, Cristy Burne, Bob Jones and André-Pierre Olivier.

About *International Science Grid this Week*

iSGTW is a weekly newsletter promoting grid computing. It shares stories of grid-empowered research, scientific discoveries, and grid collaboration from around the world. *iSGTW* appears weekly, free of charge, in the e-mail inboxes of subscribers and is available via RSS.

First beginning in 2006, *iSGTW* is now read by 12,000 people every week and has 4,800 subscribers (as of the end of April 2009). Readers come from 200 countries and territories.

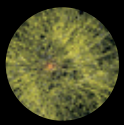
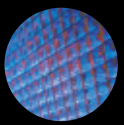
About Enabling Grids for E-sciencE

Enabling Grids for E-sciencE is Europe's leading grid computing project, providing a computing support infrastructure for over 15,000 researchers worldwide, from fields as diverse as high energy physics, earth and life sciences.

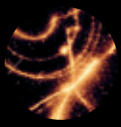
In 2009 EGEE is focused on transitioning to a sustainable operational model, while maintaining reliable services for its users. The resources currently coordinated by EGEE will be managed through the European Grid Initiative (EGI) as of 2010.

In EGI each country's grid infrastructure will be run by National Grid Initiatives. The adoption of this model will enable the next leap forward in research infrastructures to support collaborative scientific discoveries. EGI will ensure abundant, high-quality computing support for the European and global research community for many years to come.



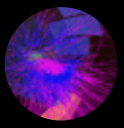



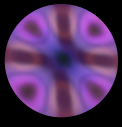


Astronomy and physics








No business like show business: final rehearsal for WLCG

Opening night . . . Like a theatrical company on the eve of a big new production, CERN's Computing Centre and its partner computing sites are preparing for their opening night.

This summer, the two stars of the show will mark their debuts: The Large Hadron Collider—largest particle accelerator built to date—will start up, and the Worldwide LHC Computing Grid (WLCG) will collect, move and process the massive amount of data the LHC generates.

In the next few months, final touches will be added to ensure that the WLCG and the people involved with it can put on their best performance, says Jamie Shiers, responsible for overall coordination of the system's practice runs—known as 'Common Computing Readiness Challenges.' To prepare, WLCG staff have run these "dress rehearsals" for years, which have become more intense as curtain time nears.

"These final two dress rehearsals are much more realistic," he says. "Before this, the testing was done in isolation; now it is everyone at the same time."

Rehearsal, rehearsal

During each run, the LHC Computing Grid receives data from four main LHC experiments simultaneously, at loads predicted for a fully operational machine. After a quick cache, the Tier-0 site at CERN sends data to Tier-1 sites, which send it to Tier-2 sites. Over 100

Before the show can go on, there must be extensive, full-scale dress rehearsals. For the Worldwide LHC Computing Grid, that means several Common Computing Readiness Challenges. Stock image courtesy of sxc.chu

auxiliary Tier-1 and Tier-2 sites around the world also aid in storage and processing.

The first full rehearsal was held in February; the final one will be in May. February's test yielded a middleware bug and a delay in storage configurations, but nothing was crippling, states Shiers.

"It was quite a success," he says. "We managed to get everyone to work together, all sites and all experiments, and we made it last. Before we started we honestly didn't know if we could support them all at the same time. We could have had to turn the volume down on the machine—and no one wanted to see that happen."

Shiers believes May's tests will probably reveal more glitches. For example, the grid may not be able to keep up with the scaling-up of computing volume. However, Shiers observes, the "LCG team" has grown adept at finding fixes or work-arounds.

Start-up itself will be a time of excitement and relief, Shiers says. "Switching on a machine like this doesn't happen very often," he says. "There is a sense that this whole process is like putting on a theatrical production—and we're the actors."

-

-Danielle Venton, EGEE

Preparing for opening night, left to right: Jan van Eldik, Miguel Coelho from CERN CASTOR operations team, Gavin McCance from the FTS operations team. Image courtesy of Jamie Shiers

(Almost) starting up the LHC: a view from the front lines

On Thursday, 8 August, CERN announced that it would make the first attempt to circulate a beam in the Large Hadron Collider (LHC) on 10 September, with the first test scheduled for the weekend of 9 August.

To find out what it is like from the point of view of those actually building the LHC and installing the equipment, *iSGTW* caught up with Mike Lamont of LHC Machine Operations for a few minutes, just before workers were about to start the initial steps of a preliminary, low-energy, "pre-startup."

iSGTW: What are your feelings, after all this time and effort?

Lamont: "At the moment, my general feeling is an overwhelming one of relief, actually. We've pulled all the bits together; we've got the ring more or less cooled now; we're underway."

"We're up to our necks with the detail—and the devil is in the details—but overall I'm feeling relieved."

iSGTW: If you had to take a guess, what would you expect the LHC to find?

Lamont: "Well, I'm just on the accelerator end, but I'm thinking the wish list is the Higgs, supersymmetry, and then the more exotic stuff." (Note: when *iSGTW* spoke with CERN theoretical physicist John Ellis a few months ago, Ellis said "We could find all kinds of weird and wonderful creatures we haven't even thought of yet.")

iSGTW: What do you see as the role of grid computing in all this?

Lamont: "There's going to be a huge, unimaginable amount of data coming out from this. I'm on the machine side, the operations side, but already I can see that. And even from my vantage point, you also get a real sense of the international aspect of the collaboration–you've got the Americans and the Russians and everyone else, sending data from Europe to Taiwan or China for processing. You really get a sense of the international level of the project."

iSGTW: Where do things stand right now?

Lamont: "This is as exciting as it gets. We've got the machine cooled, the access controlled and all the people out. We're going to start things today for a preliminary test tomorrow, and there's a pit in your stomach from the excitement."

(laughs) "This is as exciting as it gets."

Overview: what's involved

Starting up such a machine is not as simple as flipping a switch. Commissioning is a long process that starts with

the cooling down of each of the machine's eight sectors. This is followed by the electrical testing of the 1,600 superconducting magnets and their individual powering to nominal operating current. These steps are followed by the powering together of all the circuits of each sector, and then of the eight independent sectors in unison in order to operate as a single machine.

By the end of July, this work was approaching completion, with all eight sectors at their operating temperature of 1.9 degrees above absolute zero (-271°C). The next phase in the process is synchronization of the LHC with the Super Proton Synchrotron (SPS) accelerator, which forms the last link in the LHC's injector chain. Timing between the two machines has to be accurate to within a fraction of a nanosecond.

A welder works on the interconnection between two of the LHC's superconducting magnet systems. Image courtesy of CERN

After the first synchronization, or injection, test on the weekend of 9 August, for the clockwise-circulating LHC beam, the second test will follow over the coming weeks. Tests will continue into September to ensure that the entire machine is ready to accelerate and collide beams at an energy of 5 TeV per beam, the target energy for 2008.

Force majeure notwithstanding, the LHC will see its first circulating beam on 10 September at the injection energy of 450 GeV (0.45 TeV).

Once stable circulating beams have been established, they will be brought into collision, and the final step will be to commission the LHC's acceleration system to boost the energy to 5 TeV, taking particle physics research to a new frontier.

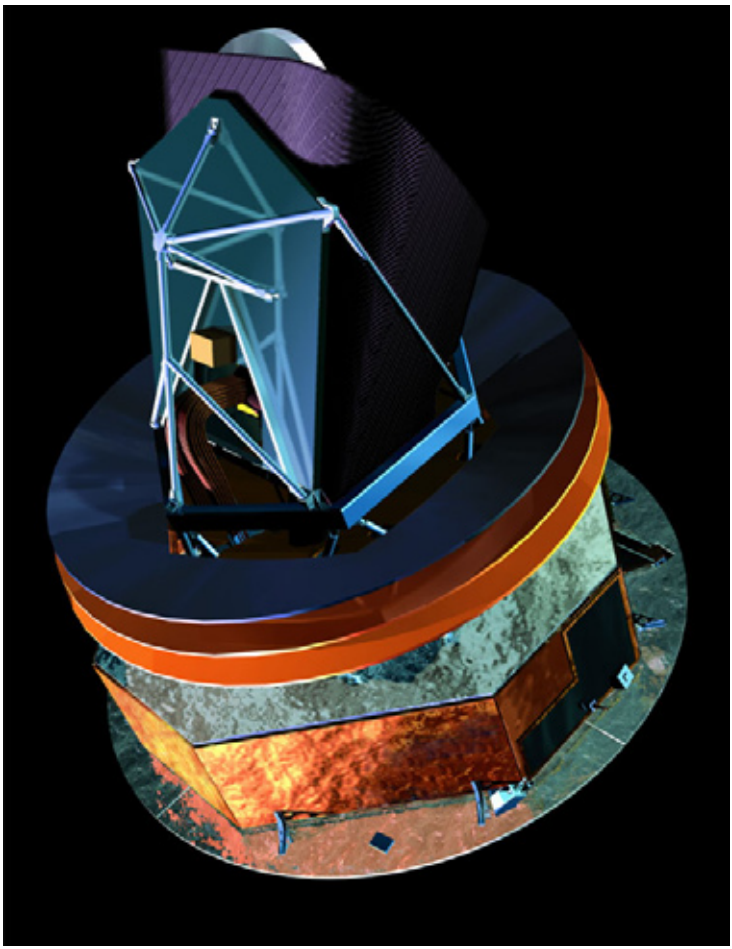
'We're finishing a marathon with a sprint,' said LHC project leader Lyn Evans. 'It's been a long haul, and we're all eager to get the LHC research programme underway.'

—Dan Drollette, iSGTW

World's most sophisticated thermometer to go into space

Next spring, the Planck satellite—arguably the world's largest thermometer—lifts off from the European Space Agency's launch pad in French Guiana.

A few hours after launch, Planck will detach from its escort and begin its six-month journey to its final orbit—a place 1.5 million kilometers away from the Earth, far enough away to avoid the emission of heat from the Earth, the moon and the sun.



An artist's impression of Planck. The Planck satellite, set to launch next spring, will collect information about the early universe. Image courtesty of ESA 2002. Illustration by Medialab.

While the Universe is cold, darn cold, at -270 Celsius, there are in fact slight temperature variations between one part of the sky and another. By taking sensitive temperature measurements of this background heat—detectable as microwaves—the Planck satellite will be recording relics from the Big Bang.

When our universe burst violently into existence, energy was initially trapped in plasma. Light was first able to escape this fireball about 300,000 years later. It

is this "first light," still around today as microwaves, that Planck will record—the oldest light anywhere in the universe.

It will give information about the infant state of our universe and measure more accurately the densities of ordinary matter, dark matter and dark energy (masses that seem to occur in puzzling proportions: respectively about 4, 21 and 75 percent).

"As the fruits of years of preparations are about to be sent into the sky 'excitement is mounting'."

From readings to discovery

The mission is expected to produce 17 terabytes of data during its 14-month run. Its data will come in a continuous stream and will require around 100 teraflops of computing power for storage and analysis. For part of this analysis, the Planck community is using grid technology. Claudio Vuerli of Istituto Nazionale di Astrofisica in Trieste, Italy, works with the grid team charged with managing the data processing one of Planck's two detecting instruments—the one specializing in lower frequencies traces.

"For the last few months we've been focusing on getting ready for the launch," says Vuerli, group leader of the astrophysics' Virtual Organization of Enabling Grids for E-sciencE.

"A mission of this scale and complexity requires a lot of ground-based testing and calibration during the pre-launch phase to fully understand the instruments and software behavior in preparation for real data-taking."

Many of Vuerli's colleagues have been working on Planck since its conception 12 years ago. As the fruits of years of preparations are about to be sent into the sky, Vuerli says, "excitement is mounting."

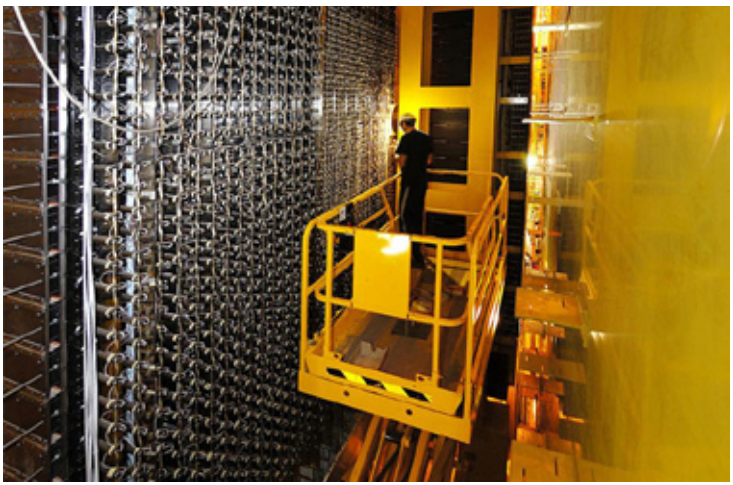
Other members of the EGEE Planck-Low Frequency Instrument Consortium work at the Instituto de Física de Cantabria, Spain. To learn more about cosmic microwave background and the Planck satellite mission visit the European Space Agency website and the Planck Operations homepage.

—Danielle Venton, EGEE

Beauty to unlock the mystery of asymmetry

Particles and their companion antiparticles differ only in the sign of their electric charge (positive or negative), and in other respects behave identically in most situations. Physicists were therefore stunned in 1964 to discover that some antiparticles behave differently from their particles.

An experiment at CERN called LHCb, for "Large Hadron Collider beauty" (where "beauty", also known as "bottom", is the second heaviest of the six known quarks), will study this difference in behavior, referred to as "asymmetry." The scientists expect to process 700 terabytes of data a year at several Enabling Grids for E-sciencE (EGEE) sites. In preparation, they've recently been running 10-20 thousand simulation jobs a day on the grid. The yearly processing power required for the simulations is equivalent to roughly 3,600 Dell Precision T7400 processors running continuously. Once data starts flowing, they'll need to add 50% more computing power.



The LHCb electomagnetic calorimeter, a device to detect electrons, their antiparticles (called positrons) and photons, particles of light. This immense array will collect immense amounts of data, requiring immense amounts of computing power and data storage. Image courtesy of LHCb.

Universe out of balance

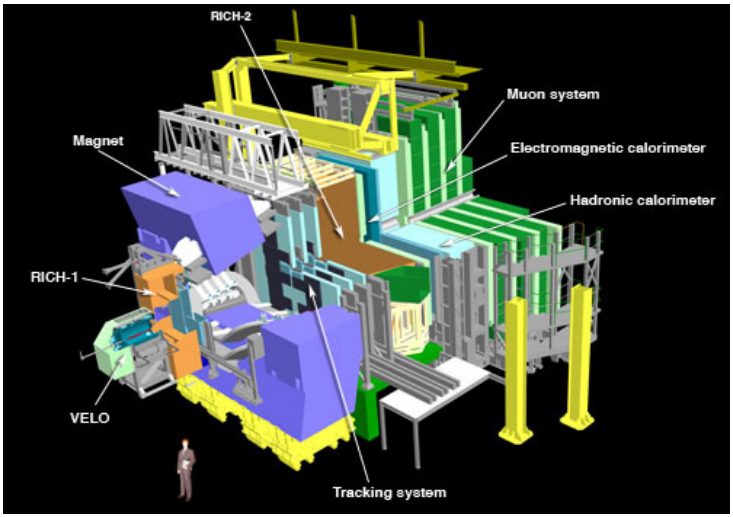
Scientists refer to the identical, or symmetric, behavior of particles and antiparticles as Charge Parity (CP) symmetry. The unexpected asymmetry that violates it is known as "CP violation". This asymmetry helps explain why the current Universe seems almost entirely made of particles with virtually no antiparticles. Although the Standard Model of particle physics does provide a source of CP violation—in weak interactions, for example, nuclear decay—it is not enough to explain the huge imbalance.

"We have a very strong indication that there are other fundamental interactions in the Universe not yet discovered

that provide additional CP violation," says Syracuse University physicist and LHCb collaborator Tomasz Skwarnicki. LHCb will look at rare processes in which heavy particles containing a beauty quark decay, or change, into lighter particles.

produce them at a much higher rate than previously possible, and will allow a much more detailed study.

"LHCb will reach a new level of sensitivity which will either lead to discovery of new forces in nature or provide new constraints on theories beyond our Standard Model," says Skwarnicki.



A schematic of the LHCb particle detector showing its components, one of which is the electromagnetic calorimeter. Image courtesy of LHCb.

Big is beautiful

LHCb chose the beauty quark because of its large mass. The heavier the quark, the better the chances of producing new types of fundamental interactions that are believed to hide at high energies, Skwarnicki explains. Although the top quark is heavier yet, it is much more difficult to produce and is less stable.

Beauty particles have been produced and studied for many years. However, the higher energy beam at the LHC will

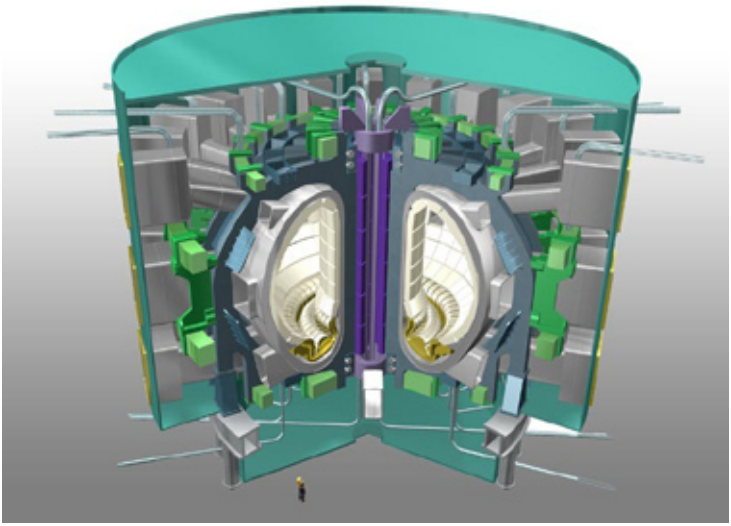
"In some channels we will quickly reach sensitivity beyond the previous experiments, and therefore may discover something new and important quickly. Grid computing is crucial for the LHCb experiment because of the amount of data to be processed and the amount of CPU needed for both processing real data and generating simulations."

—Amelia Williamson

Help create an earthbound sun

The dream of fusion power sounds so fantastic that one's initial reaction might be to dismiss it as science fiction.

Yet scientists hope to bring the power that emblazons the sun, fusion, to earthbound reactors. In this type of reaction two atomic nuclei bind—or fuse—together to form a heavier atom, triggering a monumental release of energy. The International Thermonuclear Experimental Reactor (ITER) is a joint international research and development project seeking to build a prototype fusion power plant. (The finished machine will be located in the south of France.)



A view of the International Thermonuclear Experimental Reactor (ITER), showing its main components, with a person for size reference. One of the pilot applications running in Ibercivis is devoted to ITER simulations. Image courtesy of ITER

Aiding ITER in its computational load is Ibercivis, a volunteer computing project centered in Spain, which allows computer users citizens to donate unused computing cycles from their personal computers to scientific research. While the Ibercivis resource broker and its storage elements are in Spain, project participants are from all over the world. Ibercivis launched in June 2008 and since then has been increasing its computing resources. The number of people donating their desktop resources for scientific computing has risen steadily, and is currently at about 20,000.

"Participating in Ibercivis gives people the opportunity to directly help in this scientific research," says Francisco Castejón, head of the Fusion Theory unit at the Research Center for Energy, Environment and Technology (CIEMAT),

Spain. “Ibercivis is accompanied by informative talks and material to help people understand what their computer cycles accomplish. This project is also interesting for researchers, giving them the chance to communicate their work to the general public.”

Map of the computing resources that are available for Ibercivis donated by participating citizens. Points in the sea correspond to cases where the location is unknown. The coloured lines correspond to the sent works and data. Image courtesy of ITER

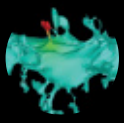
Ibercivis was developed with the cooperation of the Institute of Biocomputation and Physics of Complex Systems at the University of Zaragoza, CIEMAT, CETA-CIEMAT, CSIC and RedIris. In addition to hosting fusion applications, Ibercivis can be used for calculating other applications. For example, protein folding and material simulations also run in Ibercivis.

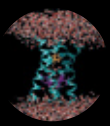
Ibercivis is not a temporary project (unlike many BOINC-based volunteer computing projects), so it will be possible to submit applications indefinitely. In addition, Ibercivis is designed to run not just one but several applications belonging to different disciplines. All applications previously ported to the grid can run efficiently in Ibercivis. This makes it possible to run a single application on both volunteer and grid resources.

For more information: The project known as EDGeS (enabling Desktop Grids for e-Science) brings together BOINC and EGEE, linking volunteer and production grids to support communities such as fusion. EDGeS and Ibercivis, currently working on a Memorandum of Understanding, hope to collaborate in the future.

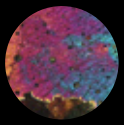
EUFORIA (EU Fusion fOR Iter Applications) is a pan-European project developing and supporting an IT infrastructure based on grids and high-performance computing for ITER applications.

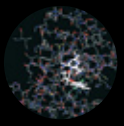
—Danielle Venton, EGEE



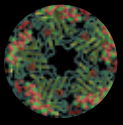
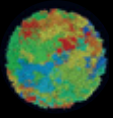
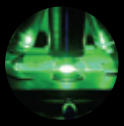
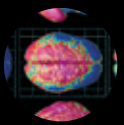


eGee

Life Sciences

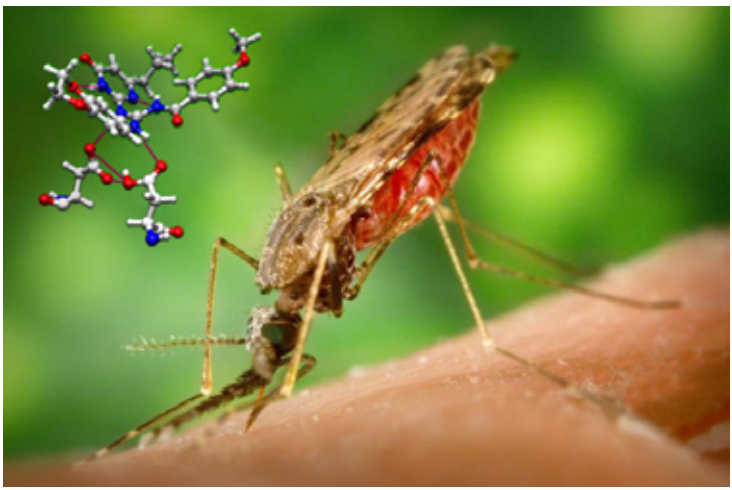
WISDOM unplugged: malaria drug-leads graduate to the wet lab

Hot on the heels of their initial success, WISDOM collaborators have high hopes as their anti-malaria drug-leads move from computer processers to petri dishes.

"We're excited to be at this stage," says Ana Lucia Da Costa, WISDOM researcher at CNRS, France. "I can't wait till we get the next set of results back."

"Disabling this protein essentially starves the parasites to death," says Da Costa.

Plasmepsin, Da Costa explains, is malaria's workhorse. Used by the parasite to attack red blood cells in the human body, it can be disarmed using the perfect weapon: a matching ligand; the lock for malaria's destructive key.



The WISDOM project screened one million molecules for their potential to bind to and inhibit "plasmepsin," a protein essential for survival of the malaria parasite. The parasite can be transferred to a human host via the mosquito vector. Image courtesy of PD-USGOV

The WISDOM project, a collaboration of eight core institutions in five countries, began searching for anti-malaria drugs in 2005. Rather than go straight to the lab, the team used a grid-powered software program to screen for potential drug-leads, searching for small molecules—called ligands—that could bind to and disable the malaria protein plasmepsin.

One in a million

The search for this perfectly matched ligand led the WISDOM team to perform 41 million "dockings," screening a million molecules and then discarding all but the 30 most promising molecules. This select handful then moved off-screen: undergoing in vitro evaluation in a wet lab at Chonnam National

University in South Korea. In a fantastic result, all 30 computer-selected ligands were able to inhibit plasmepsin in the lab, even at nanomolar concentrations.

"This was a great success," says Da Costa, who helped analyze the results. "We didn't expect all of the ligands to be active. It shows that our approach worked really well."

Now these 30 ligands have advanced to the next stage: in vivo testing in living cultures of the malaria parasite at the Commissariat Energie Atomique in France. The testing is ongoing, but initial results show at least one ligand inhibits the parasite life cycle at micromolar concentrations.

The next step will be to test for toxicity to animal cells. If the ligands are safe, the collaboration can begin to pursue new drug approval.

Classical pharmaceutical research often requires 15 years for a drug to be developed and approved. Thanks to grid-power, the WISDOM team has fast-forwarded this scenario and are already in the lab after just three years of research. They are using the same approach to expand into diabetes and avian flu research. WISDOM uses several grid infrastructures including Enabling Grids for E-sciencE.

Every year, malaria affects more than 500 million of the world's poorest people, killing around three million people a year. A child dies from malaria every 30 seconds; an African child has on average between 1.6 and 5.4 episodes of malaria fever each year. Image courtesy of hdptcar

- Danielle Venton, EGEE

In this case, it *is* brain surgery

To the edge: branching through your brain, a complex system of arteries, capillaries and veins feeds the organ that allows you to think. While the shape of each person's network is similar in the basics to everyone else's, it is unique in the particulars. This presents a challenge for neurosurgeons.

A new type of three-dimensional imaging and processing tool, based upon grid-computing, stands ready to change the way doctors perform neurosurgery. Called HemeLB, it won the Transformational Science Challenge Award at TeraGrid '08 in Las Vegas, US, last month.

New, grid-based imaging technology has the potential to change neurosurgery. Image courtesy of Phil Beard, sxc.hu

The current technology neurologists use allows for a patient's vasculature to be mapped by the processing of MRI scans. These scans, created with magnetic resonance imaging, use magnetic fields and radio waves to create pictures of bodily tissues.

Typically these are two-dimensional images: just a slice. Images in three dimensions are cutting-edge. Just beyond that 'edge' is a new type of imaging and processing that stands ready to change the way doctors perform neural surgery.

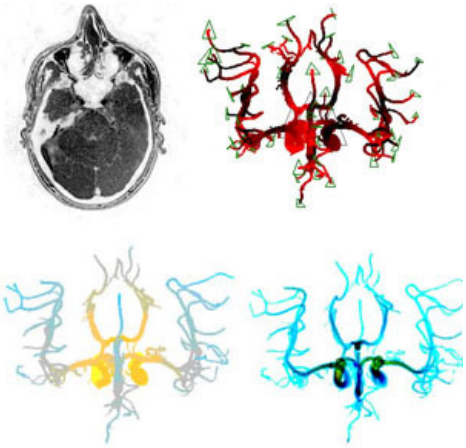
Funded by both the EPSRC in the UK and the NSF in the US, the HemeLB blood flow fluid solver is at the heart of a larger infrastructure project known as GENIUS, or Grid Enabled Neurosurgical Imaging Using Simulation, run by a group of British computational scientists and neurologists working out of University College London, University of Manchester and the National Hospital for Neurology and Neurosurgery, London.

Using HemeLB, it is now possible to visualize and predict changes in patient-specific cerebral blood flow.

"We are taking static MRI images and adding value to them. This will enhance the tools that doctors base their decisions on," says Peter Coveney, University College London.

Map-making

HemeLB processes MRI images, reconstructs a patient's vasculature and builds a map. This map shows blood vessels topography, blood velocity and pressure, and which vessels are likely stressed—possible sites of a future aneurysm. An area of abnormally low pressure is likely a site of blockage. The tool is also interactive and able to predict how these parameters change in the event of a surgical embolisation (a therapeutic procedure where an obstruction is placed in a blood vessel to redirect the flow).



Progress made with HemeLB, starting with (top left) a 2D MRI slice of a human head and then the images derived from it. Illustration courtesy of Peter Coveney, University College London

"When a surgeon blocks the site of a malformation in surgery, it is not possible to know how this will affect all regions of the brain," says Steven Manos, also of University College London. "What you would like to do is perform virtual surgery to see how pressure and blood flow are affected elsewhere. This will aid in surgical planning."

The tool is not in clinical use yet; it is still being tested and validated. However it may be launched in production quite soon.

"We would love to see it in use by the end of 2008," says Coveney. "The infrastructure to use HemeLB is in place: the tool is developed and we have the computing power."

—Danielle Venton, iSGTW

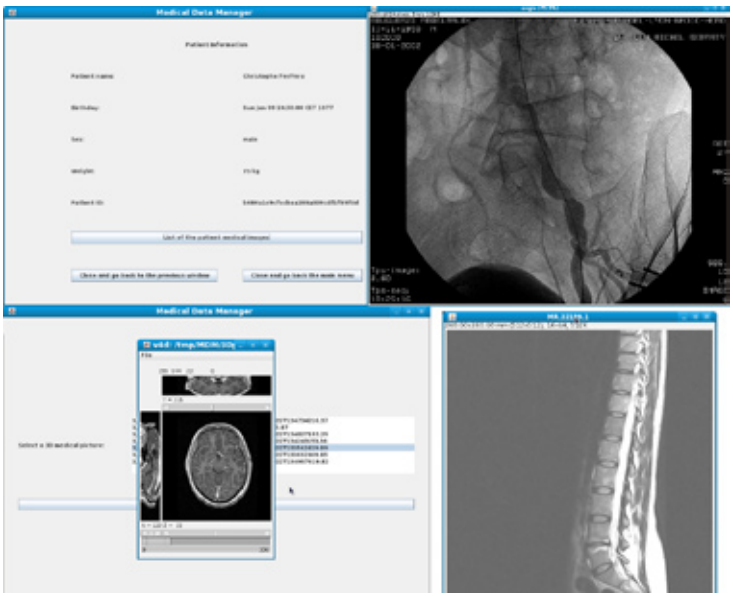
GENIUS uses grid power and, in the spirit of interoperability, is a trans-Atlantic project: computing resources come from both TeraGrid in the US and the National Grid Service in the UK.

Medical Data Manager

Doctors in Amsterdam may soon have more time in their busy schedules to spend with their patients.

A new way to access and manage patient data—the EGEE-developed Medical Data Manager, or MDM—was installed in June at the Amsterdam Medical Centre, Netherlands, in its first deployment for real use.

They hope to finish testing by the end of this year, and use the tool in production in 2009.



DICOM images in the Medical Data Manager interface. Image courtesy of MDM

In development since 2003, the MDM allows secure access to images in the DICOM format from grid resources. Standing for Digital Image and COmmunication in Medicine, DICOM is the prevailing world standard for computerized tomography (CT) scans, magnetic resonance images (MRI), positron-emission tomography (PET), X-rays and ultrasound images. MDM has several layers of access to comply with strict privacy requirements.

"This tool is expected to be a great boon for the doctors and researchers," said Johan Montagnat, computing researcher at CNRS. "They can now access data quickly—avoiding the lengthy and tedious process of manual anonymization, copying and transferring."

The data catalog they can now access has swelled—since they can now compare any relevant data stored within the system. And they can access data remotely—useful if they are in the habit of using computers at different sites or using grid resources to perform analysis of their patient data.

—Danielle Venton, EGEE

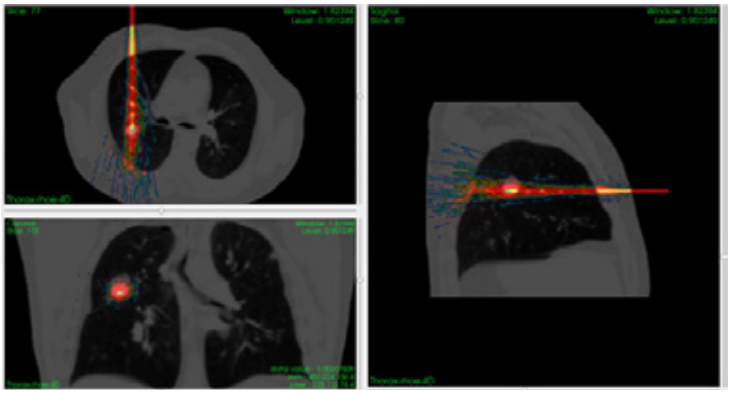
ThIS Cancer Therapy

ThIS is an acronym for Therapeutic Irradiation Simulator for cancer therapy, which simulates the irradiation of a patient with carbon ion beams in order to allow clinicians to compute the 3D dose distributions inside the patient for a computer-aided tomography (CT) scan.

Such simulations are very computing intensive, requiring thousands of hours on a single CPU. But by putting ThIS on the EGEE Grid, one simulation can be split into independent sub-simulations executed concurrently on different CPUs, speeding up the job.

The project is now being integrated into the OpenGATE project. The developers — Sorina Camarasu, Tristan Glatard, Laurent Guigues, Thibault Frisson and David Sarrut, of CNRS in Lyon, France — say "The results are promising and we are confident that this advanced submission method will help the simulator to reach a larger research and medical community. Moreover, since most users need a higher-level graphical user interface, we are currently interfacing the existing ThIS-DIANE architecture with tools like Moteur and the medical virtual laboratory of the Dutch VL-e project for a research usage and with the HOPE platform for clinical use."

—Daniel Drollette, iSGTW



Computing intensive CT scans are made a little easier to handle with the help of the grid. Image courtesy of CREATIS-LRMN & Léon Bérard cancer center

A FOOTPRINT keeps pesticides out of the water

Apple farmers in France's Vallée d'Auge — a part of Normandy renowned for its apple-based Calvados brandy — apply the insecticide Phosmet to their orchards to keep apple worm caterpillars from ruining their crops. But they don't want to pollute local water sources in the process.

Knowing how much to add is tricky. Too little, and the compound will stay localized and break down before it travels far enough to do much good. Too much, and the pesticide may contaminate surface water and groundwater. The EU estimates that already 40% of Europe's surface water has been affected by such runoff problems, as have the freshwater plankton, amphibians, fish and other organisms that dwell in it.

"A major difficulty so far has been the lack of tools allowing any pesticide user to know whether the application of pesticides may lead to a potential transfer to water resources," says Igor Dubus of the Bureau de Recherches Géologiques et Minières (BRGM), France's public geoscience institution.

An EU-funded consortium centered at BRGM, FOOTPRINT, has been seeking to develop computational tools for pesticide risk assessment and management throughout Europe.

In 2006, FOOTPRINT began running pesticide risk models for various locations, soil types, depths, slopes, and so on, in Europe. The team used the MACRO modeling software from Sweden and PRZM from the US, and collected the results in a data repository. They ran each of these models 50 million times in order to collect enough data. The models ran on computing resources

A Calvados producer may use as many as 100 different varieties of apple to produce their brandy. The apples used can be sweet (such as the Rouge Duret variety), tart (such as Rambault), or bitter (such as the Frequin, Saint Martin and Mettais varieties). Image courtesy of Sandor Fizli (www.sandor.ca)

from several projects including Enabling Grids for E-sciencE and FOOTPRINT@ work, a distributed modeling system that makes use of BRGM corporate computers at night, when machines are typically unused.

tural advisor, 'regional-scale' for the local authorities, and 'national- or continent-scale' for ministries of European Member States or EU regulators.

"This is a major leap forward: any per-

Apple worm caterpillars can swiftly ruin an apple crop. Image courtesy of the Agricultural Research Service, United States Department of Agriculture.

A reservoir . . . of data

When a user sends a query to the FOOTPRINT tools, complete with information about his crop, soil type, the intended pesticide and other relevant information, this reservoir of data provides the basis for the answer. The software drills down through vast datasets and returns information on the likelihood of the pesticide reaching surface waters or groundwater.

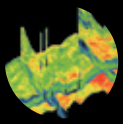
Three tools are available, each tailored for different scales and audiences: 'farm-scale' for the farmer or agricul-

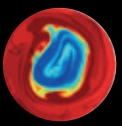
son, from decision-makers right through to farmers, will be able to do risk assessment for pesticides," Dubus says.

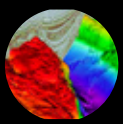
As of September 2009, FOOTPRINT will be made freely available to download. A dedicated team of FOOTPRINT scientists and technicians will provide long-term support.

"These tools will be a big help in keeping pesticides to minimal levels in watersheds," says Dubus.

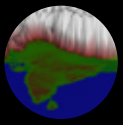
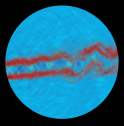
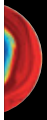
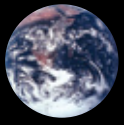
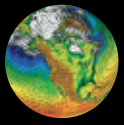
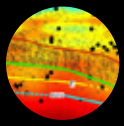
—Danielle Venton, EGEE

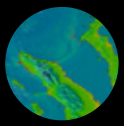






eGee



Earth Sciences

GRelC at Euro-Mediterranean Centre for Climate Change

Drowning in data: talk to any climate scientist, and they'll tell you that one of their biggest problems is dealing with the sheer volume of data they acquire when dealing with something as complex as worldwide climate change.

They don't lack for sources of data; if anything, they have too much to choose from—everything from tree ring data to ice cores to collections of air samples from hot-air balloons.

RealClimate is a commentary site on climate science, founded by working climate scientists—mostly from Columbia University's Lamont-Dohery Earth Observatory— written for the general public. Image courtesy of sxc.hu

With this in mind, managing the metadata can be a bottleneck. But distributed, peer-to-peer and grid-enabled solutions can help to provide a secure, decentralized, transparent and scalable solution for managing petabytes of datasets spread among several sites, say researchers at the EuroMediterranean Centre for Climate Change (CMCC, Italy). They have developed what they call the CMCC Grid Metadata Handling System (GMHS), which provides both access to and integration of climate metadata stored in different and widespread data sources, while allowing scalability, transparency and efficiency.

The middleware for their system is based upon GRelC DAIS, adopted as a data grid enabling technology, developed by the SPACI Consortium and the University of Salento under the Grid Relational Catalog (GRelC) Project.

The software, compatible both with gLite and Globus, is part of the IGI release, and was recently included in the EGEE Respect Program and presented at the last two EGEE events, at the EGEE '07 conference in Budapest and the 3rd EGEE User Forum in Clermont-Ferrand. It was mentioned as a runner-up during the related demo sessions—in both demos, CMCC researchers Sandro Fiore, Salvatore Vadacca and Alessandro Negro showed practical scenarios involving the distributed climate metadata management at CMCC.

Supporting science

"The CMCC Data Grid framework," says Giovanni Aloisio, Head of the Scientific Computation and Operation Division Head at CMCC, "will play a strategic and active role in the Euro-Mediterranean Centre for Climate Change infrastructure since it will support the scientific community in organizing and managing the huge amount of distributed climate datasets, and will provide enhanced metadata search and discovery services leveraging a transparent and scalable P2P/Grid approach. In addition, due to its distributed nature, it represents a perfect incubator for new data grid developments."

In the CMCC infrastructure, the GRelC DAIS will connect several distributed production centers, in Lecce (Tier-0), Bologna, Capua (Tier-1), Sassari, Venice, Milan, Sannio (Tier-2) through a P2P grid framework. It will allow CMCC climate scientists to perform distributed metadata search and discovery activities on the climate datasets that will be produced starting September 2008.

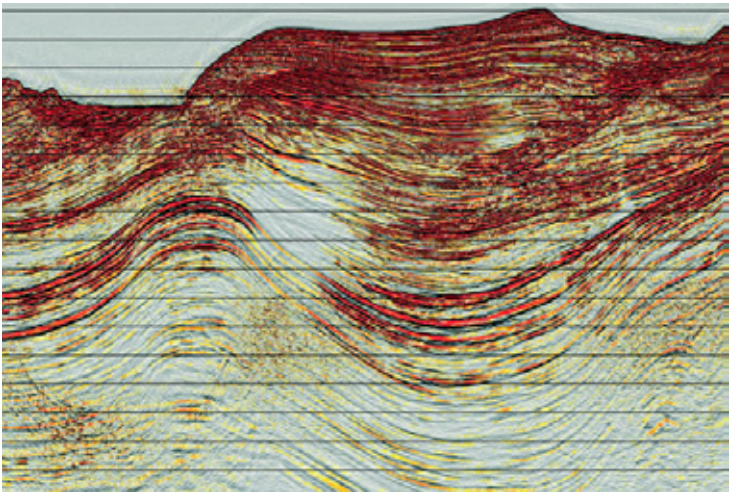
A distributed European testbed, exploiting GRelC DAIS, will be presented at the EGEE '08 conference in Istanbul in September during the poster session.

—Giovanni Aloisio and Sandro Fiore, Euro-Mediterranean Centre for Climate Change

Geocluster open for business

Last summer, CGGVeritas (Compagnie Générale de Géophysique-Veritas), the first business to develop an industrial application with Enabling Grids for E-sciencE, opened its doors to clients.

"Everything is set for our sales team," says Gaël Youinou, software manager at CGGVeritas. "I expect that the first contracts with clients will be signed before the end of the year."



Two-dimensional images can be hard to interpret. Image courtesy of CGGVeritas

The company developed software that processes seismic data, which comes from a variety of sources, including trucks that mechanically send sound-waves into the earth with the use of large, hydraulics devices, and ship-borne "airguns" that fire at regular intervals just below the water's surface as the vessel moves along pre-determined survey lines.

Regardless of the source of the sound-waves, or whether they take place on land or sea, the principle is the same. The reflected compression wave they create is detected by a network of sensors, forming a signal. It's much like radar.

And just like radar, the trick lies in the processing of the signal, distinguishing it from background noise, storing and interpreting the information, and putting it all into a format easily understood by humans.

From 2-D to 3-D

CGGVeritas' software, known as Geocluster, accomplishes this by creating three-dimensional underground maps which outline properties of the subsurface, along with the locations of oil

and gas reserves. A sister application, known as Reservoir Simulation, models how these reserves will evolve throughout the drilling process, enabling more efficient extraction. The data is then interpreted by a geophysicist who directs the exploratory company where to drill.

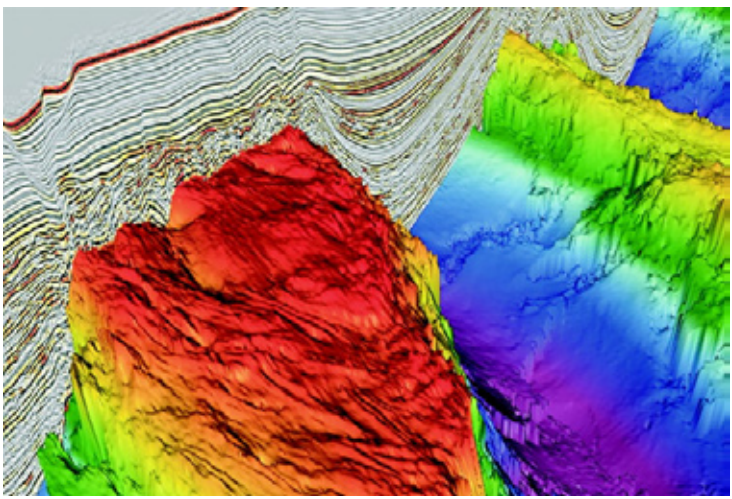
Both Geocluster and Reservoir Simulation were developed to operate in a grid environment using gLite, by the virtual organization EGEODE, using EGEE's infrastructure as a research environment and testbed. Members of the virtual organization come from both the business sector, such as Youinou, and from six academic laboratories.

businesses. They offer their service via a web portal

"This is a huge benefit to the client," says Youinou. "Beyond the ability to process data quicker, without managing new IT hardware and software, he can access data wherever and whenever and can easily collaborate with distant colleagues. In the end, their business will be more flexible and they will be able to reach the market more quickly."

The work is a partner of the BEinGRID project, which focuses on the benefits of grid technology for business.

—Danielle Venton, EGEE



Three-dimensional images can be much easier for humans to comprehend. Image courtesy of CGGVeritas

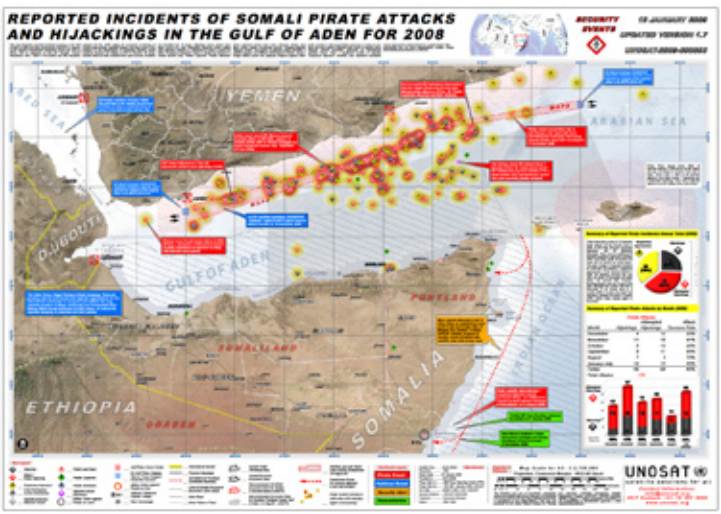
CGGVeritas will sell the application software as a package or a service. CGGVeritas has used EGEE's gLite middleware to develop a separate infrastructure for

Rough waters: fighting piracy with technology

In the past year, maritime shipping has suffered a resurgence of piracy, at a level that the world has not seen since the early 18th century.

Sailors working off the Horn of Africa have been particularly hard hit: last year, records show that 125 ships were attacked and 45 seized.

Real numbers are likely much higher, as piracy is believed to be widely under-reported. One of the world's busiest shipping lanes, about 20,000 ships annually pass through the Gulf of Aden on their way to and from the Suez Canal — carrying a tenth of world trade.



Map of reported pirate incidents in the Gulf of Aden, off the coast of Somalia. Image above courtesy of UNOSAT

Unlike the popular image of pirates seen in movies and books, modern pirates are more likely to wield machine guns than muskets; and the crime remains as difficult to fight as it ever did. Piracy is an extremely profitable and attractive occupation in a region characterized by lawlessness, bringing in multi-million dollar ransoms for the release of hostages, ships and cargoes.

Piracy has nearly cut off humanitarian aid deliveries to Somalia and has caused shipping insurance rates to skyrocket. Regional economies suffer as ships increasingly choose to go around the Cape of Good Hope. Given the number of oil tankers in the region, it seems only a matter of time before we see an environmental disaster of the Exxon-Valdez scale.

However, just as technology may have helped to promote the fall of the Robert Louis Stevenson-type of pirate, say historians (increased size and speed may have helped merchant vessels evade pursuing pirates), there is hope that technological advances will help protect cargos, vessels and crews.

UNOSAT aids monitoring, tracking and evading

Satellite-based maps produced by using grid technology are one promising anti-piracy tool. Different versions of these maps can tell the location of reported incidents and when they occurred, the identity and location of highjacked vessels, and the geographic areas with the highest density of attacks — accurate to within 100 meters. Some are offered in 3-D imagery.

Image courtesy of sxc.hu

UNOSAT, a co-operative project between the United Nations Institute for Training and Research (UNITAR) Operational Satellite Applications Program, and the European Organization of High Energy Physics (CERN), delivers satellite images to relief and development organizations. For the past five years, UNOSAT has worked on Somali-related security and humanitarian issues; it has monitored Somali pirate activity since last June as part of a UN Security Council resolution.

Typically, computer-intensive UNOSAT raw images are transferred to EGEE, where programs heavily compress satellite images for transmission over low-bandwidth connections, allowing users to access the latest maps from devices as simple as mobile phones. In this way, merchant ships, for example, can avoid areas where the latest assaults have been reported, and military vessels can know where to deploy.

Aside from pirate fighting, UNOSAT is involved with many other projects. Learn more at UNOSAT.org

—Danielle Venton, EGEE

Feature - Foreseeing floods

In autumn, the Cévennes mountain range in south-central France draws warm, humid lightning storms in from the Mediterranean. Depending on the cloud system, the storm can remain in the area, stuck churning in the mountains, dumping moisture for many hours.

In September, 2002, one such storm in the area, lasting 15 hours, caused 600mm of rain within one day. Unexpected flash floods caused the deaths of more than 20 people and economic damage estimated at 1.2 billion euros.

During the last decade, such floods have become significant natural hazards in Europe.

To help combat them, Vincent Thierion, a geo-informatics researcher, has been working on a grid-enabled flood forecasting application.

He said, "We met a few weeks ago with the civil protection service and they are very interested in this prototype, so we must work to finish it so they can use it!"

The ability to predict such situations is vital.

"There were no forecasting systems set up at that point (of the 2002 flood)," says Thierion. "Such events had been rare, and now they are more common, occurring also in 2005, non-catastrophically, and last fall 2008. You have no chance of knowing that one could happen until about one day before."

To help alleviate the problem, Thierion works on the European project CYCLOPS


Image courtesy of sxc.hu

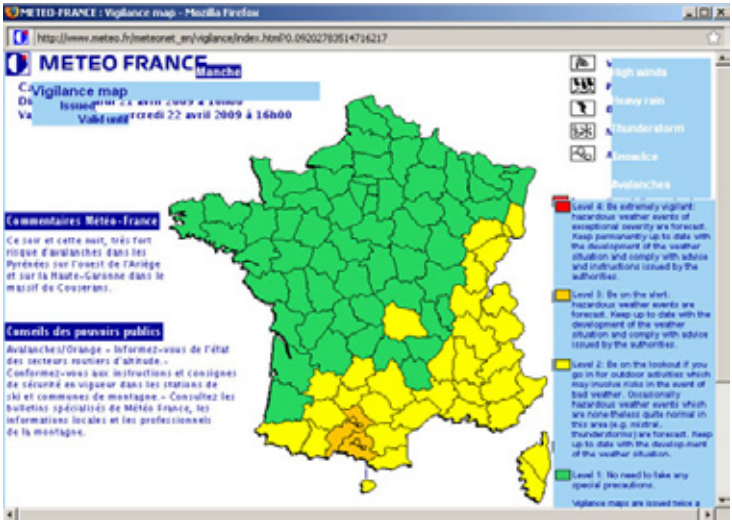
— Cyber-Infrastructure for CiviL protection Operative ProcedureS — which aims to bridge the gap between European civil protection agencies and the grid technology community. Supporting the work of the French Grand Delta flood forecasting service, CYCLOPS has developed a platform for running the forecasting application, G-ALTHAIR, on a distributed computing grid, and ported this to the Enabling Grids for E-sciencE infrastructure.

This style of computing allows the application to multi-task: researchers can run forecasts for several watersheds,

and for several rainfall scenarios at once, incorporating new data or adding additional layers of forecasting complexity in real-time. CYCLOPS researchers hope this will aid authorities in making decisions when managing crises.

possible scenarios using the grid.

"The existing system is mostly dedicated to monitoring a flash flood which is already occurring," Thierion says. "The innovation I think is important is to use



A typical forecast for France. Green represents good weather, yellow is possible rain, darker yellow (at bottom center) is Level 3 Hazardous Weather Alert. Image courtesy of Meteo France.

Advance warnings save property and lives

Currently in prototype phase, G-ALTHAIR will be in production by 2010, Thierion believes.

Using data from 170 regional monitoring stations spread around France's flood-prone zone, and radar images and rainfall forecasts from Meteo France weather service, the application gathers local data from regional computer clusters, which simulate the real-time hydrological situation, and forecasts the abilities of the grid to the prevent damage. With it we can share richer hydrological information with the civil protection services, to protect the population."

Thierion demonstrated G-ALTHAIR at the 4th Enabling Grids for E-sciencE User Forum in Italy in early March. Watch his demonstration through the GridCast video blog, learn more at GridCast.org.

—Danielle Venton, EGEE

EGEE

Editorial

Old friends and new: *iSGTW*, GridCafe, GridTalk and more

A new editorial team...

A big welcome to Dan Drollette and Anne Heavey, who will be producing all-new editions of *iSGTW* as of next week. Dan comes from a strong background in science journalism and Anne has been an *iSGTW* contributing editor since our launch. Both are keen to hear your grid-related news and announcements so please drop them a line anytime.

A new partner and project...

And welcome to EU project GridTalk. Co-funded by the European Commission, GridTalk was launched last week and aims to create a unified and cohesive approach to European grid communications. The project means the start of a new mission dedicated to deciphering grid policy, the addition of new contributions to CERN's acclaimed GridCafé, the creation of an interactive global GridGuide, and support for a full-time *iSGTW* editor.

A new funding setup...

The launch of GridTalk restructures *iSGTW's* funding such that we are now co-funded by GridTalk in Europe and Open Science Grid in the U.S., continuing the strong international nature of this collaboration. Many thanks to the EGEE project for their sponsorship over the last two years. EGEE, TeraGrid, OSG and GridPP will continue to stay involved in *iSGTW* through the provision of contributing editors.

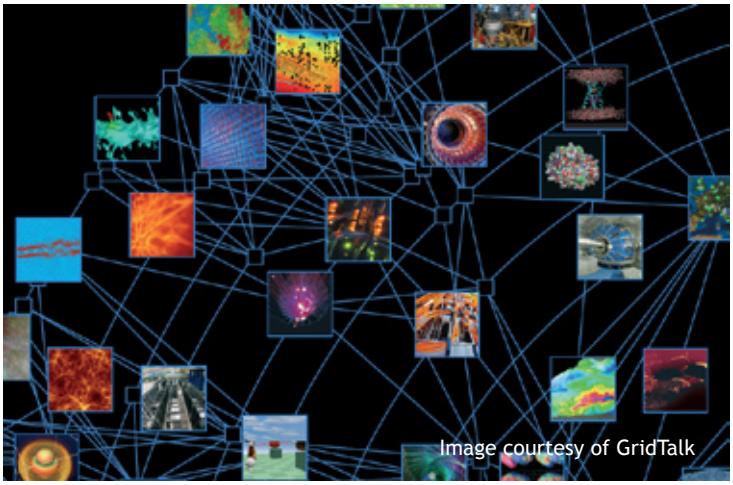
A new challenge for me...

And it's not quite goodbye for outgoing *iSGTW* editor Cristy Burne, who will be joining the GridTalk project to work on the creating of the GridGuide, an interactive map of the human face of grid computing. The GridGuide will make use of Imperial College's Real Time Monitor to showcase the people and personalities creating, using and benefiting from grid computing. Cristy will also be working to decipher grid policy, aiming to bring you the latest in grid policy updates as part of *iSGTW*. Watch this space for more information or contact Cristy with your ideas!

A little more on GridTalk...

Like grid computing itself, GridTalk combines the resources of several distributed partners: Queen Mary University of London in the UK, CERN in Switzerland, and APO, a web design company based in France. And like grid computing, GridTalk will reach out to distributed audiences, benefiting scientists and the public as well as those in government and industry.

"Ordinary people are benefiting from grid-enabled discoveries in health, chemistry, biology and more," says project manager Sarah Pearce. "GridTalk will show this human face of grid computing. We're going to put grid computing firmly in the spotlight, showcasing the behind-the-scenes computing that is making these new discoveries possible."



As well as outreach across Europe, GridTalk will co-ordinate activities with international partners in areas such as the U.S., Latin America and Asia.

Pearce says this coordinated approach is essential to the ongoing success of European grid computing efforts.

"The impact of grids has rapidly expanded beyond that which can be disseminated by individual groups," she says.

"Grid initiatives across Europe are contributing to new scientific results and changing the way science is done. The GridTalk team will co-ordinate the communication of these results, ensuring that grid-enabled scientific successes are reported widely in print and online."

A final word...

Many thanks to you all for making *iSGTW* the success it is! I have had a fabulous year as *iSGTW* editor and look forward to seeing many of you at grid-related conferences and to reading about your achievements in future issues of *iSGTW.* Thank you!

- Cristy Burne, iSGTW

Where are they now?

Updates from the three projects covered in iSGTW's *inaugural issue, 16 November 2006.*

Building the global grid

iSGTW led its first issue with a feature on the possibility of and progress toward achieving one seamless global grid. At the time of writing, basic interoperation between Open Science Grid and Enabling Grids for E-sciencE had been achieved, according to Laurence Field of EGEE.


Image courtesy of CERN

"Scientists using either infrastructure can now submit jobs to both and copy data between the infrastructures," Field said. "And if another grid interoperates with either, they'll see the other grid's resources. Through activities like these we hope to build up a homogenous grid landscape."

Where are they now?

Morris Riedel of the Jülich Supercomputing Centre in Germany, and chair of the Grid Interoperation Now (GIN) group within the Open Grid Forum, says that since then, members of GIN have tested emerging common open standard implementations deployed on production infrastructures to see if they can be used in real multi-grid scientific use cases.

Using lessons learned from this work, GIN has defined an interoperability reference model (IRM) that profiles mature standards in job and data management and security that can be deployed on production infrastructures to enable interoperability. This model represents a trimmed-down and more specific version of the Open Grid Services Architecture (OGSA), and thus represents one mid-term milestone towards full OGSA conformance as a long-term goal.

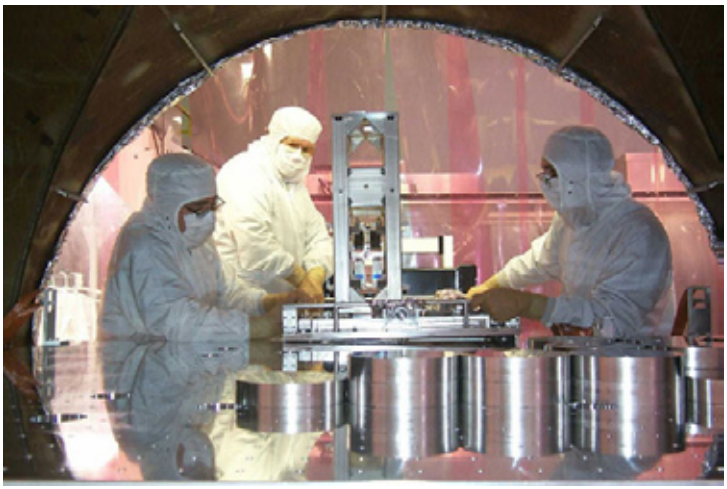
Exploring the Gravitational Wave (and Grid) Universe

Two years ago, the Laser Interferom-

eter Gravitational-Wave Observatory (LIGO) was on the computationally-intensive hunt for gravitational waves and the perfect grid. Theories predict one gravitational-wave event every 10 years. With the help of the LIGO Data Grid in partnership with the Open Science Grid, LIGO diligently searched the skies for these the raw signatures of a gravitational wave.

Where are they now?

While the collaboration installs the Enhanced LIGO hardware to enhance sensitivity for a new science run in 2009, LIGO scientists are analyzing 2007 "coincident" data from five geographically-separated detectors. Since May of this year, most days see over 1500 jobs on OSG, with a few days in the 5000's or above – a tiny fraction of what happens on Einstein@home; the recent average BOINC credit (a daily measure) is almost 13 million, second only to SETI@home!



LIGO staff installing a mode-matching mirror and suspension into a vacuum chamber during the construction of LIGO. Image courtesy of LIGO

Last May, *iSGTW* reported on LIGO's Einstein@Home volunteer computing project to attract CPU power to their search for a gravitational signature from spinning neutron stars. Einstein@HOME had reached one million BOINC credits a day, and had started running opportunistically on Open Science Grid resources, which led to significantly increased throughput.

LIGO expects to collect a year's worth of data during the new run and will look at about 10 times the volume of the universe. This will be a precursor to the next generation of "Advanced LIGO" instruments.

Where are they now? (continuted)

WISDOM vs. Malaria (round two)

In silico, in vitro, in vivo: WISDOM harnesses the power of grid computing to look for drug candidates. In 2005, the first data challenge targeted a specific enzyme from the life cycle of the malaria-causing parasite. In a second data challenge in 2006, with the support of EGEE and several related European grid projects, WISDOM tested up to 150,000 compounds per hour on up to 5,000 computers around the world.

In March 2007 *iSGTW* reported that WISDOM had tested 4.3 million potential malarial medicines during this second phase, and analyzed possible docking arrangements between drug compounds and target proteins of the malaria parasite.

In May of this year *iSGTW* reported that after having screened a million molecules *in silico*, the team selected 30 promising candidates to undergo evaluation in a wet lab. Almost all 30 successfully inhibited the enzyme, and hence the parasite.

Where are they now?

Ten of these 30 compounds have advanced to the next stage: testing in live parasite cultures for toxicity to parasite cells (while sparing human cells) and pharmacological potential. So far, these properties have not reached adequate levels in the selected compounds, so scientists are synthesizing related compounds, aiming to boost the levels.


Image courtesy of sxc.hu

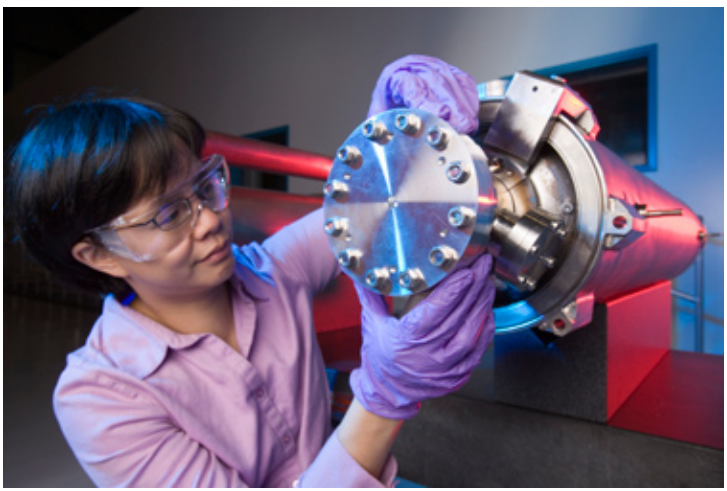
Meanwhile, two more malaria targets are moving to the laboratory for testing.

—Anne Heavey, iSGTW

Women in grid computing

In honor of International Women's Month, *iSGTW* looks at the role of women in computing, science and technology.

In a November 17, 2008 story in *The New York Times*, "What Has Driven Women Out of Computer Science?" Ellen Spertus, a graduate student at the Massachusetts Institute of Technology, tells of her experience at computer camp, in which she discovered that there were six boys to every girl. (And later, she found that only 20 percent of computer science undergraduates at M.I.T. were female.)



Engineer Mayling Wong examines an accelerator component at Fermilab. Image courtesy Fermilab Visual Media Services

The article says: "She published a 124-page paper, 'Why Are There So Few Female Computer Scientists?' that catalogued different cultural biases that discouraged girls and women from pursuing a career in the field," and notes that her paper was published in 1991. Computer science has changed since that time.

Today, there are even fewer women entering the field.

This gender imbalance is something we've noticed ourselves, when attending any grid computing conference. (Only 20% of the EGEE workforce — Europe's largest grid computing project — is female, for example.) And it's not just confined to computing; according to a study of National Science Foundation statistics by Joan Burrelli, only 7.8% of America's doctorates in physics went to women last year. On a similar note, only one percent of American college graduates are women who have studied engineering, said an editorial in *Progressive Engineer*.

Editorial: Women in grid computing (cont. . .)

Many speculate why this is the case in the fields of science, engineering and technology, especially in comparison to the proportion of women in other fields, which have seen steady increases over the past 40 years. "In 1973, only 6 percent of the Ph.D. scientists employed full time in academia, business or elsewhere were women; by 2006 the number had risen to 27 percent," wrote Natalie Angier in the 19 January 2009 Science Tuesday section of the *New York Times*.

One view—perhaps most famously attributed to then-president of Harvard, Lawrence Summers, in a 2006 address—is that the genetics and biology of gender determine the under-representation of women in science and technology. Meanwhile, others point to social factors.

The founders of the Smith College Picker Engineering Program in Northampton, Massachusetts, for example, note that American boys and girls show an equal interest and aptitude in science and technology up until adolescence.

In a newly released GridBriefing, "Facing the skills shortage: attracting more women to ICT," former *iSGTW* editor Cristy Burne notes a similar phenomenon among European 15-year-olds: Both sexes are equally interested in science as a career, but this interest does not carry through as they age — the percentage of female graduates in science and technology drops to as low as 20% in The Netherlands. (The percentage varies markedly from country to country, with 44% of science and technology graduates in the former Soviet states being female.)

"Shadowing day" in the IT department of CERN, part of EGEE's Gender Action Plan. Increasingly, there are more such programs being developed; to help keep track of them all, Fermilab established its Women in Physics site as a clearinghouse of information. Image courtesy of EGEE

What is needed

There are many possible reasons. According to *Women in ICT 2008*, a short list includes: a lack of female role models or mentors, a continuing imbalance in pay, a limited presence of women in decision-making positions, a lack of infrastructure and support that allows mothers to return to work, and unequal family responsibilities (40% of women aged 25-54 are out of the workforce due to family responsibilities, compared with only 4% of men).

In upcoming issues, *iSGTW* hopes to offer some light as to what may be happening, along with profiles of successful women in grid computing, and information on initiatives to counter the downward representation of women in the grid.

Progress will not happen overnight. As Judy Franz, Executive Officer of the American Physical Society, told *iSGTW*: "It will take a while for this effort to have its full effect — it's a bootstrap process (which used to work only for men), and these take multiple cycles to yield substantial change."

Getting more women into grid computing is more than a purely academic exercise: The European Union expects that the information and communication technologies, or ICT, field will be short by about 300,000 qualified staff by 2010, and more women must be drawn into the field to meet the needs of this growing industry. "It is unacceptable that Europe lacks qualified ICT staff. If this shortage of computer scientists and engineers is not addressed, it will eventually slow down European economic growth, and Europe runs the risks of falling behind," said Viviane Reding, EU Commissioner for Information Society and Media. "We need to overcome common stereotypes which describe ICT careers as boring and too technical for women."

And in the USA, in the face of what some have termed a coming "engineering shortage," industry has felt the need to encourage more women to get into science, engineering and technology. To cite just one example, industry gave Smith College a $73 million science and engineering center in order to ensure a supply of trained professionals for the future. Science and engineering are now the most popular majors on the campus, students told us.

If such outreach efforts succeed, women in science, engineering and technology — including computing and physics — could start to increase in number, make substantial scientific contributions, find their voices . . . and help enable the grid to grow.

—Dan Drollette, iSGTW

EGEE




Events



Image of the week - EGEE 2008 this week in Istanbul

It's tough to enjoy the ambiance of an Istanbul sunset via the Internet, but thanks to Grid-Talk's GridCast you're not too far removed from what's happening at EGEE 08...

Take some virtual tours:

* Neu Grid
* Grid CSIC
* Etics 2
* BalticGrid II
* EELA 2
* Belief II
* TR-Grid
* Health-e-Child

Read Catherine Gater's reports on the opening plenaries and thinking about grids in Asia.

Gillian Sinclair of the U.K.'s NGS will be "podded", as she calls it, on Wednesday.

"Yes, I'm being filmed for a podcast and I'm slightly nervous as this is my first one! Must remember not to use any odd Scottish words... "

Istanbul roundup

What has 545 people from 48 countries, 285 presentations, 97 sessions, 12 demonstrations, and 50 partners?

The EGEE conference in Istanbul.

"There was an excitement here," said Bob Jones, EGEE project manager. "It all came together."

Highlights, said Jones, included a Best Application Presentation award given to the CYCLOPS team of Marco Verlato (INFN), Stefano Dal Pra (INFN) and Valerio Angelini (CNR-IMAA) for their "G-RISICO: a Wild Fire Risk Assessment application running on an advanced Grid infrastructure." The team said that their approach could help civil protection authorities predict not only wildfire, but the risks of many types of natural disasters.


Image courtesy of Onno Sweers

Of EGEE and EDGeS

Another highlight was the signing of a Memorandum of Understanding between EGEE and EDGeS (Enabling Desktop Grids for E-Science). The first of its kind, "It will bring together the large, community-based approach of EGEE with the volunteer, BOINC-style desktop world," said Peter Kacsuk, director of EDGeS and president of the Hungarian Academy of Sciences. Such an approach could potentially aid researchers working in fields as diverse as fusion, biomedicine, engineering and chemistry, he added.

An outstanding range of posters covering topics from high energy physics to bioinformatics to middleware development was displayed, and the best poster competition was closely fought. This year, conference delegates voted for Javier Rojas Balderrama's "NEUROLOG: Neuroscience Application Workflows Execution on the EGEE Grid."

The demonstration competition winner was "EGEE Application Porting Support Group" presented by Gergely SIPOS. Judges said that they were particularly impressed by the speed with which this group has got up and running in just a matter of months since the start of EGEE-III, designing the application, porting it and now training other teams.

Best demo, runner-up was "Grid-Enabled Virtual Screening Service based on Grid Application Platform" from Mason Hsiung, Lee Hurng-Chun, Ueng Wei-Long, Chen Hsin-Yen, Eric Yen and Simon Lin.

—Dan Drollette, iSGTW and Catherine Gater, EGEE

Stop the presses!

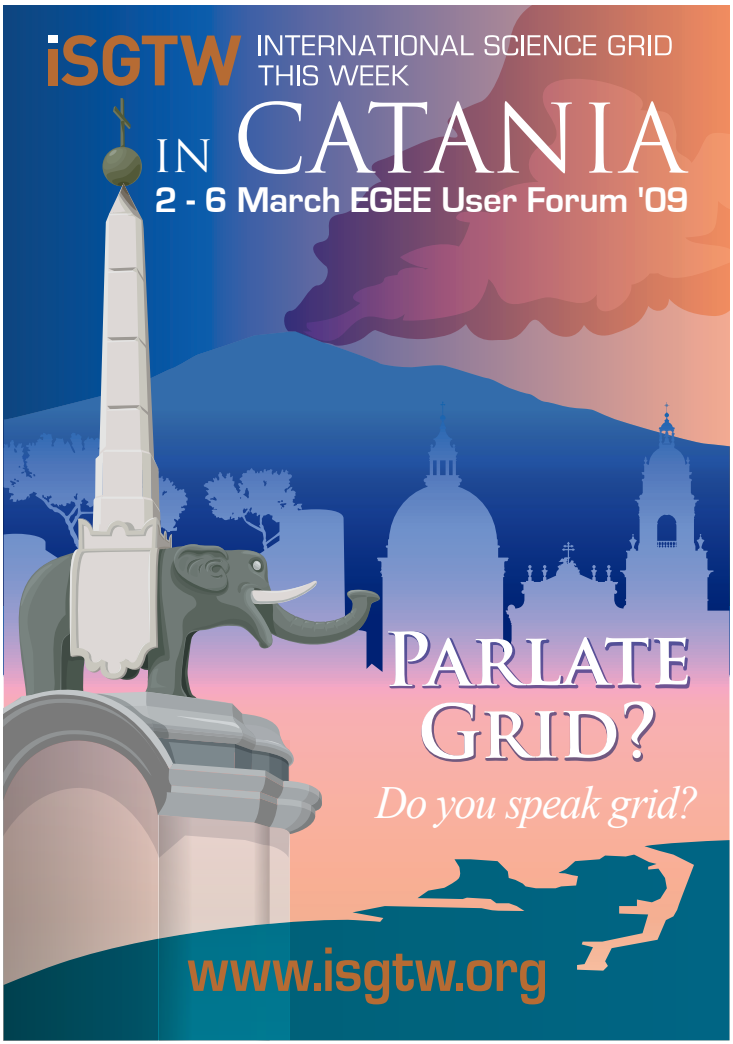
Here's a sampling of the latest headlines, blogs, twitters and "newsbites" on what has been happening so far this week at the 4th Enabling Grids for E-sciencE User Forum in Catania, Italy:

Amsterdam chosen to host EGI.org

Amsterdam was selected as the host city at the last EGI policy board meeting in Catania on Monday, 2 March 2009, ahead of seven other European cities.

"The choice of the location of the EGI.org headquarters is a further and decisive step towards the implementation of a sustainable European Grid Infrastructure," said Gaspar Barreira, chairman of the EGI Policy Board. "From now on we will be all mobilised for the real establishment of a new international research infrastructure in Europe, where a large number of countries will put together and operate the world's largest grid computing facility."

"We're very honored that the European grid community has chosen Amsterdam to host EGI.org," said Patrick Aerts, Director of the National Compute Facility



(NCF), the Netherlands, after the decision was announced.

"The decision on the EGI.org location is another major milestone for the establishment of the EGI," said Ludek Matyska, the EGI_DS Project Director.

Keeping up to date in Catania. Image courtesy of Gergely Sipos

"This decision comes just after the approval of the EGI Blueprint in Prague, on the 20th of January 2009," added Jürgen Knobloch, the EGI_DS Technical Director. "We can now proceed with the transition strategy from EGEE and other Grid infrastructures in Europe to EGI."

And in other breaking news . . .

Worldwide Grid helps in the fight against coronary artery disease (Press release)

Cardiogenics, an EU-project coordinated in Lübeck, Germany, aims to discover genetic variations leading to coronary artery disease, to uncover the underlying disease mechanisms and help to develop new treatments.

Read more on the GridCast blog (GridCast.org)!

Celebrating the development of the LHC grid

Last Friday, 3 October, grid-enthusiasts gathered—physically and virtually—from around the world for a special event: to celebrate the success of the Worldwide Large Hadron Collider Computing Grid.

"After the very successful start of the Large Hadron Collider on 10 September, it is now the turn of the Worldwide LHC Computing Grid to celebrate the transition from a development and deployment phase into operation," said Wolfgang von Rüden, head of CERN's Information Technology department, in his welcome address to the 250 grid enthusiasts gathered in CERN's 'globe' building.

Wolfgang von Rüden unveils the cermonial globe sculpture, before an audience in CERN's globe-shaped building. Image courtesy of GridFest

"I am very happy to tell you that we are ready to handle the unprecedented flood of data that will be generated by the LHC."

In his address "CERN and LHC: their place in global science" CERN's Director General Robert Aymar, highlighted the necessity of computing for studying particle physics: "This is just the first step in a long voyage. The LHC is a discovery machine, and its research programme stretches out two decades into the future," he said. "There are three needed tools to study physics. You need accelerators, detectors and computing. It is the computing that turns the flow of data into useful information."

Grids for science

While physics was the star scientific discipline of the day, attendees also witnessed the importance of a wide range of applications from a host of disciplines.

"The great thing about grids is that they have many uses," said Bob Jones, project director of Enabling Grids for E-sciencE, in his talk "The Grid Beyond Physics." He discussed applications in seismology, atmospheric research, astronomy, fusion and the life sciences.

On-site demonstrations, held through out the day, showed attendees some of these applications live. Demos included the Worldwide LHC Computing Grid, the ALICE experiment, the ATLAS experiment, the CMS experiment, the LHCb experiment, the Health-e-Child project (pediatrics), the ITER project (fusion energy), Open Science Grid and the WISDOM project (drug discovery).

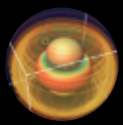
At the LHC GridFest, Bob Jones highlights the far-reaching uses of grid computing. Image courtesy of GridFest

If you were not able to attend in person, check out many of these demos online, posted by the GridTalk project on YouTube.com.

—Danielle Venton, EGEE

egee

Grid Collaboration



Securing the multi-platform grid

All locked up tight: one of the biggest challenges facing scientists who wish to make use of multi-platform grid infrastructures today is reconciling the different security systems inherent in the various platforms.

For the last two years, the Open Middleware Infrastructure Institute for Europe (OMII-Europe) has been developing a flexible framework for integrating the three dominant platforms in use in Europe: UNICORE, gLite and Globus. (OMII-Europe is a separate initiative from OMII-UK, which is often referred to as OMII for historical reasons.)

There's secure, and then there's making things really secure. OMII project member and photographer Sergio Andreozzi shot this image in Florence. He said: "Each lock is attached by a couple who just got married, as a symbol of a strong union." Image courtesy of Sergio Andreozzi, Istituto Nazionale di Fisica Nucleare (INFN), CNAF

A major part of this work has been "unpicking" the different security elements of these platforms and defining a common security profile. Such a profile would enable scientists to unlock the potential of multi-platform grid solutions in a secure environment for future research.

OMII-Europe has been able to achieve this breakthrough by means of a triangulation process that involved active coordination between the relevant security standards working groups of the Open Grid Forum and the Middleware Security Group, along with an embedded investigative process within challenging e-Science projects such as WISDOM. This was combined with a solution-driven investigation by experts from all three platforms working within OMII-Europe.

New users of UNICORE, gLite and Globus all have to go through a similar process. Most of the steps relate to becoming recognized by the X.509 structure; additional steps are middleware-specific and are needed to ensure that

the user is recognized by the middleware.

The main difference between the platforms—the barrier to secure interoperability—is that UNICORE, unlike the other two, does not use proxy certificates. Previously, this did not matter but scientists are increasingly demanding the benefits of multi-platform resources.

The key to the OMII-Europe solution is the harmonization of a core set of open standards to allow the secure transmission of user credentials among Web services on the different middleware systems.

- Steve Brewer, OMII-Europe deputy project manager

Some of the many members of the OMII-Europe team, taken at their face-to-face, all-hands meeting in Stockholm at the Kungliga Tekniska Hogskolan. Image courtesy of Sergio Andreozzi, Istituto Nazionale di Fisica Nucleare (INFN), CNAF

The EGEE-based WISDOM project is promoting the use of grid resources for drug-discovery against neglected or emerging diseases such as malaria. Virtual screening is a complex process that can benefit from both massively parallel computational tasks as well as brute force non-parallel number crunching. In order to separate the bioinformatics applications that require expensive supercomputing grids from those that can be executed more cheaply on distributed clusters, a common security profile is required.

Malaysia's grid grows in power

(Editor's note: We were curious about what people in other parts of the world have to say about the grid. Here's an excerpt from Malaysia's "Star Online.")

Malaysia's grid computing system, comprising several thousand CPUs at Mimos Bhd's Technology Park Malaysia headquarters, has been successfully linked to to a European Union-supported grid called Enabling Grids for E-sciencE (EGEE).

The Petronas Towers, Kuala Lumpur, Malaysia. Image courtesy of sxc.hu

Mimos, an applied research organization, said its KnowledgeGrid Malaysia gained a substantial increase in computing power from the link-up with EGEE, which is made up of several clusters located in Germany, Britain, Austria, Turkey, the Netherlands, Italy and other nations. It will allow Malaysian researchers to collaborate more efficiently and easily with their counterparts outside the country.

KnowledgeGrid went through a three-month vetting process with EGEE to assess its performance and security qualities recently. "It had to meet several requirements, such as data fidelity, before it got the green light to join the European network," said Ng Kwang Ming, Mimos senior manager for grid computing. "It is not a one-off process. EGEE will constantly reassess the performance quality of KnowledgeGrid to ensure that it is on par with the European grid," Ng said.

"It also had to achieve assessmentalgorithms precisely," he added.

Despite its power, KnowledgeGrid is under-utilized by Malaysian researchers and industries.

"There are very few users on the grid now, because many (businesses and industries) still don't understand how the KnowledgeGrid can help them," said Abdul Rahman Ahmad Dahlan, director of grid computing at Mimos.

To tell more organizations about the benefits of grid computing, Mimos organized its inaugural Grid Computing Conference (GCC), held in conjunction with the World Congress on Information Technology 2008 (WCIT 2008) in Kuala Lumpur.

Speakers from organizations such as Open Grid Forum, German D-Grid, Pacific Rim Applications, Thai National Grid Centre and EGEE made presentations at the GCC on May 15-16 at the Palace of the Golden Horses in Seri Kembangan.

A select audience

Attendance was by invitation only. Abdul Rahman said the GCC was also a good platform for local scientists and industry players to rub shoulders with their international counterparts and share experiences.

Mimos believes that harnessing the power of grid computing will make industries more agile and competitive, as well help keep research and development costs lower.

"Industries would be able to keep the bulk of their IT budgets for other aspects of product development because KnowledgeGrid would meet most of their computing needs," Abdul Rahman said. Mimos said the two-day conference attracted about 300 local and foreign delegates.

Worldwide appeal

For its part, WCIT is a bi-annual global ICT (information and communications technology) forum that brings together global leaders in business, government and academia.

Billed as the Olympics of the ICT industry, it hopes to encourage global economic and social development through the exchange of policies and ideas on technology.

For these and other projects, the grid world is about to become a much more connected place.

WCIT 2008 took place at the Kuala Lumpur Convention Centre from May 18 to 22. Image courtesy of KLCC

- Jo Timbuong, Star Online, Kuala Lumpur. Reprinted with permission of Star Publications (Malaysia)

Grid computing walks the standard line: thinking inside the box

"Standard" is often equated with "average" or "boring." How can you innovate or invent when you're bound by standards and regulations? How can you push the boundaries when you're stuck inside a box?

Yet how can you create something on a grand scale—something that can slot into place with other grand things—unless you create something interoperable. Something . . . standard.

In this special feature, we delve into this easily overlooked aspect of grid computing.

Staring down railroad tracks in Guthrie, Oklahoma, just before dusk. Image courtesy of Patrick Moore, sxc.hu

In 1850s Australia, budding railroad tycoons began laying train tracks across the continent. Each team of financiers, surveyers and civil engineers adopted their own preferred system, independent of the others.

The result?

Australia developed horribly incompatible train lines—some had 4 feet 8 1/2 inches between the parallel rails of steel, some had 2 feet between them, some had 3 feet 6 inches . . .
By 1917, you needed to change trains six times to get from Brisbane on the east coast to Perth on the west . . .

Will the same incompatibilities be true of data on a computing grid?

In early June, two important meetings of the grid community were held in Barcelona: the 23rd Open Grid Forum and the 5th e-Infrastructure Concertation Meeting. Both were dedicated to standardizing the grid. Debate was rife, but a strong message emerged—Europe, and the ICT world, cannot afford to repeat the incompatibilities of the early Australian train situation.

Standardizing grids: the current landscape

While "the Grid" in its idealized form is a single interconnected, interoperating computer farm, the reality of grid computing is very different. Instead of a single all-powerful Grid, there are many smaller grids, each customized to the specific needs of a user group.

With many projects involved, truly seamlesss interoperability can be a challenge. Image courtesy of NorduGrid and Vicky White

These different needs have led to different technical solutions: just as a toaster from the United States won't automatically work in a kitchen in Great Britain, grid solutions developed for one grid don't always work for another.

The challenge is clear: if grids are to be widely adopted—if they are to offer real solutions for industry and e-science—then they must be interoperable, which means the development of standardized, transferable technologies. Such technologies usually develop in one of two ways: de facto standards, like using Google for web searches, seem to develop themselves. Meanwhile, formal standards, like the meter or the kilogram, require consensus within a user community.

In the grid world, the Open Grid Forum is the largest group working towards standards adoption. The OGF provides a global opportunity for volunteers from all walks of grid computing life to contribute to developing new standards. The process sounds simple: first, a group works to develop best practices in a particular area, then, the group approaches OGF for endorsement of that work as a particular standard. Or, in reverse, an area of interest is first identified, and then an group is formed to work on a standards solution in that area. These processes may sound simple, but in practice, the path to achieving an accepted, implemented standard is long and dotted with potholes.

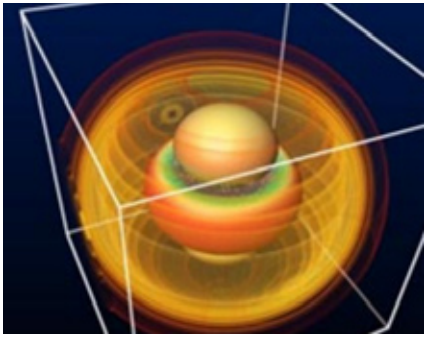
Challenges for the future

In addition to technical challenges, standardization can introduce issues such as different user requirements, incompatible policies and poor market timing. A classic example is that of videocassette standards, where Sony's earlier and ar-

Grid computing walks the standard line (cont. . .)

guably superior technology, Betamax, was outdone by VHS, a cheaper option that better served the rental movie market. (Both standards are now obsolete due to the rise of digital technologies. This highlights another difficulty facing grid computing: a rapidly changing marketplace makes it hard to pin down a strict standard.)

Despite these challenges, the benefits of standardization are very tangible. Standardization translates to interoperability, which encourages collaboration, competition and sustainability.



Physicists used the Globus Toolkit and MPICH-G2 to harness the power of multiple supercomputers to simulate the gravitational effects of black hole collisions. Image courtesy of Max Planck Institute for Gravitational Physics.

The implementation of a popular, functioning standard leads to smooth technology transfer, reliability and ease of use. In the medical industry, for example, adoption of the DICOM digital imaging standard enabled physicians anywhere in the world to interchangeably send, receive and store medical images.

The use of Latin (and the organizational system of Karl von Linnaeus) when naming biological species means all biologists speak the same taxonomical language.

And the advent of HTTP as a communications protocol has helped fuel the massive growth of the World Wide Web. Equally, the success of network standards such as Ethernet has not repressed healthy commercial competition in the network equipment market.

The way forward

No one can force a community to build middleware to a particular specification, or to adopt a particular security policy. Standardization relies on grid users and builders choosing to implement a solution that works most of the time, for most of the people.

At both OGF23 and the 5th e-Infrastructure Concertation Meeting, when the floor was opened for discussion, there were many questions yet to be answered: Who will pay to test for standards compliance? If we're not testing these standards, why bother to create them? And how can we ensure that we create standards that enforce best practice, when we're still learning what those best practices are?

Several things are certain: Grid comput-

ing requires standards at the industry level, with a validation framework that reinforces continued efforts towards software quality. Standards developers need to think long-term—beyond the lifespan of a typical project—to allow the time and energy required for a standard to mature. And standards developers need open corridors for communication, with different projects, different standards bodies and different user groups.

Wooden hinged ruler. Image courtesy of Jade Colley, sxc.hu

A standard in action: GridFTP

Grid computing provides the power needed to run data-intensive scientific applications such as drug discovery or high-energy physics. As part of this, massive amounts of data must be shunted around the world at high speed. Although there are many different ways of storing and partitioning such data, the grid community has agreed on just one way of transferring it: GridFTP.

Also known as the "grid file transfer protocol," GridFTP is the accepted method for securely and reliably transferring large volumes of data across distributed computing grids. It is based upon standard Internet FTP protocol but tailored to support the special needs of grid computing—including authentication and confidentiality features, reliability and fault tolerance, and third party and partial file transfer.

GridFTP was developed by the Globus project. The GridFTP Working Group, organized by the Open Grid Forum, continues to coordinate updates.

A "de facto" standard: VOMS

"Virtual organizations" are the human backbone of grid computing: groups of researchers from around the world who collaborate on common challenges, using grids to share and integrate their data and resources. The Virtual Organization Management System, or VOMS, is a system that allows distributed VOs to centrally manage the roles and authorizations of their members. Using VOMS, site administrators can generate local credentials for specific VO members, giving them with a single login and access to VO grid resources.

VOMS was developed by the DataGrid and EGEE projects. Although not an official standard, it is used by many VOs around the world.

—Cristy Burne, GridTalk

Bonus Feature: Readers talk back about standards

Find out what people from institutions as diverse as e-Bay and OGF have to say about the need for standards. (Followup to our special feature story this week on the roles of interoperability and standardization, by former* iSGTW *editor Cristy Burne.)

—Paul Strong, e-Bay

"For all of us, standards are a means to an end: interoperability that enables integration, collaboration, choice of vendor products/components and reduced costs. As scientists, government bodies or businesses, all of us have slightly differing priorities, but ultimately we all need the benefits that standards bring. The real challenge is how to deliver relevant standards in a timely fashion. My belief is that community-driven standards are the way forward and that standards driven through implementation are the most likely to be successful for those of us driven by quarterly and yearly results. We need interoperability and we need it fast!"

—Mario Campolargo, European Commission

"The e-infrastructures initiative of the European Commission delivers cutting-edge ICT-based infrastructures and services to solve real-world problems. Already today, there are more than 300 different organizations participating in the e-infrastructures initiative, including more than 80 scientific projects. Through their work, these researchers are helping to shape the necessary common network and service standards that will be the key to European competitiveness in the ICT domain. Such common standards are very important. They make sure that these new infrastructures and services will be broadly adopted in the future.

Although the e-Infrastructures initiative targets the scientific community initially, common standards will ensure that eventually this technology spreads to areas with very high societal impact, such as education, health-care and environmental monitoring."

—Craig Lee, OGF President

"The European Union is leading the way on effective grid adoption and sustainability through its policy of comprehensively supporting the entire adoption process. Projects such as the European Grid Initiative will also drive the development of policy on many levels, such as data access, resource consumption, and energy efficiency or green IT. Such usage

policies will have to be monitored and enforced through standardized tooling, which OGF working groups are already pursuing. OGF looks forward to partnering with the EU at all levels—from the development of commission policies to technical implementations—to make this vision a reality."

—Oxana Smirnova, NorduGrid

"Standards are the basis for innovation and progress in grid and are essential for grid technology take-up and exploitation. Over the last few years, NorduGrid has been involved in a large number of interoperability-related efforts. Condor, gLite, UNICORE, a number of lesser-known solutions—you name it, we tried it. The only way for users of other middlewares to use our services, and for us to reach out to such users, is to make sure our ARC middleware is interoperable with other major solutions."

—Dieter Kranzlmueller, European Grid Initiative Design Study

"It is clear that EGI on the one hand and OGF on the other hand need to collaborate on shaping the future European Grid Infrastructure. Mastering the landscape of standardization is a key element in the future sustainable European Grid Infrastructure. In fact, standardization is essential to the further distribution of grids as well as to interoperability between different grids, such as National Grid Infrastructures."

—Steve Brewer, OMII-Europe

"The OMII-Europe experience has shown that the adoption of standards can potentially enable scientists and other grid users to do new things that would not have been possible with resources tied to single infrastructures. The benefit of the project-based approach is that it brings together demanding users, standards-savvy developers and competing providers focused on achievable medium-term goals. There are no simple answers as business and research communities want different degrees of security and stability."

—Gaby Lenhart, European Telecommunications Standards Institute (ETSI)

"ETSI has a long history of part-

Bonus Feature: Readers talk back about standards

nership with the grid standards community, working with OGF to develop grid test cases and interoperability events in collaboration with all stakeholders. This year we will hold our 5th Grid Plugtests interoperability event, providing new opportunities for companies to test their prototypes against a standard with their partners and competitors. ETSI's mission is to succeed in involving the telecom operators and manufacturers in grid standardization. This is what is happening now in the ETSI Technical Committee GRID and we should soon have some echo of success from this committee's collaboration with OGF."

—Erwin Laure, EGEE Technical Director

"To build a seamless European and even worldwide grid infrastructure, common standards are key. While today's users have to adopt to the specific services offered by different grid infrastructure providers, common standards will soon help offering by them a seamless infrastructure. EGEE works closely with standardization bodies such OGF and other grid infrastructure projects to ensure standards are developed according to best practices originating from operational experiences."

—Ruth Pordes, Open Science Grid

"When common interfaces and best practices are adopted and begin to demonstrate value, standards follow.

In Open Science Grid we work with many different end-user communities to provide easy access to resources across local campus, national and international distributed infrastructures. Standards emerge from the agreed-upon services to meet the needs of these communities.

Agreeing on and documenting standards across a diverse set of projects is a painstaking process. I much appreciate the efforts within and facilitated by the OGF and similar standards bodies to which we contribute.

As we implement our model of federated, interoperating infrastructures in the OSG, we use, and in the process validate, such standards. This has of course been instrumental in our achieving interoperability with EGEE, Pragma and TeraGrid at the international and national levels. At the regional and campus levels we work with New York State Grid, Harvard, Clemson, Fermilab and the University of Wisconsin."

We must not be afraid of the future: EGEE to EGI

At last week's Enabling Grids for E-sciencE conference in Istanbul, the audience received a piece of sound advice during the welcome address from the European Commission, 'e-Science grids: Where does Europe stand?', delivered by Mário Campolargo: "We must not be afraid of the future."

Within the European Grid Initiative, planning is underway for the implementation of a sustainable, pan-European grid infrastructure to support collaborative eScience. The building blocks of this infrastructure are national grid initiatives—autonomous, national grid bodies, able to mobilize their own national funding and computing resources and to support their own research communities.

Time to turn off this road and transition to the next. Image courtesy of www.gerards.be

The projected transition of grids from the current short-term, project-based approach to a new long-term model will require significant reorganization for the grid community; EGEE will no longer be the central organizing body for science grids within Europe. Careful planning is required to ensure a smooth transition from today's EGEE-based infrastructure into the future EGI model.

"During the EGI sessions today I felt that while there are still issues that exist, there is a real willingness to move forward and, above all, a new sense of urgency," says Bob Jones, project director of EGEE. "This comes from the recognition that the end of EGEE-III and other collaborating projects is really just around the corner."

Agreement by the EGI Policy Board to begin accepting bids for the location of the EGI organization marks a significant step forward. The EGI.org is expected to start operation in 2010. "Between now and Christmas, we need to come up with our transition plan," says Jones. "This plan will outline how EGEE will move from where it is now to where it needs to be to fit in with the EGI planning."

EGI Design Study project director Ludek Matyska notes that the transition must ensure continuation of the present grids and that it should also encourage wider adoption.

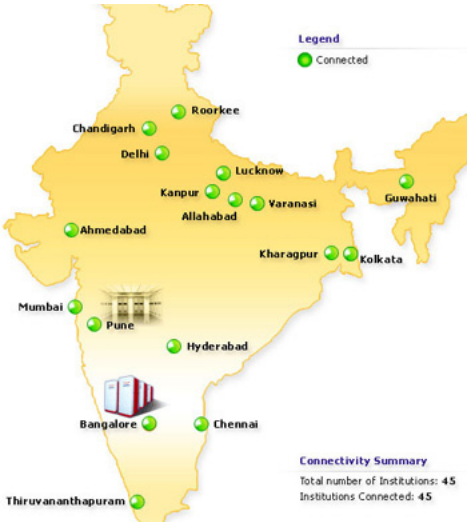
—Danielle Venton, EGEE

Prepared with the help of the GridTalk project.

India gets its "passport" to the grid

On the 5th of November, 2008, the Indian Grid Certification Authority (IGCA) was granted accreditation by the Asia Pacific Grid Policy Management Authority, also known as APgridPMA.

Indian researchers can now request user and host certificates to the Center for Development of Advanced Computing (C-DAC), which is located in Bangalore, and get access to worldwide grids. (C-DAC spearheads the Indian National Grid Initiative and is a project partner of EU-IndiaGrid, which joins European and Indian grids.)



GARUDA had linked 45 institutions in 17 cities via a high-speed network to promote science and computing in India. Image courtesy of GARUDA

Grid Certificates provide a secure key that allows researcher to access the grid, much the same as a passport allows you to enter a country. They are fundamental to guarantee worldwide grid access.

For this reason, supporting the creation of an internationally recognised National Grid Certification Authority represented one of the main goals of the EU-IndiaGrid project.

Since the very beginning C-DAC, in its role as coordinator of the Indian National Grid Initiative, has been identified as the candidate for the Indian Grid Certification Authority. The procedures to obtain the recognition by the International Authority are, for understandable reasons, extremely long and complex and therefore a temporary solution was worked out.

Thanks to the cooperation with the Academia Sinica in Taiwan, and the EGEE Regional Operation Center in Asia, the EU-IndiaGrid made a trusted certification authority available on a temporary basis right after the project start-up. This way access to worldwide grids for Indian researchers was immediately granted. However this clearly represented a temporary solution.

Today the establishment by C-DAC of the Indian Grid Certification Authority

and its recognition by the APgridPMA represents a major achievement not only for C-DAC and the EU-IndiaGrid project but also for the sustainable development of e-Infrastructures in India and their international integration.

In our opinion, the establishment of the IGCA has come at the right time, just a few months after the approval of GARUDA project phase two (foundation phase) and when the National Knowledge Network plan (NKN) has started its first steps. NKN is a recently approved plan for a multi-gigabit, low latency, e-Infrastructure and is a major step forward for Indian Research, as it represents the first step towards developing an ICT-based infrastructure capable of connecting major laboratories and research centres across India.

Hyderabad is the capital city of the Indian state of Andhra Pradesh and an emerging IT and bio-technology hub. Image courtesy of charminar, sxc.hu

—Alberto Masoni, EU-India

Get it all with GridGuide

Want to know what science is on the grid, who the scientists are and where they work? Help is at hand with a new website launched today. GridGuide is an innovative introduction to the sites — and sights — that contribute to global grid computing, a technology that connects computers from around the world to create a powerful, shared resource for tackling complex scientific problems. The launch of GridGuide comes as part of the 4th Enabling Grids for E-sciencE User Forum.

While still a work-in-progress, the GridGuide website already allows visitors to explore an interactive map of the world, visiting a sample of the thousands of scientific institutes involved in grid computing projects. Sites from 23 countries already appear on the GridGuide, offering insider snippets on everything from research goals and grid projects to the best place to eat lunch and the pros and cons of their jobs.

GridGuide's clickable map is an interactive guide. Image courtesy of GridGuide

GridGuide is an EC-funded project, and most of the sites included so far are European. But the GridGuide team are very keen to extend this — for example, they are working closely with OSG and TeraGrid to add more sites in the Americas, and anticipate adding more sites from Africa, Asia and Oceania.

"We're thrilled to see the GridGuide bringing grid sites to life," said Bob Jones, head of EGEE. "This site shows that people from all over the world are contributing to the success of grid computing."

GridGuide has been developed by the EU co-funded GridTalk project. Sarah Pearce, GridTalk project manager, says that sharing this human face of grid technology is essential to the continued success of e-Science.

"Grid computing is powering research into issues that affect us all: global warming, renewable energy, drug

discovery and more," said Pearce. "GridGuide tries to show what it's like to be part of that, encouraging each of us to be aware of and involved in the global research effort."

Using satellite imagery from NASA, the Real Time Monitor displays the Grid as it is geographically spread over the world. In near real time (3 minute delay) it shows running and scheduled jobs, job transfers and detailed information on resource brokers and computing elements for each site. Image copyright of GridPP's Real Time Monitor.

"Grid computing helps modern science facilities make an important contribution to large-scale international collaborations," said Mihai Petrovici, head of the Hadron Physics Department of the National Institute for Physics and Nuclear Engineering, Bucharest, one of the sites featured in the guide. "By being part of the GridGuide, we also help to substantiate the initiative for a 'grid of excellence' at the European level."

The GridGuide is non-profit and inclusion is free for institutes using grid computing for their research. The site aims to increase its global coverage over the next twelve months and is working with GridPP's Real Time Monitor to create an interactive 3D version of the site.

Other GridTalk projects include GridCafé, a website that acts as a beginners guide to grid computing, created in 2003 and recently relaunched to keep the public informed about advances in grid computing; *iSGTW* — a joint project between Open Science Grid in the US and GridTalk in Europe; and GridBriefings — jargon-free articles that provide timely summaries of policy-oriented issues in grid computing. The briefings target non-technical policymakers in government and industry, as well as scientists and the public.

—Cristy Burne, GridTalk

Anyone wishing to add a profile of their institution to GridGuide should contact gridtalk@gridtalk-project.eu for details.

Standards are the GLUE 2.0

In a major step forward, the Open Grid Forum, or OGF, announced on Tuesday, March 3 that they endorsed the GLUE 2.0 specification as the international grid standard.

The specification delivers the long-awaited common information model of grid entities. This document is a product of the international grid community, with contributions from the largest grid infrastructure projects and their middleware providers, such as EGEE, Open Science Grid, TeraGrid, NorduGrid, NAREGI and practical experiences from the science collaborations around the Large Hadron Collider (WLCG).

"During recent years, the grid community has been working very hard to reach convergence on how grid entities are modelled and described. The non-existence of a common information model has always been a major obstacle to interoperability. The release of the GLUE 2.0 specification as an OGF proposed standard is a major achievement of the grid community. As one of the founders of the GLUE working group, the Nordugrid Collaboration is naturally committed to adopting GLUE 2.0 through its ARC middleware," said Balazs Konya, NorduGrid Technical Coordinator and co-chair of the GLUE Working Group. "This will allow us to provide standards-based interoperability for our users with several grid infrastructures, including EGEE, the world's largest multidisciplinary grid."

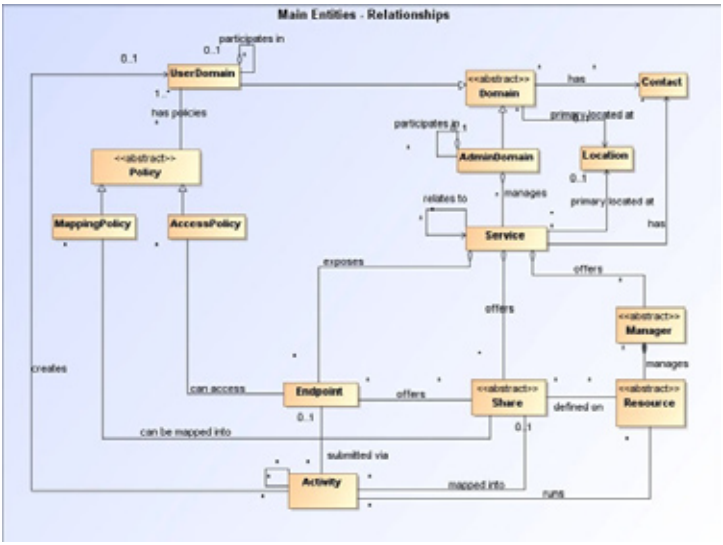
The pervasiveness of the Internet and World Wide Web in modern life would not be possible without the underlying standards that everyone understands and uses. For grids to reach this level of ubiquity, the global community needs to agree on and codify the protocols that should be used across the board. OGF is the largest body working towards these necessary standards within grids. Built from users, vendors and developers, OGF allows everyone associated with the technology to have their say on the standards adopted.

Image courtesy of sxc.hu

"The Information Service within any grid project is vital in allowing users and their applications to discover the existence of services and the capability of their underlying resources," said Steven Newhouse, EGEE's Technical Director.

The GLUE 2.0 specification was presented at OGF25 in Catania, Italy in the first week of March.

Over the week there were numerous sessions on defining, discussing and debating the standards which need to be in place for the future growth of grids. The sessions included lessons to be learned from the high throughput computing community, the benefits of grid standards to other computing infrastructures like volunteer projects based on BOINC, and discussions on which path to take in various aspects of the grid technology from workload management to metadata.



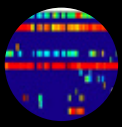
A conceptual model of the relationships among the main entities in a grid environment, showing resources, activities, policies and domains, among other items. Image courtesy of GLUE 2.0

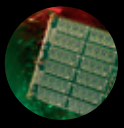
"The GLUE 2.0 specification represents the most recent phase of sustained effort between the major grid infrastructures around the world that has taken place over several years to build interoperability between our systems. EGEE will be adopting and deploying the GLUE 2.0 specification within our Information Service over the next year to provide an interoperable base that our user community can use to exploit services from our own and our collaborating infrastructures."

For more information on grids and standards, download GridTalk's GridBriefing.

—Dan Drollette, iSGTW

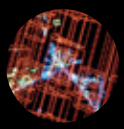
EGEE



Grid Technology

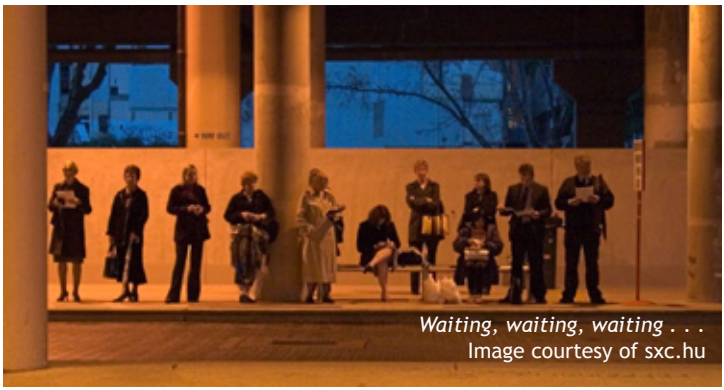




Let's be fair: a new approach for grid scheduling

No one likes waiting in line but—as long as no one cuts in front—we're usually happy to wait our turn. However, that changes as soon as there's the slightest hint we are being treated unfairly. In the research world, there are reports of individuals ditching their otherwise sunny disposition when they feel their jobs have queued longer than they should have.

Emmanuel Medernach noticed this among his colleagues at the LPC laboratory at the University Blaise Pascal in Clermont-Ferrand, France, where he is a member of the lab's computing management.



Waiting, waiting, waiting . . .
Image courtesy of sxc.hu

Medernach is now developing a new approach to grid scheduling to fix the problem. His approach has caught attention of many—including the EGEE awards committee, who gave Medernach first place in this spring's User Forum Poster Competition.

"Something has to be done to avoid inducing starvation of user computation." says Medernach. "In a queue we all expect the same treatment—we wanted the same thing for grid scheduling."

Although the lab's previous scheduling program was designed for efficiency, with the goal of maximizing mean throughput per job, the program focused on the jobs as a total—not on the users behind the jobs.

That solution might be good for the average person, but it could be bad for others. For example, a schedule that computes small, short jobs first will irk the researcher waiting for the solution to a big job. A schedule that gives priority to long jobs may never get around to the calculations of someone with many short jobs.

"We began by stating the properties we wanted our scheduling to have: impartiality and equity," says Medernach.

Impartiality means that scheduling is based on unbiased decisions. Equity means that, in allocation, the priority should be given to the neediest individ-

uals. Medernach's approach to scheduling is inspired by the writings of John Rawls, the American political philosopher who penned *A Theory of Justice.* This work attempts to marry the principles of liberty and equality, and present a way to distribute resources justly and fairly.

Emmanuel Medernach is working on fair grid scheduling for his PhD thesis at University Blaise Pascal in Clermont-Ferrand, France. To learn about his work view his prize winning poster. Photo courtesy of Emmanuel Medernach.

What is "fairness," really?

Medernach and his PhD advisor, Eric Sanlaville, from the LIMOS laboratory at University Blaise Pascal, found that fairness involves a multi-dimensional assessment of each user criterion. Then the only way to design an order based on impartiality and equity is known as a Leximin order, which ranks vectors by examining first the least well-off user, then the one who is just above, and so on. This approach, well established in the realm of economics, is novel to grid scheduling.

Medernach and Sanlaville are currently studying how to theoretically schedule simple job batches. Once they master this they will begin work on scheduling more complicated batches. In time they want to see this scheduling approach become real working software. To help with the development and testing of scheduling models, they are hoping to attract the attention of several grid projects interested in forming partnerships.

—Danielle Venton, EGEE

Guardian angels at the ready

Life is a complicated, tricky business and we sometimes need a little guidance. In the same way, novice grid users adding their applications to the grid—also a complicated, tricky business—often benefit from a helping hand.

"We are here to be the equivalent of 'guardian angels' for the grid," says Gergely Sipos, porting expert at the Hungarian Academy of Sciences in Budapest. Sipos is the team leader of a new group within the Enabling Grids for E-sciencE project. "We want to simplify a process that—in some instances—can be quite difficult."

Image courtesy of sxc.hu

The Application Porting Support group is designed specifically to make it easy for new or advanced users to get their applications on the grid. Experts from the porting team work closely with application owners to understand their requirements and to identify suitable approaches, tools for the porting process and to set realistic and feasible porting scenarios. The group started their activities this last May as part of EGEE's efforts to welcome and attract new users.

"Grid application programmers usually feel lost in the jungle of manuals, tutorials and presentations that are available on the Web," says Sipos. "We organize this knowledge according to our clients' needs and present it to them in a well structured way. We cut down their development effort and increase their efficiency."

Intensive workshops and personalized training events organized by the team

ensure that application owners become expert on porting tools that are relevant for them and can quickly and efficiently perform the porting process with the help of the porting team.

"It is very rewarding for me," Sipos says, "because I can see that real end users benefit from the grid and can achieve scientific results that could not be achieved without grid computing—and probably without our support."

Want to sign up for your own grid angel? Get involved

Gergely Sipos, left, working with a client. Image courtesy of sztaki.hu.

The Application Porting Support welcomes any academic or industrial application developer with demands for large computational and data storage resources, a platform for distributed applications.

At the webpage of the EGEE application porting team you can find further information about current applications, past success stories and how to apply. You can meet the members of the support team at the EGEE'08 conference, where they will exhibit their results in the "Application Porting Support Group" demo booth and will give presentations in the parallel sessions.

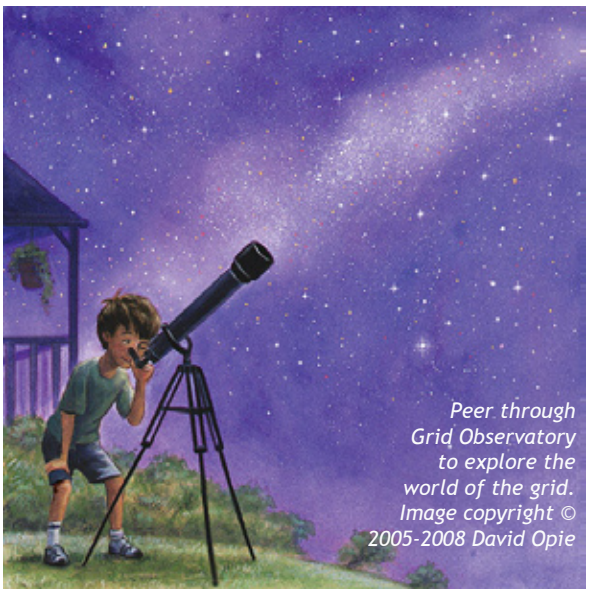
—Danielle Venton, EGEE

Observing the grid

What was once lost has now been found—and stored.

Thanks to the Grid Observatory, people in the esoteric field of studying the behavior of large, distributed systems have been given a gift: a trove of data.

Like astronomers who peer through a telescope to explore the solar system, researchers in grids and complex systems are able to examine Grid Observatory's data repository to find new patterns in the global behavior of the grid.



Peer through Grid Observatory to explore the world of the grid. Image copyright © 2005-2008 David Opie

The Grid Observatory, a sub-set of Enabling Grids for E-sciencE's applications group, opened its doors—via a Web portal—this autumn. Through the portal, researchers can access anonymous grid traces, collected and stored through the Grid Observatory application, and information about those traces, for an overall picture of the grid.

Grid traces, explains Cécile Germain of Grid Observatory, are signatures of grid usage. These take many forms: for example, which sites are up, how long a job takes to complete or which scheduling model is being used.

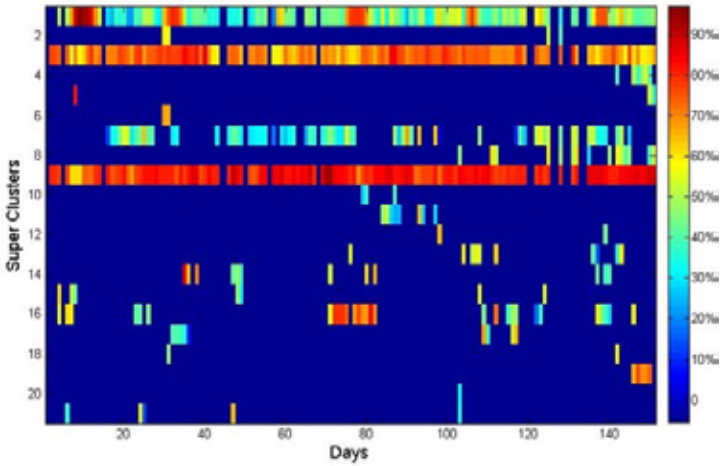
"This information was always collected through EGEE's monitoring tools," says Germain, "now it is stored, archived and can be used in study."

Not so esoteric after all

While this field of research is interesting in its own right and is related to studies of other complex systems (such as ant colonies, the economy, nervous systems and the climate) the group also expects that in time it will yield practical benefits as well.

"Before improving a system, you must observe it," says Germain, "then you can strategically place your adjustments. I hope that in time the Grid Observatory will help the grid be more scalable and reliable."

Germain believes that the Grid Observatory will help turn the grid into a complex adaptive system: one that is able to organize itself, change and learn from experience.



A monitoring tool used by Grid Observatory to visually describe the success of job clusters over several months. Image courtesy of Grid Observatory

"This goal of autonomic computing (one able to self-manage) is highly relevant at this time," says Germain, "where production grids are moving to sustainable infrastructures, are experiencing increased usage, and reducing the manpower dedicated to daily operations."

The Grid Observatory is an open project, keen to work with computing researchers.

—Danielle Venton, EGEE

g-Eclipse: easier interface to both grids and clouds

The g-Eclipse Consortium released the g-Eclipse framework, which developers claim will provide an easy-to-use workbench for accessing both grid and cloud infrastructures.

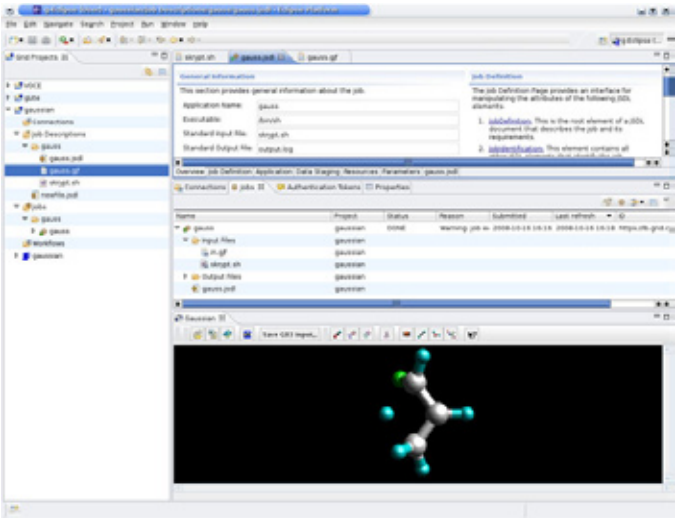
The software provides a graphical workbench that enables seamless access with the same simplicity as accessing the Internet from a browser. "This enables interoperabilty between different grid and cloud infrastructures on the client side," said Harald Kornmayer, a researcher at NEC Laboratories Europe, who led the project.

It currently supports EGEE's gLite grid middleware (aimed for scientific domains), and the GRIA middleware (used by industry and commerce), as well as AmazonWeb Services' cloud computing and storage offerings.

A key feature of the g-Eclipse workbench is its independence from the underlying grid and cloud technology.

"We started with the goal of accessing available scientific grid infrastructures, such as the one built by physicists at CERN to analyze the large amount of data generated by the LHC. With the help of g-Eclipse, scientists can use this infrastructure more intuitively and improve their productivity and scientific results," said Kornmayer.

"g-Eclipse is a great tool for scientists, who want to work efficiently on the grid. It can help with many day-to-day tasks, supporting the user by relieving her/him from tedious basic level tasks. It is therefore an important ingredient of any future grid infrastructure,"



A monitoring tool used by Grid Observatory to visually describe the success of job clusters over several months. Image courtesy of Grid Observatory

says Dieter Kranzlmüller, director of the Leibniz-Rechenzentrum high performance computing center in Munich, and strategic director of the European Grid Initiative.

Since autumn 2006, g-Eclipse is also an Open Source technology project. The g-Eclipse Consortium was supported by the European Commission in the 6th Framework Program in the field of research, technological development and demonstration. Support for the cloud offering from Amazon Web Services is also available.

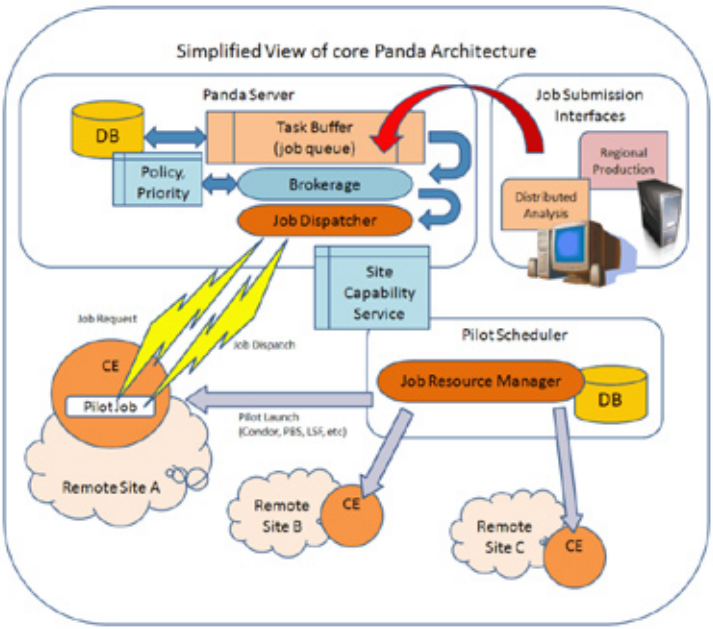
Munich's Leibniz-Rechenzentrum hign performance computing center. Image courtesy of LRZ

The g-Eclipse Consortium is formed of eight European partners, coming from both academia and industry.

—Dan Drollette, iSGTW

PanDA makes huge job sets more bearable

Keeping track of huge job sets processed on hundreds of compute clusters around the world through the LHC Computing Grid might send the most organized of logical thinkers into a tizzy. The PanDA (Production and Distributed Analysis) system, developed for the ATLAS collaboration at the Large Hadron Collider, lets scientists stay cool while it takes charge of distributing jobs, collecting results and managing workflow.



Simplified view of core PanDA architecture. Image courtesy of BNL.

in turn locates available resources and sends the collected jobs to run based on their relative priorities. The pilot system manages the workflow efficiently, providing a quick response time. And it frees users from tedious decision-making, said Kaushik De, a PanDA developer and University of Texas physicist.

PanDA was initially developed in 2005 for U.S.-based ATLAS production and analysis on the Open Science Grid, but it has since been adopted by the global ATLAS collaboration as its primary system for distributed processing. ATLAS uses a total of three different systems —OSG, EGEE and Nordugrid—but PanDA is the interface to them all.

An important feature of PanDA is that it allows the user to submit one job, called a pilot job, which coordinates a series of jobs that the user has put together and configured. When launched, the pilot job contacts the PanDA server, which

ATLAS has also developed a separate data management system, variously called DQ2 or DDM (for "Distributed Data Management"), that catalogs the tens of millions of ATLAS files distributed worldwide at hundreds of storage locations. PanDA works seamlessly with DQ2/DDM to match user jobs to the input data required, either by sending the job where the data already resides or vice-versa.

At the moment, PanDA's jobs produce and analyze simulated data, which physicists can use to fine-tune their analyses in preparation for real data once the LHC is operational.

As of January 2009, PanDA had processed more than 25 million simulated data jobs. Its current daily rate is split into about 50,000 data production jobs and between 3,000 and 5,000 analysis jobs. Once real data starts coming in, scientists estimate job counts to approach 500,000 jobs a day.

"PanDA makes it possible to use huge amounts of computing resources distributed all over the world," Kaushik De said. "Without a system like PanDA, it would be almost impossible for physicists to do the type of large-scale processing necessary to analyze their data and quickly get results."

—Amelia Williamson, for iSGTW

This panda may not be so productive or widely distributed, but he knows his priorities. Image courtesy of sxc.hu.

Application Porting Support Group celebrates its first birthday

The Application Porting Support Group, EGEE's service that helps end users get their existing applications to work on the grid, is now entering its second year. A lot of work went into the group's first year successes — they even won "best demo award" at the EGEE'08 Conference. What kind of an experience has it been for them?

Gergely Sipos (MTA SZTAKI) and Jose Luis Vazquez-Poletti (UCM) with the certificate for the best infrastructure demo of EGEE'08. Image courtesy Gergely Sipos

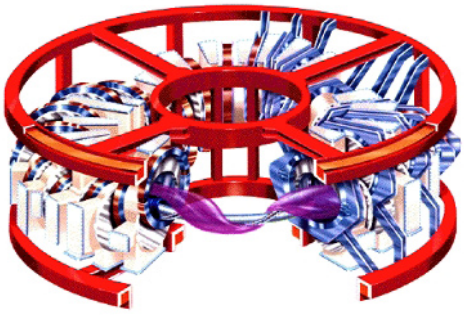
"It was a busy year for us," said Gergely Sipos, coordinator of the porting group at the Hungarian Academy of Sciences in Budapest. "We had to set up the group and integrate members from Hungary, Italy, Spain, France and Taiwan. We needed to define the working environments, infrastructures, protocols and the support cycle itself."

Sipos said that the group tries to fulfill every application porting request.

"As our growing set of success stories shows, we've ported quite a few applications so far." Of about 30 requests over the year, the group successfully completed 22. The rest, Sipos said, were either better suited to supercomputing environments or required licensed software unavailable on EGEE.

The variety of user groups with very different backgrounds and needs posed challenges to the group. Some researchers simply could not cope with the complexity of grids by themselves, and others — mainly representatives of more established grid user and development

projects — requested support for more complicated scenarios of their applications already running on EGEE.



Simulation of a microwave beam (purple) in a magnetic confinement fusion device, as rendered on the grid via Gridway. Image courtesy Grid Application Support Center

In order to meet this wide range of requirements, the group took advantage of available tools from the RESPECT program (Recommended External Software for EGEE CommuniTies), which provides extensions of core EGEE gLite middleware. These extensions allow for enhanced functionality such as complex workloads with parameter sweeps (running a model many times with different values for the various input parameters), workflows (sequences of operations), MPI (message passing interface) programs, portals and interactive jobs.

Learning from experience

The group itself reaped benefits from the sheer variety of applications they ported.

"The more applications we worked with, the more experienced we became," said Sipos. "Moreover, because the providers of several RESPECT tools were also involved — most notably the science gateway extension of P-GRADE Portal and gLite's CREAM Computing Element in Gridway — their work advanced as result of the direct feedback from clients."

The group works with new user groups as needed, and will continue to offer clients support during the transition from EGEE to EGI, expected to begin in autumn 2009.

—Danielle Venton, EGEE

egee

Profiles

Paola Celio: ARGO, ATLAS and growing the grid

Paola Celio, like many "grid people," got her start in particle physics.

Now a technical engineer at the Department of Physics at the Roma Tre University, and associated with the Italian National Nuclear Physics Institute, Celio has moved from physics to grid hardware and software, attracted by the tremendous potential of the field.

The first Grid Application School of EUMedGRID Project, April 2007, Cairo. Image courtesy of Paola Celio

"I started working with grids because I was interested in using something that involves many people, and because it is a field that is evolving very quickly," Celio explains.

"Grids are still in their infancy, but with them we can create a very different future. If we are successful, ultimately the whole scientific community will be able to use grids as just one more tool."

Celio is very keen to pass on her knowledge of grids.

"In addition to developing applications, I help other people put their applications on to grids—mainly for physics, but also for other fields such as biology or archaeology," she says.

"Particle physicists are very experienced with computer codes, so we hand on our experience and the components we develop to support other groups. Grids are not only for physicists."

Bringing people together

In addition to the physics experiments ARGO-YBJ, which analyzes cosmic rays, and ATLAS—one of the LHC experiments—Paola is also involved in several other grid projects, in particular EUChinaGRID and EUMedGRID.

"Collaboration in these projects is an exciting experience," she says. "We work with many new communities in the Mediterranean area, and there is just so much enthusiasm in these new groups. They are keen to work with this new field. Seeing this enthusiasm from others also motivates me for my own work."

Celio says that grids bring people together from many different cultures, crossing boundaries and spanning continents.

"Recently I spent 15 days in Cairo, Egypt, for a school on applications," she says. "This kind of work really is more than science—it is a unique social experience, which makes it very satisfying."

Celio was also astonished by the number of women she met at this and other events.

Some of the places Paola Celio has been are quite remote, as in the site of the Argo-YBJ gamma-ray detection experiment, located in Tibet, 90 km north of Lhasa. Image courtesy of Paola Celio

"Countries such as Egypt, Turkey and Tunisia have more women in this field than in Italy! In this way they are more advanced than some of the so-called 'developed' countries."

- *Hannelore Hämmerle, EGEE*

People behind the LHC grid: Jamie Shiers

Jamie Shiers is part of grid support for the Large Hadron Collider, and has been working on the LHC for 16 of his 25 years at CERN. Dressed in sneakers and jeans, and looking tan after a recent grid conference in Brazil, he talked with *iSGTW* at CERN's Restaurant 1.

iSGTW: What first got you interested in physics?

JS: It could be something in the genes, I suppose. I and two of my three brothers became physicists—and we've all passed through CERN.

For me, what helped to get me hooked was a book about physics, called *Mister Tompkins in Wonderland*, that I read at age 14 or 15, by George Gamow—a Russian physicist who worked with Niels Bohr. He wrote about a world in which light was slowed down enough that you could see the effects of relativity. It was a great book that turned me on to physics . . . that, and the old cult, UK sci-fi "Dr Who" television series.

iSGTW: Do you remember when you first started on the LHC?

JS: I can look at my calendar and find the exact week. It was in September 1992, at the same conference in which we discussed the uses of the World Wide Web, applications, and object oriented computing. So it was from very early days.

iSGTW: How did you first get into computing?

JS: It sort of just happened. I was an experimental physics PhD student here at CERN, from Britain, when I realized: "Right, I am in computing."

Image courtesy of CERN

iSGTW: How does the grid look now, so far as the startup goes?

JS: In May, we met or exceeded all of the metrics in our last Common Computing Readiness Challenge, even though I have to say that not all functionality was fully tested and the overlap from the experiments was somewhat limited. So I feel confident.

But we know that the coming of the real data will be the real test; there may not be as tight a synchronization as we expect, for example. From my experience with the Large Electron Positron Collider, I know that there's always something that you planned for that didn't

work the way you thought, or a solution that comes up that works so much better than what you previously had in mind. What works is what flies.

iSGTW: So the Common Computing Readiness Challenge was a help?

JS: Absolutely. From that experience, we know that there will not be any showstopper. There's always a way up, over, under, around or through any problem.

iSGTW: What do you do in the LHCGrid project?

JS: My role is to keep it all running smoothly. It's easy to see your own little bit of this enormous undertaking, but not the overall picture. You might find a solution here that breaks something over there. You are also dealing with different time zones, different cultures, in institutions with their own priorities.

iSGTW: After 16 years with the LHC, what's going through your mind?

It's an exciting time. We don't know exactly what we will find, but we know that there will be big headlines in physics, and we'll be sitting in that auditorium upstairs when it's announced.

iSGTW: What do you see as the value of grids?

JS: At the petascale level of computing, there are very few places with resources on the level of a Los Alamos or a Lawrence Livermore, that can afford their own supercomputers. But with grids, you can get access to big resources, using commodity computers, for a price that is an order of magnitude lower.

And with that infrastructure, you inherently have a more equal collaboration. The Tier-0, Tier-1s, and Tier-2s are all equal in this project, with the small institutions more than pulling their weight. I'm willing to bet that when we do the first pass at the Tier-0, we won't find anything definitive, not only because there will be little time but also because the calibrations and even algorithms will not be fully tuned. The findings will likely be at the Tier-1s on the re-processed data with refined calibrations and algorithms, and from analyses performed primarily at the Tier-2s.

As I wrote in the paper I just presented in Brazil, ("Grid Today, Clouds on the Horizon") with grids, you're actually building something—a long-term and sustainable e-Infrastructure for the future, that helps institutes do science and research.

And who knows where that might lead. Lightning may strike twice.

—Dan Drollette, iSGTW

People behind the LHC Grid: A Bird's eye view

Ian Bird is project leader for the Large Hadron Collider Computing Grid, where he is responsible for grid deployment and all of the technical work involved in setting up the grid infrastructure for the LHC.

iSGTW: If you were at a cocktail party, and someone asked "What do you do," how would you answer, in lay person's terms?

IB: Well, I'd try to change the subject. (Laughs)

Seriously, I try to coordinate the overall activities of the LCG: the experiments, the service providers, the money and the people. I report to CERN management and oversight, to make sure all the nuts and bolts are in the right place. And, of course, I think about all the technical aspects.

When the LCG first started, the idea of a grid, with its tiered hierarchy, was still new. Image courtesy of Ian Bird

iSGTW: How do things look now?

IB: Well, we're very upbeat. Of course, we always knew that there would be surprises, but we are handling them. As has been the case for the past several years the grid is in constant use for simulations and is used for evaluating the data coming in all the time from the naturally occurring cosmic rays that are hitting the detectors—and this is helping to calibrate the equipment.

iSGTW: Did you always have an interest in physics and computing?

IB: Believe it or not, but when I was in school, my headmaster was not very encouraging about my getting into physics. He told me to get into something else. But that didn't deter me—I always wanted to know how the world works. That's the beauty of physics, that it goes into the how and the why.

So, I pursued physics—and computer science—anyway. And that's the message I try to pass on to my children about their choice of careers: do what you are interested in. As for computing, ever since I started as a physicist, I was always involved in the software for the experiments I was working on. In the

beginning of the '90s, we had one of the very first SUN clusters, and we were one of the earliest to use CONDOR .

iSGTW: How long have you been at CERN?

IB: I first came to CERN in 1977, as a graduate student, to work in the BEBC Bubble Chamber. Then I was back in 1981 as a post-doc with the European Muon Collaboration experiment. I then did a series of collaborations into the 1990s, on projects such as NOMAD, and then I went to Jefferson Lab in the States in 1996. In 2002 I came back to CERN, and then took over from Les Robertson after he stepped down earlier this year.

And Les was really one of the key people behind LHC computing and the grid. He got the funding and the instrumentation, and got all the governing bodies on board.

I've now been associated with LHC computing for about nine years, first in the Particle Physics Data Grid project in the US, one of the precursors to OSG.

iSGTW: What was it like in the early days?

IB: In the year 2000, we had a model for handling the vast amount of data we knew the LHC would generate, which already featured the structure of tiers and high speed networks. At that time the new grid technology suggested a way to implement the model, while advances in network technology solved the throughput problem; we could just buy off-the-shelf products for the network.

iSGTW: What do you think the grid needs to do to become more widely accepted?

IB: The current grid middleware is not very user-friendly, it's a bit baroque not only for end-users but for site managers as well. That's probably the biggest problem. But any new technology has to go through this phase; the difference in our case being that we're trying to use it now in large-scale production. But over time, it should become easier to use and more robust.

iSGTW: In your mind, what do you think we will find?

IB: I can't even imagine it. Whatever it is, we're ready.

iSGTW: Any parting thoughts?

IB: This scale of collaboration has never really been done before in computing. The technology is one thing, but getting all these centers to work together is quite an achievement.

What CERN has going for it is that it has always been about international collaboration. The science transcends the politics.

—Dan Drollette, iSGTW

People behind the grid: Erwin Laure

Erwin Laure was technical director for Enabling Grids in E-sciencE, and worked for CERN for the past six-and-a-half years. *iSGTW* caught up with him during his last full week at EGEE, before he was to take up a position in Stockholm at the Royal Institute of Technology, known by its Swedish acronym "KTH."

iSGTW: If you had to describe what you do at EGEE in two or three sentences, what would you say?

EL: I help keep the project's technical end running, on a short- to medium-term basis. I make sure that we meet the requirements of users, so far as middleware, infrastructure, training and coordination of our 100-plus partners goes.

iSGTW: How did you come to EGEE?

EL: Just as I was finishing up my PhD in high-performance computing and business administration at the University of Vienna in the late 1990s, this thing came along called the "grid." It sounded interesting, and I knew there was something starting up on the Swiss-French border, at CERN, called the European Data Grid (EDG). I started out on a 2-year fellowship at CERN, working on data management tools. Then Bob Jones asked if I'd be interested in working with him on the overall coordination of EDG . . . I thought it would be a fun project, in a nice area, with interesting people from diverse backgrounds, and I could learn a lot. EDG evolved into EGEE, and I was on board from the earliest days, focusing on EGEE's middleware. I later became technical director.

I also travelled a lot—to every country in Europe, to Asia, Australia and America. They made me see the world. And, of course, I got to see Switzerland.

iSGTW: Where will you be going?


Image courtesy of Erwin Laure

The high-performance computing center at KTH, where I'll oversee servers and storage for applications in research and education. KTH has a first-rate international reputation, and provides services to organizations such as the Karolinska Institute, a well-known medical university.

It will be interesting. I've never been to Sweden long-term before, although I have taught classes at a summer computing school there. But I will have to get used to Swedish dishes such as "surströmming," or fermented herring. I think fermented is another way of saying "rotten." (Laughs.)

iSGTW: How did you become interested in computing?

EL: It was something encouraged at the high school and college where I went, back home in Austria. I always liked programming, dealing with languages such as BASIC and FORTRAN. And remember that this was during the '80s, when just about every computer game could be taken apart and reprogrammed.

iSGTW: What is your opinion on the nature of the grid?

EL: The grid is fundamentally an academic, research and science structure in a non-profit setting. Only a small part of it is about technology. The rest is about

The grounds of the Royal Institute of Technology. Image courtesy of KTH

science in a collaborative environment, and sharing resources to reach a common goal. In the early days, I think the term "grid" was oversold, as if it were some kind of silver bullet that would eliminate the need for coding, managing machines, or setting up algorithms. Instead, you still must do all those things, but with the grid as an added service level that lets you expand the reach of your computing power and data. The good thing is that with the hype over, people now have more reasonable expectations, and can do real work.

iSGTW: What does the grid need to do to gain more acceptance?

EL: Two things: It needs to be easier for users to access, and easier for resource providers to join. The grid was developed by experts for experts, and they wanted as much functionality as possible, without paying as much attention to ease of usage. Meanwhile, if you are a resource provider and you want to join the grid, there's still a high threshold to overcome. We've improved it a lot over the years, but it is still complex and it needs a lot of babysitting.

So, we have to make it easier on both sides—for the end user and the resource provider.

iSGTW: What do you think you will miss, once you leave here?

EL: The people, obviously, but also the warm autumn days. I may miss some of the food as well. I'm bringing my fondue pot, just in case.

—Dan Drollette, iSGTW

People behind the grid: Simon Lin

Simon C. Lin is currently coordinator of the Asia Federation in EGEE and is responsible for the ASGC—Academia Sinica Grid Computing Center—the only WLCG Tier-1 Center in Asia. He is deputy project manager of the EUAsiaGrid project, coordinated by ASGC and INFN in Italy.

iSGTW: How did you get involved in particle physics?

I am a theoretical physicist by training. Like many of my generation, I was inspired by the great names in physics of the first half of the 20th century. I was fortunate enough to do my Ph.D. at Edinburgh University under Peter Higgs. Firstly, I worked on applying group theory to supersymmetry but then turned to field theory on surfaces and interfaces later. I still feel theoretical particle physics is excellent training for physicists, independent of what they go on to do.

Simon Lin at the Academia Sinica Grid Computing Center. Image courtesy of Kevin Wang

iSGTW: How did you come to computing?

In Edinburgh, I had Dave Wallace as a co-supervisor. Apart from his important work on renormalization group theory, he was also a pioneer in scientific computing. For example, he introduced massively parallel processing in Edinburgh at a very early stage. Later, he helped people like Tony Hey start the e-Science program in the U.K.

Dave's enthusiasm for scientific computing was a big influence on me. So when I went back to Taiwan I pursued a keen interest in computing for physics. As a result, I became director of the computing center at Academia Sinica for 15 years.

iSGTW: How did Academia Sinica become the Asian Tier-1?

By about 2002, many different flavors of grid middleware were appearing in Asia. These were basically regional variations of Globus. I was attracted by the more pragmatic approach being taken for LHC computing, where the idea was to take existing middleware from different sources, scrutinize it, fix potential weaknesses and deploy the result. The fact that LHC physicists would rely on this middleware for at least a decade convinced me that it had a good chance of survival. As a result, Taiwan became one of the first Asian countries to participate in the LCG Grid Deployment Board meetings. At first, I didn't think that the enormous

resources for a Tier-1 would be feasible in Taiwan. But, it became clear to me that if LCG were going to be a 24/7 service, being in Asia was an advantage when it came to supporting that service, because of the time zone difference.

This was an opportunity for our staff at Academia Sinica to become experts and contribute to this new technology. So, with the help of the National Science Council, we moved quickly, sent a number of young researchers to CERN for training. Thanks to this early start, and with the support of our Japanese colleagues, Academia Sinica became the Asian Tier-1. We were the first Tier-1 to sign a Memorandum of Understanding with CERN, in 2005.

iSGTW: What are the main challenges of running the only Tier-1 in Asia?

A big technical challenge has been getting bandwidth at a reasonable price. When we started there were no 10-gigabit lines available between Taipei and CERN. Even today, trans-Pacific links are twice the cost of trans-Atlantic ones, and going via the Indian Ocean is eight times more expensive. Fortunately, I had been involved in the '90s in promoting major network upgrades for Taiwan, and have been able to negotiate good deals for setting up the necessary connectivity for the Tier-1.

Traditionally, in Asia we are playing technological catch-up with the West. So we tend to collaborate with the United States or Europe, rather than Asian countries, which we may see as regional competitors. I think that LCG and the Tier structure has provided us with a strong motivation for better regional collaboration.

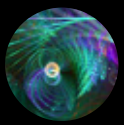
iSGTW: Looking forward, what do you see as the opportunities for grid computing?

I think grid software has great room to improve. Some sort of simple, universal protocols need to evolve, to overcome the current challenge of interoperability between different grids.

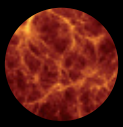
At Academia Sinica, we are involved in other grid activities, including the WISDOM virtual drug-docking challenges we have collaborated on for malaria and avian flu, which showed how quickly and effectively a grid infrastructure could respond to real research needs. Grid technology for preservation and long-term digital archiving as well as disaster mitigation, are areas where I see a lot of potential, if we can convince different regional institutions to share data.

I've always felt that the development of computers would have a long-term impact on science. Inspired by visionaries like Richard Feynman, scientists have long speculated on what computing might be like if it were based on new paradigms, like quantum computing. In the same way, I think that grid computing is a conceptual change that may have a huge long-term impact.

—Francois Grey for iSGTW

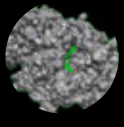
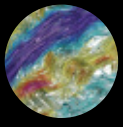
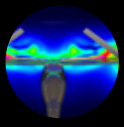


EGEE




Also in the news



Enter the age of computer merchants

Modern researchers are not short of ambition. Physicists at the Large Hadron Collider at CERN, for example, will need to catch and sort through an estimated 15 petabytes of data each year, the equivalent of 20,000 years of music on an MP3 player.

Collaborations like WISDOM, a global initiative for discovering new medicines for neglected and emerging diseases, test tens of millions of chemicals in computer models.

Goods for sale: companies like Digital Ribbon can quickly and easily connect clients with computing resources tailored to their needs. Image courtesy of sxc.hu

This is an opportunity for companies like Digital Ribbon in the United States, which seeks to be a kind of clearing house for computational resources. They call their model a "service registry," connecting resource consumers with the right resource providers. It could transform the way users run jobs on clouds or grids.

"Computing is moving towards becoming a commodity," says Erik Weaver, Digital Ribbon's CEO. "However it is much more complex than the oft-cited example of electricity."

Computing consumers care about more than power, they also care about bandwidth, interconnectivity (how fast individual processing cores can talk to each other) and memory. It is important to have not just enough resources, but also the right kind. As Weaver puts it:

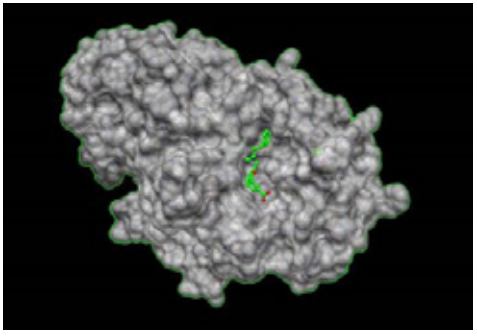
"If you are making an apple pie, you might want to use Granny Smith, not Fuji apples."

Having more options could be a great advantage to users.

"We are always looking for ways to link user communities, applications and re-

source providers," says Bob Jones, project director of EGEE.

As a step towards smoothing the way for EGEE and Digital Ribbon collaboration, the WISDOM collaboration ran a test run early this summer, using gLite middleware on Digital Ribbon resources.



WISDOM uses grid-powered software to screen for potential drug-leads. The image shows a simulated docked compound (green) that inhibits the alpha amylase enzyme as a means of treating diabetes. Courtesy of Jean Salzemann

Quantity, quality . . . commodity?

"This successful test with WISDOM shows that applications within EGEE can run well on Digital Ribbon resources," says Jones. "Now we can see if other user communities might benefit."

Data used for the test consisted of 750 chemical dockings—potential candidates for diabetes drugs. In about 12 hours they ran calculations corresponding to 55 days on a single processing unit.

"While this was small for us in comparison to previous sets of calculations," says Jean Salzemann, computing researcher with the WISDOM collaboration, "it tested new ways of sending jobs. When we are preparing future data challenges, we could consider splitting the data load between EGEE and Digital Ribbon resources."

In their last data challenge, WISDOM tested 4 million potential drug candidates using 5,000 processing cores. With these they were able to run in two months what it would have taken a single CPU four centuries to complete. If they had 10,000 cores, they could have done the calculations in one month.

"We test millions of potential drugs in our data challenges," says Salzemann, "ideally we would have access to one core per potential drug or 'ligand.' With this we would have the complete results in about an hour."

—Danielle Venton, EGEE

Digital Ribbon's service registry covers a multitude of platforms, processors, architectures, offering resources to public and private organizations.

For the love of movies: recommendations from the grid

Alfred Hitchcock once said, "A good film is when the price of the dinner, the theatre admission and the babysitter were worth it."

Tired of wasting your time sifting through DVDs at the rental store only to end up with a mediocre film? Help is on the horizon for all interested in movie recommendations.

In a one-man project, Leandro Neumann Ciuffo, offers grid-powered film recommendations through a simple algorithm, his Web portal and EELA-2, the E-science grid facility for Europe and Latin America.

"This is a very simple application," he said. "It's not going to cure cancer, but it does show that average people can use the grid."

Custom-tailored suggestions

To find a good rental, a movie buff would visit Ciuffo's Web page "Cinefilia," create a user profile and evaluate 20 movies. The tool looks at the user's tastes, and makes recommendations of its own, using a classic algorithm known as "collaborative filtering."

In this algorithm a new user's profile is compared with all other users in the system, similar to the tool that Amazon uses to recommend books. (Or that "Jester," a site from the University of California at Berkeley, uses to recommend jokes.)

Marilyn Monroe gets fitted for a scene with the help of Hollywood costume designer William Travilla, who custom-made Monroe's outfits for eight of her films. Image courtesy of The Travilla Tour

Ciuffo, an application manager for EELA-2 working at INFN, Catania, Italy, enjoys a good flick like everyone but began developing Cinefilia for the most practical of reasons:

"When I started working I didn't know how to port applications to the grid. I decided to take an application I developed during my master's degree and port it. So this is like a toy application for me to learn how to help others."

Cary Grant in North by Northwest. Image courtesy of MGM

What was recommended for this writer? The Godfather, Cinema Paradiso, Se7en and Casablanca.

Ciuffo presented his application at the 4th EGEE User Forum in Catania, Italy. See a video of it, filmed on site and prepared by GridTalk (on GridCast.org).

—Danielle Venton, EGEE

DANTE dances to the volcano

Domenico Vicinanza, an engineer at DANTE (Delivery of Advanced Network Technology to Europe) takes recordings of the seismic activity of volcanoes and transforms them into sound waves, in order to help predict eruptions.

Doin' the volcano dance. Image courtesy of CityDance Ensemble

A musician as well, Vicinanza went one step further, converting sound waves from multiple volcanoes — Mount Etna, Mount Tungurahua, and the Mountains Pinatubo and Mayon — into a melody, which he then composed into music.

On March 29, the American modern dance company CitySpace Ensemble, on tour in Cambridge, England, (see image above) performed to his "earth-shaking" music . . .

Researchers have long sought to predict a volcano's eruptions by looking for patterns in its seismic behavior. With the use of complex sonification algorithms, Vicinanza found he could take recordings of a volcano's seismic behavior and translate what he found into audible sound waves, thereby transforming the raw seismic data into something easier to use for predicting eruptions.

Intrigued, Vicinanza went one step further, using the computational power of the grid to convert audible sound waves from multiple volcanoes—Mount Etna, Mount Tungurahua, and the Mountains Pinatubo and Mayon—into a melody, which he then composed into music for a dance performance.

And on March 25, City Dance Ensemble performed to his volcanic music. The dance, titled "The Mountain," was part of the ensemble's *Carbon*, a larger

work about climate change. Originally presented in sold-out performances on March 14 and 15 at the Music Center at Strathmore, Maryland, it was later repeated on March 28 and 29.

Songs of the Earth

"As a scientist, it was my priority to develop tools to help us predict eruptions and ultimately reduce the loss of lives," said Vicinanza. "As a musician and artist, it was a natural step for me to take these seismic sonification sounds and apply them to the arts. I am delighted that the results, or songs of the earth, are being created into a dance performance."

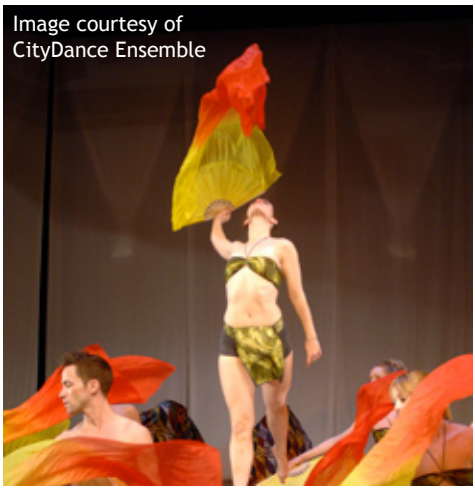


Image courtesy of CityDance Ensemble

First of its kind, the computations for the event were run on DANTE, EGEE, and the E-science grid facility for Europe and Latin America, or EELA. The complex sonification algorithms that convert the seismic data into sound melodies require the power of the grid, as the process would be nearly impossible using standard bandwidth networks or computing resources.

Research and education data communications networks GÉANT2 in Europe and TEIN3 in the Asia-Pacific, both operated by DANTE, as well as Latin America's RedCLARA operated by CLARA, underpin the immense computing power provided by EGEE in Europe and EELA in Latin America.

For his part, Paul Gordon Emerson, CityDance Ensemble choreographer and Carbon curator, said: "High bandwidth research and education internet networks together with grid computing power played a vital part in making this project a reality . . . the fact that this work uses the voices of the earth from three continents is a very powerful metaphor for *Carbon* as a project and as a concept."

—Dan Drollette, iSGTW

Link of the Week - The Epigonion is heard once more

As we reported last September in "Ancient musical instrument comes back to life," a harp-like instrument from ancient Greece can now be heard again for the first time in centuries, due to the grid and a computer modelling project known as ASTRA (Ancient Instruments Sound/ Timbre Reconstruction Application).

Music historians believe that the Epigonion's closest relative, was a "psaltery," shown above, a musical instrument that was plucked or hammered. Image courtesy of Wikimedia Commons/ Public Domain Image.

In the first week of March, after the 4th EGEE User Forum in Catania, Sicily, BBC Radio 4 treated its listeners to an interview with Domenico Vicinanza of ASTRA, in which he described the process of reconstructing this long-lost instrument.

And you can now hear a full-blown "concert" done in Naples, using virtual epigonions on the grid at www.astraproject.org/files/concertdemo.mov.

—Dan Drollette, iSGTW

Web: www.eu-egee.org
Contact: projectoffice@eu-egee.org
ISBN 978-92-9083-330-7

EGEE-III is co-funded by the European Commission under contract number INFSO-RI-222667