Grid TALK

GridBriefings

Grid computing in five minutes

Introduction to grid computing

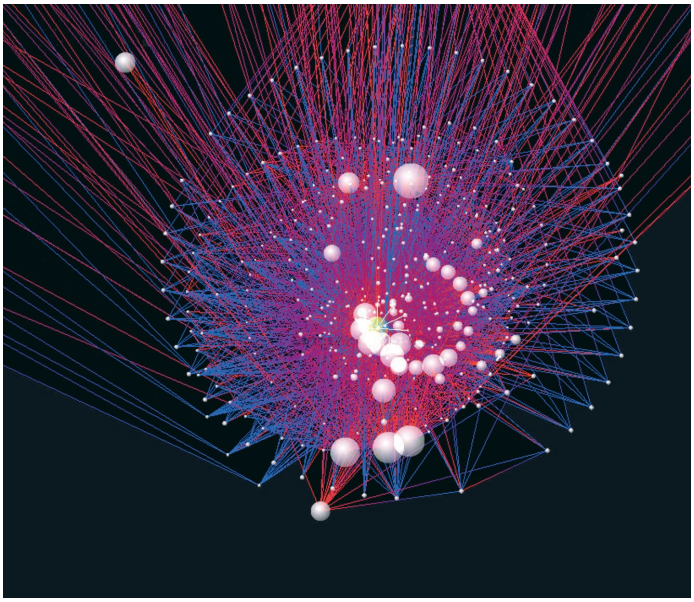
What is grid computing?

Grid computing connects computers that are scattered over a wide geographic area, allowing their computing power to be shared. Just as the World Wide Web enables access to information, computer grids enable access to computing resources. These resources include data storage capacity, processing power, sensors, visualization tools and more. Thus, grids can combine the resources of thousands of different computers to create a massively powerful computing resource, accessible from the comfort of a personal computer and useful for multiple applications, in science, business and beyond.

Why do grids matter?

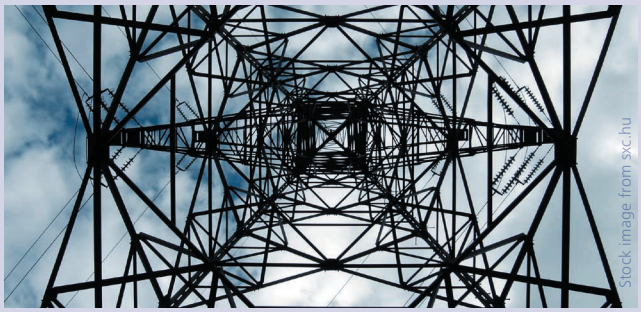
For scientists trying to solve highly complex problems, grid computing provides the power to help solve some huge questions: What happened just after the Big Bang? How will global warming affect our lives in the future? Is there a cure for malaria or cancer? Grids are helping researchers find answers to these questions. Grids also speed things up: a simulation that might take weeks on a single PC can run in hours on a grid. This means grids can react quickly to changing needs, providing an invaluable resource during crises like natural disasters or epidemics.

Further, since grids cannot work without people, the development of computing grids also develops communities. Grids encourage and require people from different countries and cultures to work together to solve problems.



Plugging in to "the Grid"

The term "Grid" was inspired by the "electrical grid," stemming from the vision that plugging into a computing grid could be as simple as plugging into an electrical grid. The idea was that, like an electrical grid, users could simply access as much computing power as they required, without knowing where it comes from or how it was produced. This is not yet the case, but has also led to the term "utility computing," where computing power is viewed as a utility, available on a pay-as-you-go basis, like gas or electricity. "Cloud computing" is another form of utility computing that is gaining popularity.



Stock image from sxc.hu

What are the challenges of grid computing?

Grid computing works because people participating in grids opt to share their computer power with others. This opens many questions, both social and technical. Who should be allowed to use each grid? Whose job should get priority in the queue to use grid power? What is the best way to protect user security? How will users pay for grid usage?

Answering these questions requires all-new technical solutions, each of which must evolve as other grid and information technologies develop. Since grids involve countries and regions all over the world, these solutions must also suit different technical requirements, limitations and usage patterns.

Grid computing in the future

Grid computing isn't simple, but its potential is huge. Grids are providing technology that can transform the world's computing resources into seamless computing powerhouses, allowing new ways of doing science and enabling new virtual research communities. They are helping communities across the globe in their quest for improved knowledge of our planet and our universe, which means boons for our heath, economy, environment and future.

Case study: grids and high energy physics



Image courtesy of CERN

The challenge: improving our understanding of the Universe - the LHC

The Large Hadron Collider is a particle accelerator built as a collaboration between more than 50 countries. The world's largest machine, it accelerates atomic particles to nearly the speed of light and then steers these particles into 600 million collisions every second. Data from these collisions is expected to change our basic understanding of antimatter, dark energy and more. The LHC will produce 15 million gigabytes of data a year: the storage capacity of around 20,000,000 CDs. Thousands of physicists all over the world want timely access to this data.

The solution: the LHC Computing Grid.

The LCG combines the computing resources of more than 140 computing centers in 33 countries, aiming to harness the power of 100,000 computers to process, analyze and store data produced from the LHC, making it equally available to all partners, regardless of their physical location. In 2007, during the lead-up to LHC start-up, the LCG ran 44 million computer programs. This number is set to double to reach 100 million programs in 2008.



"As a particle physicist in the 90s, I witnessed the birth of the World Wide Web: I still remember the trouble I caused in my experiment when I suggested we create a web page to facilitate interaction between international members. I have since witnessed the birth of grid computing, and am deeply involved in the Large Hadron Collider Computing Grid. I can't imagine how LHC physicists could produce science without it. Grid technology enables the science to be worldwide and collaborative." *Faïrouz Malek, LHC ATLAS experiment physicist and scientific project leader for the LHC Computing Grid in France*



"Through LHC research, we hope to uncover the mechanism that causes particles to acquire mass, and to confirm or modify the longstanding standard model of physics. To contribute to these discoveries, the CMS experiment produces a walloping five petabytes of data a year, equivalent to 45 MP3s a second. Moving and analyzing this data is a challenging job, but by using grid technology, we can distribute data to physicists around the world, drastically improving response time and developing local computing infrastructure that contributes to a large and international collaboration. Historically, advances in particle physics have also had significant impact in medical imaging, electronics, radiation therapy, and more" *Malina Kirn, LHC CMS experiment physicist, LHC Computing Grid and Open Science Grid*



"Each day, biologists, chemists, physicists, mathematicians and engineers are working together to create new nanoparticles that require exhaustive structural characterization. These analyses require intensive use of computational chemistry, which can be very computer-time expensive. Such revolutionary milestones in pharma, medicine, and other fields are only possible with collaborative grid computing and the use of emerging new field like nanoinformatics." *Fernando Danilo Gonzalez Nilo, Center for Bioinformatics and Molecular Simulations at the Universidad de Talca, ACTION-Grid.*

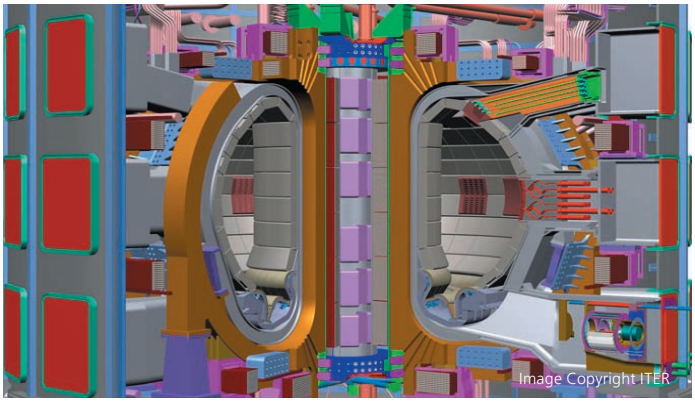


"Grid applications in the humanities tend to be data-centric. For example, datasets concerning ancient inscriptions and papyri may be of limited use individually, but we are using grid middleware to turn this scattered material into a powerful resource for research into the ancient world. The fuzziness and uncertainty of the data can also lead to computational challenges. For example, analyzing historical periodicals that have been converted to text using Optical Character Recognition would be infeasible for single machines." *Mark Hedges, Centre for e-Research, King's College London*



"Thanks to this grid-enabled application, we now benefit from unused resources in a simple way. This implies reduced processing time. Saving time during the processing phase actually enables us to concentrate more on the test phase and on reducing post-production related issues. We raise the customer's satisfaction and avoid a delay, which, at the end, makes it a really cost-effective solution." *Nicolas Hubaux, Art & Build, BEinGRID's 'Virtual Reality for Architects' application.*

Case study: grids and nuclear fusion



The challenge: safe, sustainable energy - ITER

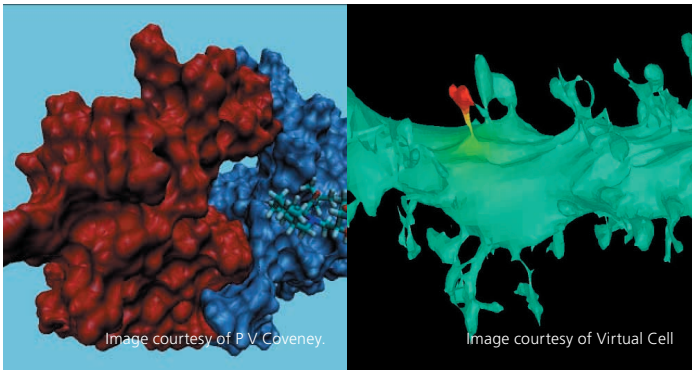
ITER, or the International Thermonuclear Experimental Reactor, will be built in France and is a collaboration between seven countries aiming to demonstrate that fusion energy can meet the world's changing energy requirements. A prototype fusion power plant, ITER will inject around two grams of fuel at a time, recreating the same process that fuels the sun and stars. The ITER device should generate 500 megawatts of power for extended periods, demonstrating the potential of fusion as a safe, clean and reliable source of energy.

The solution: grid computing, such as the Russian Data Intensive Grid

The Russian Data Intensive Grid is just one of many national grids providing ITER researchers with extra computing resources. Igor Semenov of the Russian Research Center Kurchatov Institute, Russia's (RRC KI) lead organization for nuclear research, said grid computing has been a boon for his team's research. "The load on RRC KI computing resources was growing, but thanks to our work on the grid infrastructure, we could involve the computing resources of other institutions. As a result, we can solve our problems about 50 times faster." RDIG are part of the Enabling Grids for E-sciencE (EGEE) grid infrastructure.

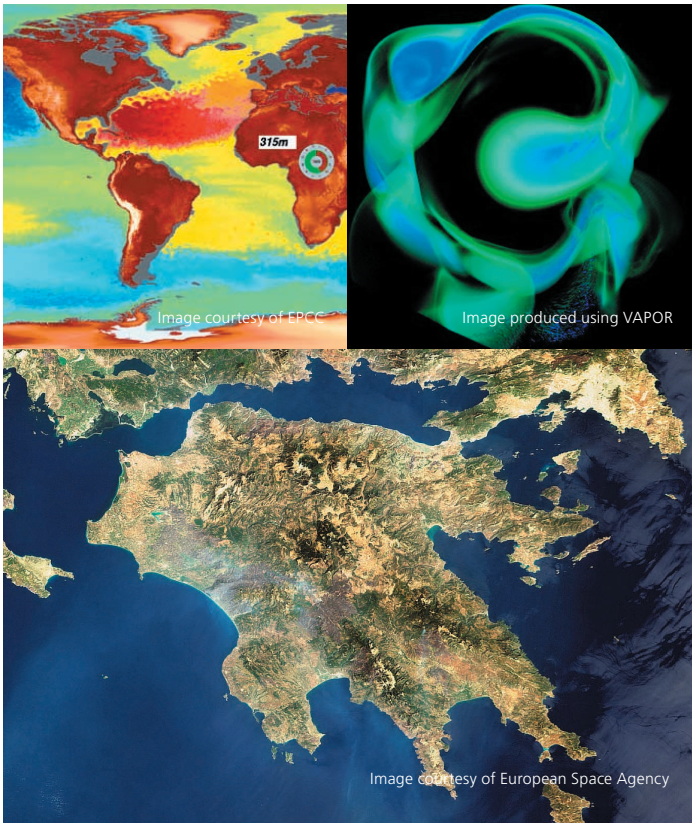
Case study: grids and human health

Grid computing assists researchers working in drug discovery for HIV, cystic fibrosis, cancer and more, as well as medical imaging, gene sequence analysis and pediatric medicine.



Case study: grids and earth science

Grid computing is contributing to studies of global climate change, air pollution trends, natural disaster management, coastal aquifers, ozone levels, flood forecasting and more.





"In earth science, grid technology is a very challenging tool for exploring new domains and helping to face key requirements of the civil sector. Grid is very useful for the exploitation of large sets of data which are currently not enough exploited due to a lack of tools, for running simulations and for statistical approach used for event prediction. Grid emphasizes the collaboration among various scientific and technical domains, a very fruitful point." *Monique Petitdidier, Enabling Grids for E-sciencE, IPSL and CNRS.*



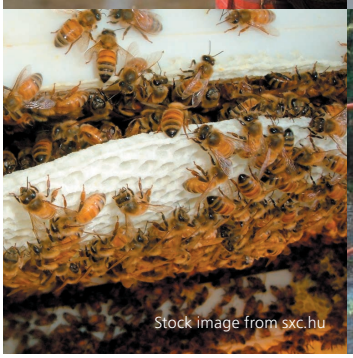
"I work on developing statistical software for studying the molecular origins of genetic diseases, such as Cystic Fibrosis. A large element of my research is focused on building a distributed data-management and processing facility for molecular biologists. Grid technologies have proved to be absolutely essential in implementing a reliable, secure and powerful system. The facility will go live following the start of a gene-therapy clinical-trial in Cystic Fibrosis and I am confident that, given its strong foundation, it will be a great success." *Rob Kitchen, National e-Science Centre, University of Edinburgh*

Case study: grids and agriculture

Grid computing helps to power research projects aiming to optimize farming practices, improve animal health, and fast-forward the search for agriculturally useful genes.


Image courtesy of the University of Washington.


Image courtesy of Wikipedia


Stock image from sxc.hu


mage courtesy of PISCES

The different layers of grid computing

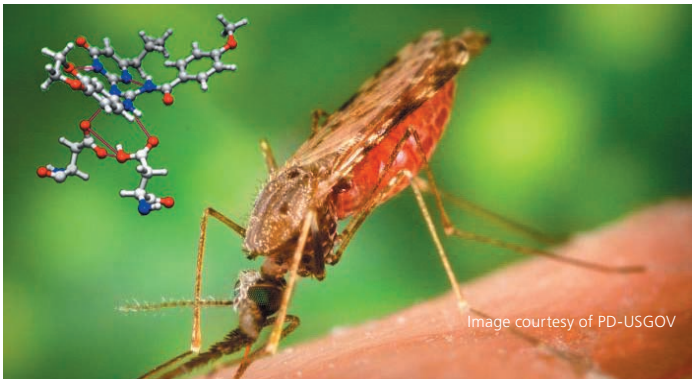
There is no single "Grid." Rather, there are many grids, created by groups of people who want to share their resources to increase their individual access to these resources. Some grids may contain only ten or twenty computers; others comprise many thousands. Some grids use supercomputers, others combine the might of a myriad of ordinary PCs. A computing grid has three basic layers:

Top level: grid users—for example medical researchers—interact with a top-level user interface or grid portal. From this, they can submit jobs for the grid to process—such as the analysis of hundreds of potential disease or drug proteins.

Middle level: the middle level is comprised of "middleware": layers of specially written grid software. Middleware automates scheduling of jobs, job allocation to different computers, grid security and permissions, and otherwise enables the various elements of a grid to cooperate.

Bottom level: just like the World Wide Web and email, grids rely on the Internet to link the computing resources of many different computers. These underlying networks allow communication between individual computers.

Case study: grids and medicine


Image courtesy of PD-USGOV

The challenge: narrowing the search for an anti-malaria drug

Malaria kills more than one million of the world's poorest people every year. Yet the development of new drugs to combat malaria—a disease that kills a child every 30 seconds—costs about one billion dollars in research time, equipment and personnel.

The solution: grid-powered WISDOM

The WISDOM project, a collaboration of eight institutions in five countries on three continents, began searching for anti-malaria drugs in 2005. Rather than go straight to the lab, the team used grid-powered software to screen molecules for their ability to disable a crucial malaria protein. In just six weeks, WISDOM crunched through the equivalent of 80 years of work, producing a short-list of 30 promising drug leads from a million initial candidates. Vincent Breton, WISDOM coordinator, said grid technology significantly cuts the costs of developing new drugs. "By finding good drug candidates for neglected diseases in such a rapid and cost-effective way, we can help to move away from the sad economic situation that led to their neglect in the first place."

For more information:

International Science Grid This Week:
www.isgtw.org

GridCafe:
www.gridcafe.org

International Thermonuclear Experimental Reactor:
www.iter.org/

Worldwide Large Hadron Collider Computing Grid:
http://lcg.web.cern.ch/LCG/

Russian Data Intensive Grid:
http://egee-rdig.ru/

WISDOM:
http://wisdom.eu-egee.fr/

GridTalk is co-funded by the EC under FP7