PDC Newsletter

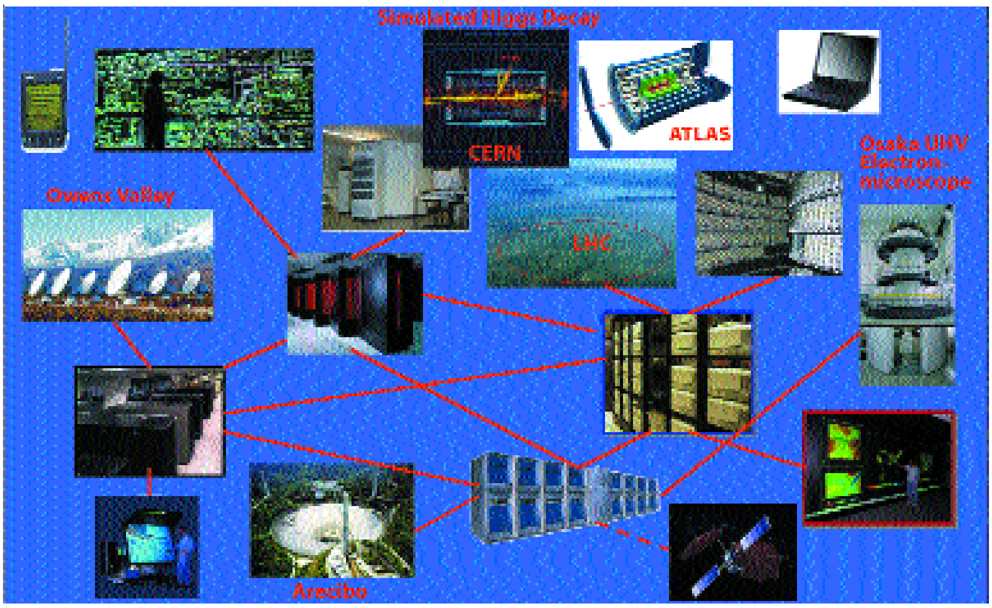
Grids – status and options

The STARTAP map (se article on page 4) and lecturers at the Annual Meeting 2001 at PDC, to the right Lary Smarr and Tom DeFanti, below from left Bill St Arnaud, Carl Kesselman and Tom DeFanti.

Grids are rapidly developing into an important prerequisite for advanced computing and data handling in research institutions, industry and many other organizations. Grids are assemblies of computers, scientific and medical instruments, digital databases, equipment for visualization and mobile communication together with a set of services for the effective and convenient use of the assembled resources. The importance of Grids is perhaps best measured by the commitments and investments by industry, with 12 companies last year announcing billion dollar investments in Grid technologies and application support based on the Globus toolkit. IBM, Sun Microsystems, Compaq, SGI, Cray, Fujitsu, Hitachi, NEC and Microsoft are all committed to Grids. Those utilizing Grids require ubiquitous access to the resources within the Grids, secure storage and transport of data and programs, and convenient means of discovering what resources are available under what conditions at various times. These basic requirements imply that a system for authenticating users and supplying services on the Grid must exist. There is also a pronounced need for management tools, such as tools for resource discovery, resource reservation, scheduling, task initiation and termination, file and account management, and monitoring of resource use. These services are often part of a software layer called "middleware" and represent what discriminates Grids from the Internet. *(… continued on p 6)* ➤➤

In This Issue

PDC and the development challenge



Lennart Johnsson is director of PDC since January 1, 2001. The PDC Newsletter is a quarterly publication providing brief information on interesting applications of high performance computing (HPC) and high performance networks, data mining and visualization, and information on hardware and software development for HPC, as well as relevant news for users of PDC resources. It is also available from the PDC web site <**http://www.pdc.kth.se**>

High-Performance Computation (HPC) in science and engineering is moving into a new era due to the rapid developments in communications networks, in particular wide area networks, with the deployment of all optical networks, Dense Wave Division Multiplexing (DWDM) technology, and Ethernet technologies for up to 10 Gbps at very low cost. Though at the moment the economy of high-speed and broadband networks seem bleak, close to 10 Tbps new transatlantic capacity has been deployed within the last year or so. Technology exists in commercial laboratories that would increase the capacity of fiber networks more than a 1000-fold. The prospect of practically unlimited bandwidth in core networks together with broadband wireless access will change the way we all work and live, and the way science is pursued and education carried out. Grids will be commonplace in this new environment. Today, dozens of projects and hundreds of engineers, scientists and graduate students are busy developing software supporting the vision of Grids. Worldwide annual investments in software for Grid middleware and Grid application development amounts to at least $50M. Application driven projects dominate, while five years ago middleware projects were the only recognizable Grid projects. This decade will be the decade of Grids in the way the 70-ties through early 80-ties were dominated by vector computing and the mid 80-ties through the 90-ties were dominated by parallel computing in computational sciences and engineering.

PDC recognized the potential impact of Grids early and initiated the development of expertise in the area in 1996, and has since then participated in several national and international efforts in development and validation of Grid software, both middleware and applications, including distributed visualization and steering.

Clearly, with the many formidable challenges to realize the vision, no one group or organization can be expected to provide all the necessary software, nor to have the winning concepts and ideas. Thus, much in the spirit of the Internet Engineering Task Force (IETF) that develops standards for the Internet, representatives from the research community in academia and national laboratories, and industry, have engaged in a community building effort to develop a forum for developing interoperability standards. PDC participated in the formation of the European Grid Forum and hosted its first workshop in December 1999. The European Grid Forum has since merged with the US Grid Forum and the Asia-Pacific Forum to form the Global Grid Forum. The recently proposed Open Grid Services Architecture (OGSA) is a significant step in the direction to an open architecture for Grid and Web services.

The need for Grids is strong in many disciplines, such as fundamental physics, a field in which experimental data is collected in very few locations, like the LHC project at CERN, while the analysis is carried out by a worldwide community. Astrophysics is another discipline in which increasing amounts of data are generated through instruments with increased resolution, often in remote locations and rarely where scientists are located and computational resources for data analysis and visualization may exist. There are many other similar examples in the traditional sciences and engineering. But, the life sciences and medicine provide perhaps even more challenging and rewarding applications since both data generation and users are distributed, and privacy as well as security issues are of paramount interest. Digital mammography is expected to generate at least as much data in the US alone as the LHC project, but unlike that project where all primary data is generated in one location, thousands of units will generate mammography data in the US, with the possibility of several tens of thousands of units generating data worldwide.

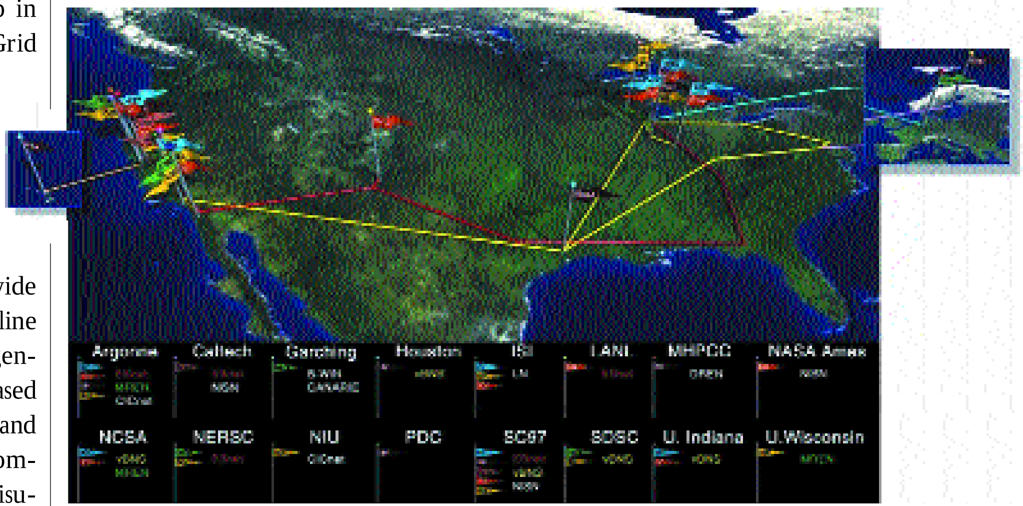
PDC has been fortunate to have the opportunity to partner with the international community in establishing testbeds and development of software and expertise. We have also been fortunate to have had the trust of PDC users to engage in some early Grid-like service deployment, like remote data and file service for SSF's Large-Scale Genotyping Laboratory in Uppsala and the Swedish Space Consortium's Odin project, as well as been entrusted to offer cluster based compute and file storage service for SBC, the Stockholm Bioinformatics Center, with participants at the Karolinska Institute, Stockholm University and KTH. The partnership with IBM for Vasaloppet demonstrates several of the unique elements of Grids. In this demonstration project sensor data is integrated with high-performance computing and data bases in real time. Grids should provide increased opportunities for the user community, and for collaboration with other centers providing resources for the academic community, industry and agencies of local and national government.

We view development of knowledge of the unique opportunities offered by Grids as an important challenge. One way in which we address this issue is by offering Workshops and Symposia. Another is by incorporating aspects of Grids in course offerings, such as our Summer School that is offered in August (see page 11) in collaboration with the National Graduate School in Scientific Computing. In this issue we briefly describe a few of the Grid projects PDC has undertaken. The hope is that this will stimulate our user community to review how the new opportunities offered by Grids can be beneficially used in its endeavors.

Lennart Johnsson

GUSTO Computational Grid Testbed

as of November 1997



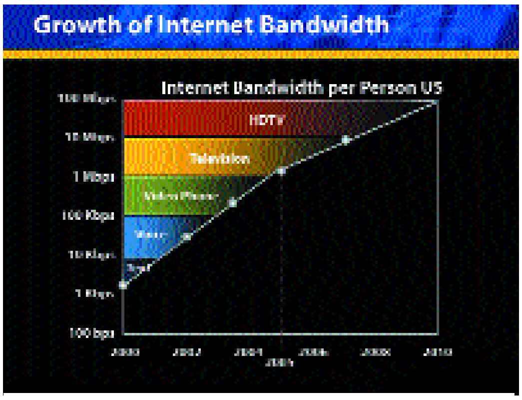
PDC is taking part in the efforts to develop open standards for Grids, tested since 1996 in the GUSTO Computational Grid Testbed

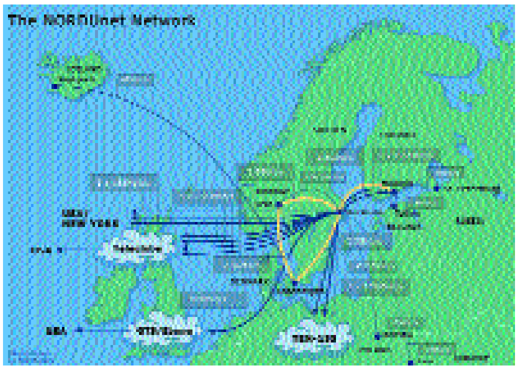
The STARTAP meeting 2001 at PDC

Figure 1. Prediction of bandwidth use over the next decade (courtesy Lawrence Roberts, Caspian Networks).

Figure 2. Nordunet connections to the US and STARTAP.

Figure 3: Starlight, the TeraGrid and the National Transparent Optical network, NTON.





The Science, Technology and Research Transit Access Point, STARTAP, is a project focused on innovations in the network infrastructure, its management and operation, and low level software support for applications required for a successful realization of Grids.

STARTAP provides a persistent infrastructure for interconnection and interoperability of advanced (international) networks in support of application development, performance measurements and technology evaluation. Tom DeFanti is Principal Investigator for STARTAP. He early realized, as indicated in the article on page 6, that several aspects of the computing and communication infrastructure for science and engineering research and education could and should change to create an entirely new information architecture where bandwidth, along with computing and storage, is the enabling, rather than gating, technology.

The focus for STARTAP and similar initiatives is on transforming the network architecture, protocols, operations and management to better serve current and future applications. The ambition is not only to speed up the adoption of new technology in the network infrastructure through (commercially) compelling applications, but also to motivate competition among telecommunications carriers based on open standards in a way similar to the transformation that has taken place in the computer industry and Internet Protocol based products and systems.

The driving vision is that science and engineering increasingly is pursued by teams based on skills, interest, experience and resources regardless of locations. An element of the vision is that the interactive sharing of virtual worlds will significantly enhance the productivity of such teams. Sharing of visual experiences requires broadband networks, but also places stringent quality of service requirements with respect to, for instance, latency and jitter. Figure 1 illustrates the expected dominance of various forms of video related technologies over the coming decade, as predicted by one of the fathers of the Internet, Lawrence Roberts.

The real-time aspect of high-performance networks is also critical in many applications involving control and data acquisition from instruments and sensors. Quality-of-Service must be a pervasive property of the network infrastructure to serve many current and most future applications. New detectors for instruments in many science and engineering disciplines and medicine are revolutionizing experimental data capture and rapidly producing large amounts of data through improved resolution in time and space necessitating scalable and distributed approaches to data analysis and visualization. The infrastructure supporting the correlation of instrument and sensor data with large-scale simulations, analysis, interrogation and interpretation must have high assured capacities and qualities.

The STARTAP effort received its first funding in 1997 from the US National Science Foundation and holds annual meetings. PDC hosted the 2001 meeting. STARTAP, a Next Generation Internet Exchange Point (NGIX), is managed by the University of Illinois at Chicago. Some peering arrangements are through consortia. Thus, for instance, Nordunet is one of four original charter members of Euro-Link. CERN has since joined the Euro-link consortium. Similarly, TransPAC handles the peering for Asia-Pacific Advanced Network

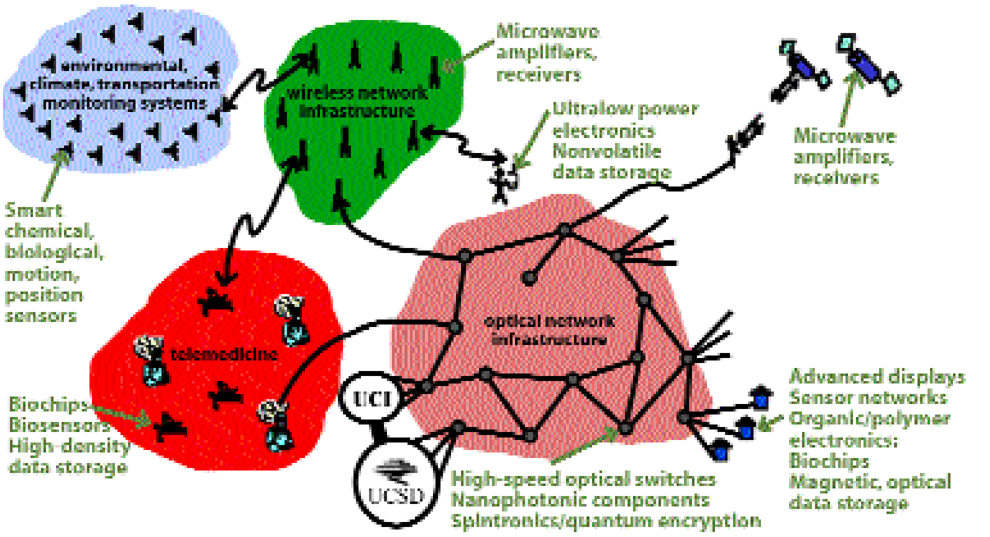


Figure 4:Pervasiveness of Grids as seen by Lary Smarr.

(APAN), and MIRnet handles the peering for Russia. The Network Operations Center's (NOC's) for STARTAP, Euro-Link, TransPAC and MIRnet is handled by Indiana University that has established a Global research Network Operations Center (GNOC). The Nordunet connection to STARTAP is illustrated in Figure 2.

At the 2001 Annual STARTAP meeting several representatives of networks peering with STARTAP, some network operators and a program manager from the US National Science Foundation gave brief presentations. Tom DeFanti described the StarLight project that will establish an optical NGIX in Chicago with initial connections as shown in Figure 3.Initially, Starlight will support four 10 Gbps Dense Wave Division Multiplexing (DWDM) channels from Chicago to other places in the US, with plans for 40 Gbps per wavelength and much higher wavelength counts by 2004. To achieve the goals of STARTAP/Starlight a concentrated effort is required on (a) management of high-bandwidth wide-area connections; (b) creation of wavelength scheduling, allocation, and security middleware; and (c) design of new protocols for managing DWDM networks, including Grid middleware for provisioning bandwidth.

Charlie Catlett, the Chair of the Global Grid Forum and also responsible for the network aspect of the TeraGrid facility, described the Illinois I-wire project, a project deploying DWDM technology between major research institutions in Illinois.

The developments in Holland and the Dutch SURFnet connection to StarLight with a 2.5 Gbps optical circuit since the fall of 2001 was described by Erik Jan Bos.

Bill StArnaud gave a status report on the CA-net, the first all optical national network, and its development. One innovative use of the expected abundance of bandwidth in Wide Area Networks (WAN) that has been pursued by CA-net is the Wavelength Disc Drive (WDD) that exploits the high bandwidth times delay product of continental scale optical networks.

Larry Smarr gave his current view of Grids, one aspect of which is captured in Figure 4. The Figure captures the vision of the integration of sensor and wireless networks into a wired optical communications infrastructure with substantial computation, storage and visualization capabilities.

Maxine Browne gave a brief overview of applications covering both telecollaboration and applications involving data collection from scientific instruments and large-scale simulation and analysis.

In summary, the annual meeting showed that great progress in the networking area is made towards enabling the vision of Grids. The meeting was attended by several of the leaders of the Grid revolution. The STARTAP meeting provided evidence of the i) emergence of data collections as fundamental to scientific research, ii) emergence of community scientific projects, iii) explosive growth of high-resolution scientific instruments, iv) broad deployment of ultra-high-speed networks. ■

…Grids – status and options

➤➤ (… continued from p 1)

The Grid has been described by Bill Johnston as tools, middleware, and services

- providing a uniform look and feel to a wide variety of computing and data resources
- supporting construction,management,and use of widely distributed application systems
- facilitating human collaboration and remote access and operation of scientific and engineering instrumentation systems
- managing and securing the computing and data infrastructure.

Bill Johnston is NASA's Information Power Grid technical project manager and manager of the US Department of Energy's (DoE) Grid effort, the DoE Science Grid.

The "World Wide Grid" concept was conceived in 1993 by Tom DeFanti, Rick Stevens and Larry Smarr. DeFanti at the Electronic Visualization Laboratory at the University of Illinois, Chicago, is perhaps best known for his contributions to computer graphics, the founding of the SIGGraph conferences, and inventing the CAVE virtual reality environment. Stevens, well known for his interest in visualization and human-computer interfaces is a computer scientist at the Mathematical Sciences Division at Argonne National Laboratories, and Division leader. Smarr, well known for his vision of large-scale scientific and engineering computing, is an astrophysicist that at the time served as director for the National Center for Supercomputer Applications (NCSA) at the University of Illinois, Urbana-Champaign.Larry Smarr is now the director of the California Institute for Telecommunications and Information Technology. The motivation for the World Wide Grid concept was the opportunities offered by the rapid technological evolution to carry out distributed visualization and simulation. NCSA's successful Mosaic web browser (the prototype for the Netscape browser) derived from ideas of Tim Berners-Lee at CERN was a great motivation as well.

The occasion driving the conception of the Grid was the expo in conjunction with Supercomputing '95 for which Tom DeFanti had accepted the responsibility. He wanted to show a vision of the future enabled by high-performance computing and communication. The demonstration project, named I-Way, involved 17 different organizations and peering between 11 different computer and telecommunication networks. About 60 applications were successfully demonstrated. The software that tied the various resources together and enabled application programs to run across the network and the assembled compute, storage and visualization resources was developed by Ian Foster's group at Argonne National Laboratories under the name I-Soft. The great success of the I-Way demonstration lead to the now very well known Globus project lead by Ian Foster and Carl Kesselman, an I-Way collaborator at California Institute of Technology in Pasadena.

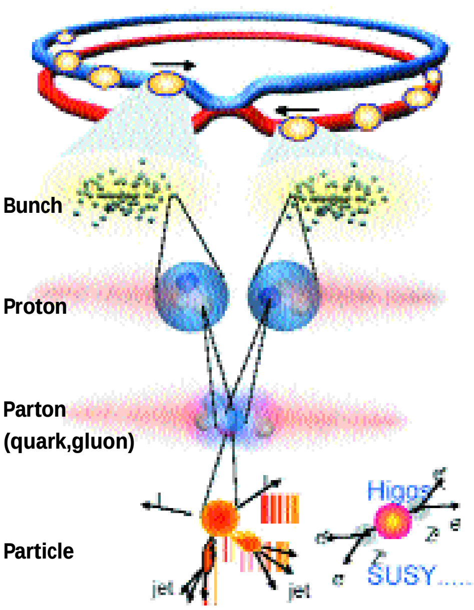
The development of the communication infrastructure is now very rapid and Grid resources are expanding fast. There are now a large number of projects in industry, national laboratories and academia spread across the USA, Canada, Europe and Asia with the aim to provide national or global environments for access to data collected by unique scientific instruments or laboratories, data bases, compute and storage resources. The shifting of the focus towards data is in part a consequence of the world becoming digital. In fact, the investments in data storage grows more rapidly than the investments in compute capacity, and the capabilities of storage technology improves more rapidly than the processing capability. Storage capacity doubles close to every 12 months, while processing capability doubles every 15 – 18 months at constant cost.

Interactive, distributed visualization for collaboration between team members in different locations is an essential element of many of these projects. Demanding applications are present in several areas, such as high energy physics, earth sciences, biomedical imaging and genome research. The challenge in collecting, storing and analyzing measurement data from the Large Hadron Collider, (LHC), at CERN (described on page 7–9) is an illustrating example. Significant amounts of data are produced in steadily increasing numbers of projects both in academic research, industrial applications, commerce and government as indicated in other articles in this Newsletter. ■

EU DataGrid Project

The EU DataGrid is a large-scale project in every respect except for the size of the objects for whose study the project was initiated. This field of physics, which is the scientific pursuit of the fundamental constituents of matter and their interactions, is known as both high energy physics and elementary particle physics. The ultimate goal is to create a unified theory of nature through a unification of forces. The Large Hadron Collider (LHC) is being constructed at CERN in Geneva represents a large step towards that goal by enabling experiments to probe interactions at an energu scale in excess of 1 TeV (10^{12} electron volts). Such energies allow for thorough studies of several pressing questions in fundamental physics, such as issues related to the Higg's particle, which is thought to be responsible for mass, evidence of supersymmetry, and extra dimensions of space-time. The LHC will begin operating in 2007.

The issues have deep connections to the large-scale structure of the universe and imply significant overlap with disciplines such as astrophysics, cosmology, and nuclear physics.



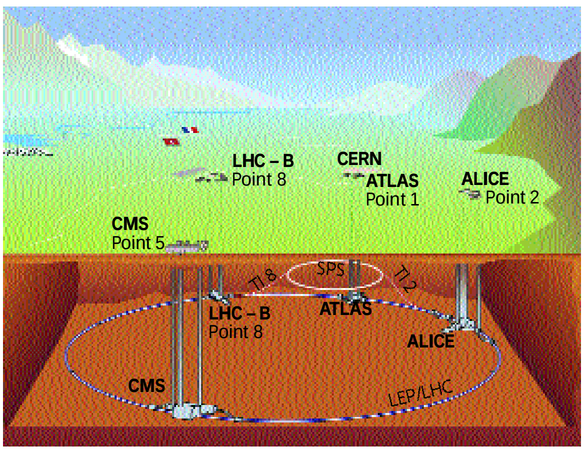
Proton – Proton	**2835 bunch/beam**
Protons/bunch	**10^{11}**
Beam energy	**7 TeV ($7x10^{12}$ ev)**
Luminosity	**10^{34} $cm^{-2}s^{-1}$**
Bunch crossing rate	**40 MHz (every 25 nsec**
Collision rate	**~10^9 Hz**
(Average ~20 Collisions/Crossing/Experiment)	
New physics rate	**~10^{-5} Hz**

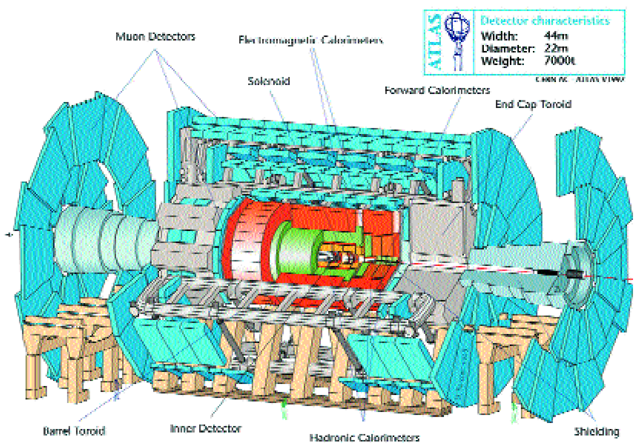
Selection: 1 in 10^{13}

Figure 1 (top): The collision and measurement selection procedure as it will operate in the LHC from 2007 and onwards (Courtesy Paul Avery, University of Florida)

Figure 2 (bottom left): The layout of the different projects at the CERN accelerator

Figure 3 (bottom right): The ATLAS detector (low left). The diameter of the detector is 22 m and its length is 44 m.





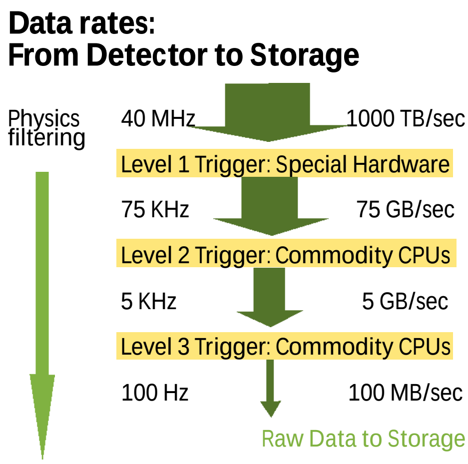


Figure 4 (above): Data reduction at CERN from detector to storage(Courtesy Paul Avery, University of Florida.)

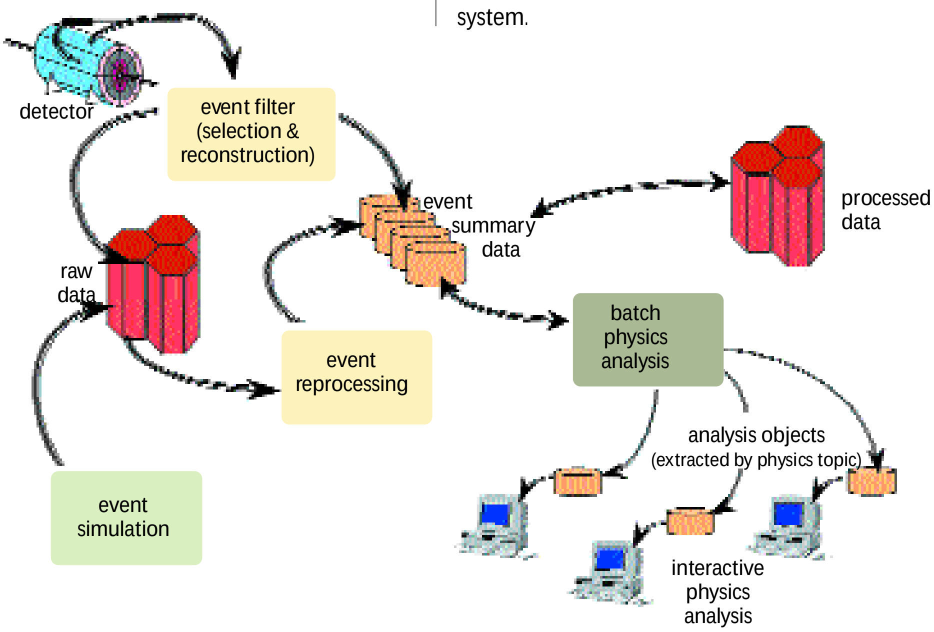
The LHC detectors will capture events that occus at distances down to 10^{-19} m. The events result from beam collisions in which the interesting events are obscured by many uninteresting collisions and the selection process is illustrated in Figure 1. Finding the interesting events are indeed like looking for a needle in a haystack. The ratio of interesting events is 1 in every 10^{13} collisions. The data generation is enormous.

The LHC has four different detectors along the accelerator ring, as shown in Figure 2. The accelerator, which lies partly in France and partly in Switzerland, has a diameter of about 9 km and a circumference of 26.7 km. The scale of the detectors is illustrated by the ATLAS experiment, with Figure 3 showing the physical size, and Figure 4 illustrating the data rates. The majority of work in the Swedish physics community is related to the ATLAS detector. The data saved from an event may be 1.5 MB; with a billion events a year, the archival storage need is estimated at 5 Pbytes in its initial year of operation, 2007. By 2012, about 100 Pbytes of data are expected to be accumulated. The scale of the scientific collaboration is also huge. The ATLAS experiment involves about 1,800 scientists in about 150 research laboratories and universities in 32 countries. The way to handle these large amount of data and the computing power required to analyse them in a world wide distributed manner has large implications also for other fields of science, e.g. bioinformatics.

The EU DataGrid **<http://www.eu-datagrid.org>** and DataTag **<http://www.datatag.org>** projects as well as the UK GridPP **<http://www.gridpp.ac.uk>**, Italian INFNGrid **<http://www.ge.infn.it/com4/grid/grid.pdf>** and similar projects in other European countries together with the US GriPhyN **<http://www.griphyn.org>** and iVDGL **<http://www.ivdgl.org>** and Particle Physics Data Grid, **<http://www.ppdg.net>** projects seek to address the data management, analysis, and access problems for LHC-produced data and a user community throughout the world. The sustained required computational rate is estimated at 100 – 1,000 Tflops.

The system architecture that has been proposed to address these needs is a hierarchical Grid. High data rate filtering and other data-acquisition operations will take place at CERN, where a data reduction of seven orders of magnitude will take place, as illustrated in Figure 4. Most of the data analysis will take place at regional and local centers. The data storage is also distributed. The data management and analysis architecture is illustrated in Figure 5. The Grid software architecture is shown in Figure 6.

Figure 5 (below)The distributed data processing system.



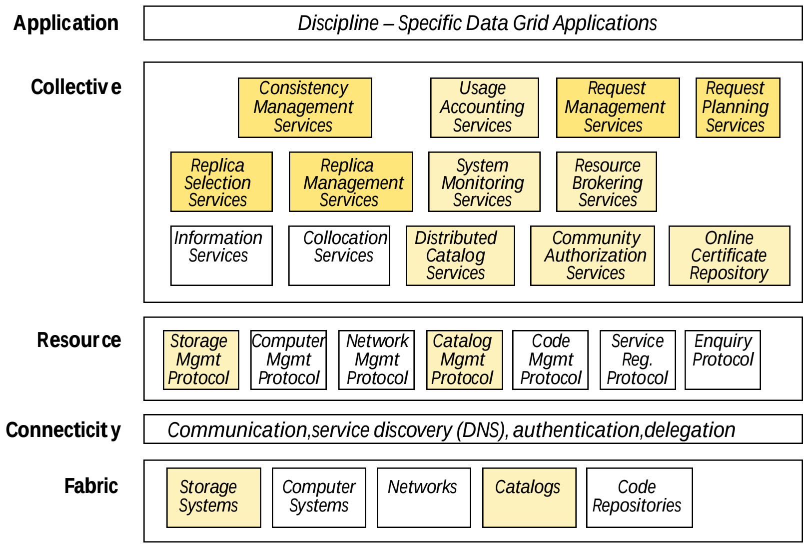
The EU DataGrid project uses the Globus toolkit as a point of departure for its software development, which is organized into 12 work packages. Seven of those address the required infrastructure, three address applications (high-energy physics, earth observation sciences, and biological sciences), one dissemination, and one project management. Two of the seven infrastructure packages address network services and fabric management; three address workload, data, and storage management; one monitoring; and one a testbed for verification and evaluation. Sweden is represented by PDC in the Grid Data Management WP (nr 2) and by Karolinska Institutet in the Biology Science Application WP (nr 10). Figure 7 illustrates the structure of the project and the interactions between the work packages.

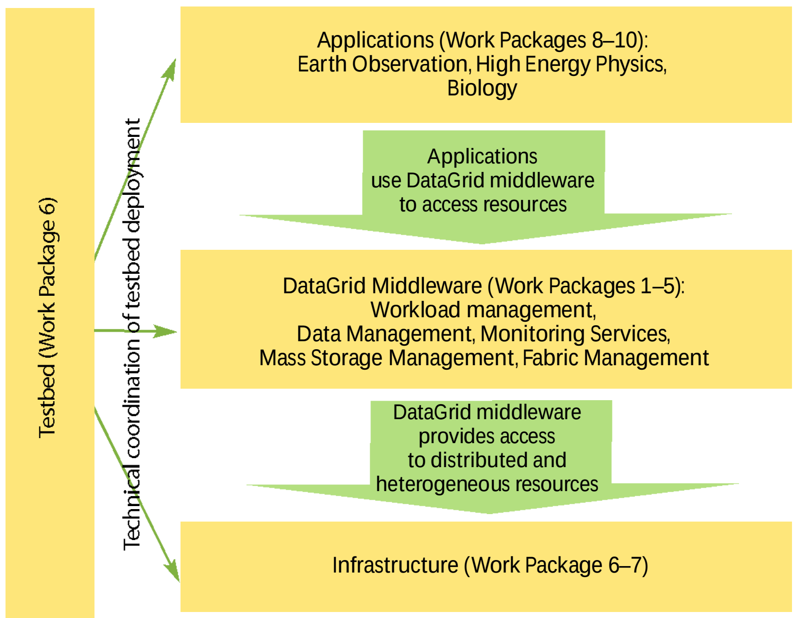
The EU DataGrid is a fascinating project with great potential to develop the tools and infrastructure that will fundamentally change how a number of science and engineering disciplines work. The tools and infrastructure will facilitate formation of virtual communities and remote access to data and computational resources regardless of location of community members, data, or physical resources.

Welcome to my brain

Teresa Wennberg, a well-known pioneer in video art, created the inauguration multi media installation at the VR cube in 1998. She has now created a new installation which illustrates recent findings in brain research. Her work Brainsongs, which is performed at the VR cube until the end of May 2002, is reflecting the activities in the brain called neurogenesis that are fundamental to out daily life and social adaptation. Please call 0708 620 143 if you are interested to visit the VR cube and experience the Brainsongs.

Data Grid Reference Architecture





Organisation of the technical work packages in the DataGrid project

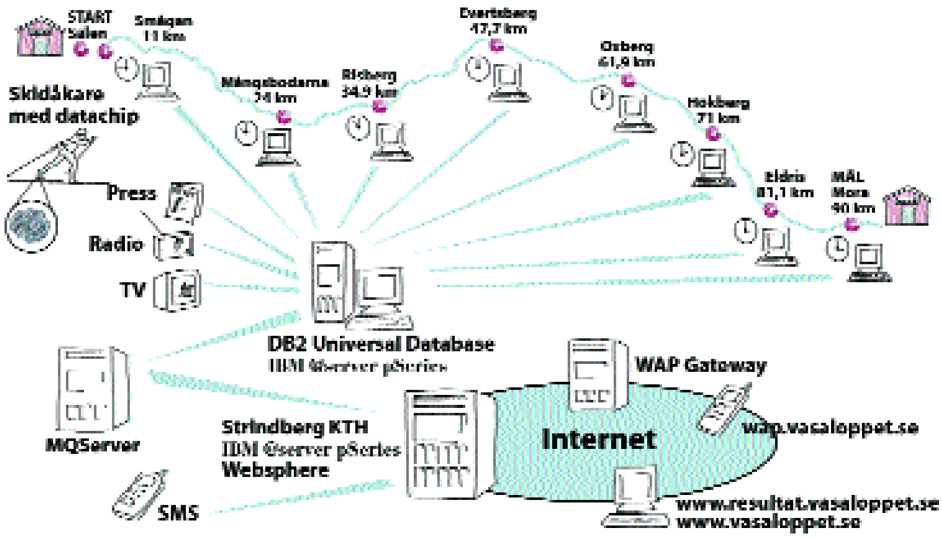
Figure 6 (top): The LHC Grid architecture.

Figure 7 (bottom): The interaction between work packages shown in Figure 7.

PDC on the Grids

An example of an early, simple, PDC Grid service is the storage service provided for data transmitted from the Odin satellite designed to support studies of star formation or early solar systems, depletion of the ozone layer in the earth atmosphere, and the effects of global warming. The Odin satellite is developed by the Swedish Space Corporation (SSC) **<http://www.ssc.se>** on behalf of the Swedish National Space Board (SNSB) and the space agencies of Canada (CSA), Finland (TEKES) and France (CNES). Another example is the remote file service provided to SSF's **<http://www.stratresearch.se>** Large-Scale Genotyping Laboratory at Uppsala University **<http://www.genome.uu.se/genotyping>**. A more complete remote service including computation and data management is provided to the Stockholm Bioinformatics Center, SBC, **<http://www.sbc.su.se>** that involves a number of scientists and graduate students that typically perform many tasks on a cluster at PDC in combination with local desktop calculations.

A service with many elements of what future Grid environments may support is the detailed tracking of all participants in the famous Swedish ski race Vasaloppet. This service has been provided in collaboration with IBM since 1999.



The main features of the track ing system of the rae Vasa-loppet

The famous Swedish ski race Vasaloppet **<http://www.vasaloppet.se>** was supported with detailed tracking of all participants. This made Vasaloppet a high-tech event as well as a major sport event. Wireless transponders communicated the location of the skiers to base stations and via a wired infrastructure furthered the data to PDC's computer Strindberg where not only time and positions were calculated, but optionally also estimated finish time of the skiers, and other data, could be asked for. Data was processed and transferred to the mobile phone operators who offered the services to their SMS and WAP customers. The information was of course also made available on the Internet.

In the main competition on Sunday March 3, almost 15,000 skiers started. During the race almost 90.000 hits per minute were recorded. In the other races during the week some 21,000 men, women and children participated and were also tracked during their skiing. About 68 millions visits were counted during the week of the race and more than 32 millions of them on the main race. This corresponds to an increase of about 50%, which clearly shows that the services were widely appreciated.

PDC is also developing Grid middle-ware in several fields. The EU project Neurogenerator **<http://www.neurogenerator.org>** is aimed at producring a database for the collection and processing of images and other data from neuroscience experiments.

The need for distributed computing and storage is rapidly increasing as exemplified by the Large Hadron Collider, LHC, at CERN, which is described on pages 7–9. PDC contributes in the Grid security area through the EU DataGrid project **<http://www.eu-datagrid.org>** .

Examples of other projects addressing the needs of the LHC project as well as astronomy projects with large data sets are the EU-DataTag project **<http://www.datatag.org>**, the UK GridPP project **<http://www.gridpp. ac.uk>**, the Italian INFN Grid project: **<http://www.ge.infn.it>** and the GriPhyN **<http://www.griphyn.org>**, iVDgL **<http://www.ivdgl.org>** and PPDG projects **<http://www.ppdg.net>** in the US.

The European DataGrid project, like the projects just mentioned, also aims to support other projects generating very large data sets, such as projects in astronomy and imaging projects in the life sciences. Examples of such projects are the UK Astro-Grid project for astronomy research **<http://www.astrogrid.ac.uk>**, the BIRN project for biomedical informatics research **<http://birn.ncrr.nih.gov>**, the NEES project for earthquake engineering simulation **<http://www.eng.nsf.gov/nees>**, the National Ecological Observatory Network, NEON, project **<http://www.sdsc.edu/NEON>** and the National Virtual Observatory, NVO, **<http://www.srl.caltech.edu/nvo>** projects in the US.

Other projects, such as the EU GridLab project **<http://www.gridlab.org>** and the US Globus **<http://www.globus.org>**, Condor **<http://www.cs.wisc.edu/condor>**, GRIDS **<http://www.grids-center.org>** and GrADS **<http://grads.iges.org/grads>** projects address middleware and application development tools.

For a more extensive annotated list of Grid projects, see PDC's web site.

Summer School on "Introduction to High-Performance Computing (HPC)"

PDC in collaboration with the National Graduate School in Scientific Computing (NGSSC) and the Parallel and Scientific Computing Institute (PSCI) organizes intensive two week summer school courses on "Introduction to High-Performance Computing (HPC)". Cost-effective parallel computers has enabled researchers to carry out complex simulations for an ever wider range of applications during the last decade. The emergence of Grids will enable new science. This course will give an introduction to the skills needed to utilize high performance computing and visualization resources, and insights into the use of computational and data Grids. Last years course was arranged in August 2001 on KTH campus with 39 participants. "We are happy to see the course growing, and to see an increase in the initial abilities of the participants," said Per Öster at PDC. "This indicates to us an increase in the interest and use of HPC in upcoming research."

This years summer school will take place August 19–30, 2002. Registration and information at **<http://www.pdc.kth.se/training/2002/SummerSchool>**

README

- **Access PDC** For MS Windows users a new possibility for secure login with Kerberos5 and X-windows is described at **<http://www.pdc.kth.se/support/cygwin-install.html>**
- **IBM SP-NH** If you are running an MPI job with several processes on each Night-Hawk node (N,H or K node), remember to set the environment variable MP_SHARED_MEMORY to *yes.*
- **Numerical libraries** The numerical libraries are continuously updated. Recent updates are PETSc (to 2.1.1) and PSPARSLIB. Let us know your requests for updates or new libraries.
- **Python** Python 2.x is now available on AIX and Linux computers at PDC. In addition several scientific python modules are installed, such as Numeric, numarray, ScientificPython and NetCFD.

B

Sverige
Porto
Betalt

Some glimpses from the history of Grids

Grid computing was originally conceived and developed by and for scientists working in international research projects, similarly to how the World Wide Web emerged from Tim Berners-Lee's work at CERN. Like the development of the web, the move of the Grid from the realm of scientific computing to the offices of companies and to individuals will rely on the development and implementation of open-source software. The challenge is to produce a software that require minimal changes to application codes. This is a middle-ware challenge of great magnitude.

There are complex requirements for managing Grids and a number of resource and connectivity protocols must be set up to facilitate the sharing of individual resources. The challenge is to establish an international interoperability standard for Grid resources and operations in order to make distributed computing and visualization easily available. To facilitate this goal, early developers and users of Grid testbeds and prototypes have engaged in community building efforts that now have merged into the Global Grid Forum **<http://www.gridforum.org>**. The success of the Globus development that has made the Globus toolkit **<http://www.globus.org/toolkit>** the base for the vast majority of Grid projects whether in industry or academia, and the commitment of the Globus leadership to the Open Grid Services Architecture gives rise to optimism for the emergence of open standards for Grids in the not too distant future. PDC is taking part in these efforts in several ways, with a focus on Grid testbeds, such as GUSTO in the Globus project, security and file system aspects of middleware, and applications in the life sciences, fundamental physics, and engineering.

The GUSTO testbed was used when PDC successfully presented calculations of electromagnetic fields around antennas for mobile communication at the Supercomputing '97 conference. The exemplary cooperation of the staff of both Sunet **<http://www.sunet.se>** and Nordunet **<http://www.nordu.net>** contributed to the success. These networks are critical for the viability of any Swedish national and international Grid effort. The good collaboration with Sunet and Nordunet has continued and been critical to the success of PDC's Grid efforts whether for production or development.

PDC has been an active partner in many Grid projects already from the beginning of the Grid development. Co-scheduling of computational resources enabling an authorized user to simultaneously use resources at more than one site in a cooperative manner was performed by HPC2N at Umeå University **<http://www.hpc2n.umu.se>** and PDC in 1997. This effort was an outcome of contacts already in 1996 with the Globus project and its testgrid GUSTO.

Visualization represents an effective tool for analyzing complex issues in research as well as in design. PDC has demonstrated the potential in a number of projects. Alliance98 responded to Caterpillar's **<http://www.pdc.kth.se/projects/alliance98>** wish to analyze how to design large vehicles as a cooperative effort between design teams in different locations, such as for instance their teams in several US locations and Sweden. The JaCo3 platform **<http://www.psci.kth.se/prog.rep./23.pdf>** makes it possible for two competing corporations to cooperate in a joint project without disclosing their own specific knowledge to their partner and competitor. EnViz **<http://www.pdc.kth.se/projects/envis>** demonstrated the flow of gases in a jet engine in parallel virtual environments on both sides of the Atlantic using the Globus infrastructure. ■

PDC Newsletter

Published by Parallelldatorcentrum at KTH.

The center operates leading-edge, high-performance computers as easily accessible national resources. These resources are primarily available for Swedish academic research and education.

PDC receives funding from the Swedish Research Council (VR),and KTH.

Editor: Lennart Johnsson

Editorial staff: Mats Höjeberg

E-mail: pdc-newsletter@pdc.kth.se

ISSN 1401-9671

To contact PDC

Visiting address: Lindstedtsvägen 3, Plan 5, KTH, Stockholm

Mail: PDC, KTH; SE-100 44 Stockholm; SWEDEN

E-mail: pdc-staff@pdc.kth.se

WWW: http://www.pdc.kth.se/

Phone: +46 8 790 78 00

Fax: +46 8 24 77 84

Graphic design: Cecilia Frank, Frank Etc. *Print:* Ekblads Tryckeri, Västervik 2002.