From web to grid

Professor Steve Lloyd discusses the development of a Grid large enough to accommodate future scientific data streams…

The worldwide web was developed at CERN in the early 1990s to help solve the problem of how particle physicists, working in large collaborations distributed across the globe, could efficiently exchange electronic information about their research. Today, particle physicists and many other researchers face another, even more challenging problem. New generations of particle physics experiments, astronomical satellites and telescopes, genome databases, collaborative engineering projects, digitisation of paper archives, etc. are expected to produce huge increases in the amount of data to be stored and processed in the next few years by increasingly dispersed groups of scientists and engineers.

In particle physics, the Large Hadron Collider (LHC) is currently being constructed at CERN and is due to start operation in 2007 to probe fundamental questions, such as the origin of mass in the Universe. The LHC will house four experiments, two with general purpose detectors designed to detect anything and everything produced in the collisions between very high energy protons, and two with more specialised detectors. The general purpose detectors contain over a hundred million individual electronic channels, each producing digital data at a rate of 40 million cycles a second. After selecting potentially interesting events to be analysed in detail later, the total amount of data produced is likely to be several petabytes a year (one petabyte is a million gigabytes), equivalent to about 10 million CDs. This huge volume of data has to be made available for analysis by hundreds of physicists around the world looking for a handful of very rare events. The proposed solution to this and other computational and data intensive problems is the 'Grid'.

The Grid takes its name from the Electricity Grid that provides a ubiquitous supply of electricity through a standard interface (plug and socket) throughout the country (and with suitable conversion across the world). The complexity of the power stations, sub-stations, power lines, etc. is hidden from the end-user, who simply plugs in his appliance. In a computational Grid, the power stations are collections (farms) of computers and data storage centres and the power lines are the fibre optics of the internet. Special software, known as middleware, provides the interfaces through which users can submit their own programs to these computers and access the data stored. The user does not know or care where his program actually runs or where his data is actually held as long as he gets his results back as quickly and reliably as possible. Since the computing resources will have many different owners, economic models need to be established, whereby either, like electricity, real money is charged, or credits are exchanged within 'Virtual Organisations', such as the worldwide particle physics community. Although the component parts, computers, disks, network, etc. have existed for many years, the challenge is to seamlessly integrate thousands of them together into one distributed system, which, to the user, looks very much like one enormous PC.

In the UK, particle physicists have formed a collaboration called GridPP, to develop a UK Grid for particle physics. GridPP consists of 18 UK universities plus CERN and RAL, and is currently funded for three years by PPARC with a budget of £17m. Although the ultimate aim is to develop a Grid for the analysis of the LHC, GridPP decided to try and include as many existing experiments as possible in the prototype stage.

The GridPP project currently has three main areas – building the underlying computing infrastructure, computers, disks, network, etc., developing the middleware that allows transparent use of this infrastructure and developing particle physics applications that will make use of this Grid. The first two areas are generic in the sense that they should allow any applications to use the Grid, not just particle physics ones.

In the organisation of the computing facilities themselves, one can image two extreme models, whereby either all the computers are in one place or they are spread out very widely with little clustering into groups. The first model is sociologically and politically unacceptable, as everyone who contributes to the project would have to put all their investment in one place, eg. CERN, rather than in their own country or institute. The second model would be impossible to manage efficiently. Clustering the resources ensures that they remain under local control, where local policies can be applied. They can be switched in and out of the Grid at will, without breaking the Grid. Clustering at the local level reflects the normal funding mechanisms and divides the hardware into maintainable chunks. In particle physics, it is natural to envisage a hierarchical system with a large computing centre at CERN, where most of the data originates, large 'Regional Centres' in each of the major countries involved, with smaller centres in small countries, large institutes or groups of institutes. Such a hierarchical model is not, however, necessary and in other disciplines, such as biology, probably not desirable.

In the UK, GridPP is developing a prototype of a UK Regional Centre at RAL and four smaller regional centres in Scotland, Northern England, Southern England and London. The UK centre is being developed around an existing robotic data tape store, with new disk servers and computers, while the regional centres are using existing computing facilities. The latter are important because they are being developed in conjunction with local e-Science centres and will provide a focus for local support, interactions with other disciplines and industry, and hopefully allow other resources to be utilised more efficiently, for instance, making use of university student PC systems that might otherwise be idle at night. A vital component in the management of such clusters of machines are so-called 'fabric management' tools that are being developed to allow large numbers of computers to be remotely configured and new versions of software to be automatically installed.

The middleware that controls the operation of the Grid is being developed as part of the European DataGrid Project. The project is funded by the EU and has 21 partner organisations, including PPARC in the UK, with a budget of €9.8m. This middleware can be thought of as the distributed equivalent of the operating system of a single PC, controlling the logging in of users, the management of disk storage resources and the scheduling of the running of the programs. As in a single PC, the operating system is independent of the applications that actually run, allowing this middleware to be used by all other disciplines as required. The software is all Open-Source.

GridPP is concentrating on grid data management, information services, network monitoring and security. Grid Data Management involves using metadata (data about data) to keep track of where files are on the Grid and optimised file replication to make copies of data elsewhere when necessary. Information services are vital in providing the 'Job Brokers' with reliable information as to where best to send a particular program to run, based on the storage and CPU time available, the location of the data that the program needs and the network bandwidth available between different parts of the Grid. Network monitoring is essential to smooth and efficient running of such a widely distributed system. Security is paramount and is based on public key digital certificates. Authentication, ('who am I?'), and authorisation, ('what am I allowed to do?'), are kept separate, with the latter being based on rights inherited from membership of Virtual Organisations. Ultimately, a matrix of international trust relationships need to be established. This was recently tested when particle physics collaborators successfully transferred files between the US and the UK, authenticated by digital certificates issued by the US DOE and the UK Certificate Authority.

The development of particle physics applications to run on this Grid has two main strands; the development of new data reduction and data analysis programs for the future LHC experiments and the adaptation of programs for existing experiments, most of which are currently based in the USA but with UK collaborators. The former test the Grid using large quantities of simulated LHC data, while the latter tests the Grid with real data for real physics analysis. Together, these should ensure that the prototype Grid is able to scale to the size ultimately required by the LHC but also that it is stable and comprehensive enough for real production use today.

Because of the open ended nature of the project and the fact that the developers are distributed across many institutions and are not employed directly by the project management team, new project management tools have been developed. These involve a 'Project Map', which provides a dynamic high level view of the whole project with the ability to drill down though sub-components to the detailed deliverables and milestones.

Whether the Grid really is the next IT revolution will be answered in a few years' time when it will become clear whether the technologies being developed by collaborations such as GridPP become widely adopted outside of data intensive scientific research fields and whether Virtual Organisations common in scientific collaborations can really be transplanted into the world of politics and business. What is already becoming clear is that the Grid has the potential to bring fundamental changes to the way large-scale computing is handled both in academia and in industry. Particle physicists will build a Grid to enable them to analyse data from the LHC and hope that this can be of much wider benefit.

Steve Lloyd
Professor of Experimental Particle Physics

Chair of the GridPP Collaboration Board
Queen Mary University of London
London E1 4NS

Tel: 020 7882 5057
Fax: 020 8981 9465

s.l.lloyd@qmul.ac.uk
www.gridpp.ac.uk