OPINION PIECE

The Large Hadron Collider and Grid computing

BY NEIL GEDDES*

Science and Technology Facilities Council, Rutherford Appleton Laboratory, Harwell Oxford, Didcot OX11 0QX, UK

We present a brief history of the beginnings, development and achievements of the worldwide Large Hadron Collider Computing Grid (wLCG). The wLCG is a huge international endeavour, which is itself embedded within, and directly influences, a much broader computing and information technology landscape. It is often impossible to identify true cause and effect, and they may appear very different from the different perspectives (e.g. information technology industry or academic researcher). This account is no different. It represents a personal view of the developments over the last two decades and is therefore inevitably biased towards those things in which the author has been personally involved.

Keywords: Large Hadron Collider; Grid computing; worldwide Large Hadron Collider Computing Grid; Grid; computing

1. Introduction and background

When the Large Hadron Collider (LHC) project was adopted into the CERN programme (1991) and approved by CERN council (1994), it was recognized that this would bring a major, but at that time not fully quantified, computing challenge. Over the years immediately following the approval, as the requirements and designs of the experiments matured, the full scale of this challenge gradually became clear. At the same time, computing itself was undergoing a transformation. It is easy to forget how different computing was at the beginning of the 1990s and how networking and the World Wide Web, with its inherently distributed and connected view of the world, transformed the way that we view access to computing and information technology resources over the last two decades. The first web server was deployed at CERN in 1990. Thereafter, things developed quickly, for example,

— the first web server in the USA was deployed at the Stanford Linear Accelerator Center in December 1991;
— the first commercial web publication, Global Network Navigator, began in 1993;

*neil.geddes@stfc.ac.uk

One contribution of 15 to a Discussion Meeting Issue ‘Physics at the high-energy frontier: the Large Hadron Collider project’.

This journal is © 2012 The Royal Society
— Yahoo and Amazon were founded in 1994;
— Google formed in 1998;
— 2000 saw the ‘dot com’ crash and first touch-screen smartphone marketed by Ericsson;
— 2001 saw the first Global Grid Forum meeting;
— 2003 saw Apple iTunes launched; and
— 2007 saw Cloud computing emerging and BBC iPlayer and Apple app store launched.

Individual computers gained in power over the same period, with typical CPUs running at 30 MHz in 1990 and 1.5 GHz in 2000; a factor 50 increase in 10 years. The first Top500 supercomputer list was published in June 1993. Top of the list was a computer from Thinking Machines Corporation, rated at 60 GigaFlops. By June 2000, the top of list was the Intel ASCI Red machine, rated at 2.4 TeraFlops, a factor 40 increase in 7 years. Today, the fastest machine is 2.5 PetaFlops.

These rapid developments were repeated across computing and information technology and opened up new opportunities across a wide range of applications. In particular, the ability to access and coordinate distributed remote resources at huge scale was directly relevant to the challenge facing the LHC physicists. The emerging world view in the 1990s was effectively captured in a seminal book edited by Ian Foster and Carl Kesselman and published in 1998 [1]. In ‘The Grid’, Foster and Kesselman argued that the Grid was an emerging infrastructure that would ‘fundamentally change the way we think about—and use—computing’. The book highlighted a range of opportunities and technological developments in distributed computing, and how these could increasingly be exploited in large-scale scientific endeavours. The vision of Grid computing was that users would no longer have to concern themselves with where their files were stored, or where the computer processing their data was located, they would have seamless transparent access to globally distributed resources. This integration would cross organizational boundaries, thereby facilitating new research collaborations. The Foster and Kesselman book struck a chord with many people. Not least with Sir John Taylor, then Director General of the Research Councils in the UK and looking for opportunities to tackle the computing and data-processing demands across a growing number of research fields. The Grid blueprint lay at the heart of the vision behind the United Kingdom’s e-Science programme, which ran from 2001 to 2006 and contributed directly to the LHC computing developments.

2. The beginning

(a) The technical challenge

In the late 1990s, projects such as MONARC began to model some of the challenges related to supporting the worldwide LHC collaborations in an attempt to quantify the computing requirements. In 2000, the LHC experiments peer-review committee (LHCC) launched a wide-ranging project to look in detail at the computing issues. This project, led by Prof. Siggi Bethke from Munich, covered the LHC experiment data-processing models, worldwide collaboration, software and coordination issues.
The LHC computing review reported in March 2001. The review:

— accepted the scale of resource requirements presented by the experiments;
— recommended adoption of a distributed, hierarchical computing model as developed by MONARC:
  (i) Tier-0: at CERN; raw data storage; first pass reconstruction and calibration;
  (ii) Tier-1: regional/supranational centres; perform analysis, Monte Carlo generation, storage;
  (iii) Tier-2: national/intranational centres;
  (iv) Tier-3: institutional facilities; and
  (v) Tier-4: end-user workstations;
— recommended Grid technology to be used (efficient resource usage and rapid turnaround); and
— identified the need for well-supported research networking with links of 1.5–3 Gb s\(^{-1}\) by 2006.

The recommendation to adopt the Grid approach was perhaps not a surprise. As well as gaining traction in the academic field, Grid computing was already becoming more widely known. In June 2000, the Deutsche Bank published a technology report entitled ‘New Economy, Forget the Web, Make way for the Grid’. This report highlighted the business opportunities likely to arise from the growth in networked information technology infrastructure, increasing automation and computer readable documents.

What was the scale of the challenge that the LHCC review acknowledged? It was summarized in the report as:

— four experiments each with a 50–200 Hz data-taking rate;
— raw event sizes ranging from 0.1 to 25 MB, with total raw data of 7 PB yr\(^{-1}\);
— total simulated data storage of 3.2 PB yr\(^{-1}\);
— worldwide tape storage of 28.5 PB yr\(^{-1}\) (40 million CD ROMs);
— worldwide disk storage of 10.4 PB yr\(^{-1}\) (100 000 disks at 100 GB per disk);
— worldwide CPU capacity of 360 000 of today’s (year 2000) PCs; and
— wide area network bandwidth (Tier-0/-1): 1500 Mb s\(^{-1}\) per experiment, 5 Gb s\(^{-1}\) in total.

This should be compared with the state of the art at the end of the Large Electron Positron collider (LEP) era. In 2000, each LEP experiment had access to approximately 2000 ‘CERN units’ of batch computing at CERN, equivalent to roughly 20 fast PCs at that time. In addition, each experiment had roughly 2.5 TB of disk storage. This is already more than 100 times the estimated need at the start of LEP operation (1989), and 1000 times the estimated need at the time of the LEP computing review (Green Book) in 1983.

The 2001 LHCC review also noted that the core software and information technology teams were severely understaffed, that CERN should sponsor the move to object-oriented programming, and that the Tier-0 plus Tier-1 installations were likely to cost 240 million Swiss Francs, plus 80 million Swiss Francs per year to sustain, ignoring the cost of staff.
Along with recognition that the technical solution to LHC computing was not yet available ‘off-the-shelf’, came the acknowledgement that the investment required to realize any solution would be significant, and was not included in the existing plans for the collider or the experiments. In response to the review report, CERN member states initiated the LHC Computing Grid (LCG) as a special CERN project in 2001. This project stimulated a number of contributions from member’s states in order to kick-start the required developments. A single project was formed to avoid unnecessary duplication of work across the LHC experiments. The CERN project collaborated directly with a number of national and international initiatives in Europe and beyond and funded the early years of the LHC Grid developments. In the UK, the GridPP collaboration was formed in 2001. Funded under the UK e-Science programme, GridPP brought together all of the UK experimental particle physics groups and explicitly included CERN as a partner. The GridPP project aimed to build a prototype Grid for particle physics computing in the UK and to contribute directly to the development of the LCG. Between 2001 and 2005, GridPP supported 20 researchers at CERN as core members of the LCG project.

Grid computing beyond the Large Hadron Collider

Grid computing of the form exploited by the LHC started in the USA. The National Aeronautics and Space Administration (NASA) Information Power Grid is often regarded as the first serious attempt to coordinate, from 1998 onwards, large-scale computing across multiple sites in order to provide an integrated service to scientists and engineers. In 1999, the PPDaGrid (Particle Physics Data Grid; PPDG) project started in the USA. This was quickly followed by the International Virtual DataGrid Laboratory (iVDGL), which brought together physicists and computer scientists to develop the tools required to exploit the distributed infrastructures for large-scale scientific endeavour.

Opportunities arising from Grid computing were also recognized at a European level. Not only did the emerging ideas support world-leading scientific endeavour based in Europe, but they also provided a flagship initiative to raise the profile of computing infrastructure investments across Europe; a key theme in the European vision of a knowledge-based economy. The European Commission co-funded a number of Grid deployment research projects during 2000, including the EU-DataGrid (DataGrid, as it came to be called) and CrossGrid projects. These projects included a number of partners from the LHC community as well as from the fields of Earth observation and biomedical science. The aim of the DataGrid project was ‘to build the next generation computing infrastructure providing intensive computation and analysis of shared large-scale databases, from hundreds of TeraBytes to PetaBytes, across widely distributed scientific communities’. The DataGrid project formed the basis of the generic LCG developments from 2001 to 2004.

The goal of the particle physics Grid projects can be summarized as deploying and operating a software and service infrastructure to coordinate the movement, storage and analysis of the extremely large volumes of data expected from the LHC. It was quickly realized that the scale of computing required precluded putting all of this in one place, e.g. at CERN, and with the geographical
distribution of the LHC collaborations, required a truly worldwide infrastructure. The challenge then quickly becomes as much organizational and procedural as it is technological, requiring agreed solutions to user and data management that are widely accepted and will work at the scales required. The nature of the system, spread across many countries and organizations, means that any services have to be developed to cope with an inherent heterogeneity and instability. The latter arises from both unplanned faults and errors and planned but uncoordinated system evolution.

An illustration of the profile of Grid computing during the early years of this century is a 2002 quotation from the then Prime Minister, Tony Blair. He said that the Grid ‘intends to make access to computing power, scientific data repositories and experimental facilities as easy as the Web makes access to information’.


(a) The Grid becomes real

In Europe, the DataGrid project quickly became the focus for initial developments towards LHC computing. The commitment to working towards generic solutions opened up opportunities for support and engagement across a range of agencies. IBM (UK) were full partners in the DataGrid project. Their direct contribution focused on the information systems, but they were equally interested in gaining experience in operating large-scale distributed data centres. In the UK, HP worked with the GridPP collaboration, with similar interests to IBM. However, the breadth of stakeholders made identification of the priority requirements more difficult. This, and the research nature of the early projects, perhaps tended to inflate the ambition of the project and hamper the quality of the software and services delivered. Nevertheless, it became clear during the DataGrid project that it was possible to deliver distributed computing Grids at the scale required by the LHC. In 2004, IBM acknowledged their Grid experience in the information technology support for the Wimbledon Tennis Championship, relying on ‘Linux and Grid computing’ to keep everything running smoothly (http://news.bbc.co.uk/1/hi/technology/3798393.stm).

The DataGrid partners developed the idea of a full production Grid infrastructure in Europe that could be deployed to support the LHC and other scientific fields with similar problems. The European Commission was receptive to the idea. The name, derived from ‘Exploiting Grids for E-sciences in Europe’ (EGEE), was originally suggested as a working title by David Williams from CERN. David suggested that it was not a good name, and someone should come up with a better one soon before we got stuck with it! EGEE lasted 6 years, through three phases. The project was coordinated by CERN, with a commitment that the underlying infrastructure required for LHC (LCG) and other communities (EGEE) would be one and the same in Europe. At essentially, the same time, the iVDGL and PPDG projects in the USA evolved into the Open Science Grid (OSG), explicitly aimed at delivering computing for the LHC and other large-scale science projects. The worldwide nature of the LHC endeavour was recognized explicitly through the evolution of the CERN LCG project into the worldwide LHC Computing Grid project (wLCG), exploiting the underlying infrastructure provided through EGEE and OSG. The wLCG...
was set up as a formal collaboration, based on the organizational models used for the LHC experiments. The requirements for membership were captured in a detailed Memorandum of Understanding (MoU) that specifies the commitments to be made by collaborators, in terms of resources and service levels, and the organizational and governance structures. The model described in the MoU retained the tiered structure originally proposed by MONARC, but merged the regional and national centres into a single Tier-1 structure.

(b) Data challenges

The evolution to EGEE and OSG, offering production Grid services, coincided with the start of large-scale ‘data challenges’ by the LHC experiments. These large-scale tests were recommended by the 2001 review as a means of building confidence that the experiment’s computing models and the infrastructure supporting them would work. The experiments also needed to generate and process increasingly large simulated data (Monte Carlo) samples in order to debug the detectors and software. The data challenges began at a relatively small scale in 2002, but by 2004, all LHC experiments were carrying out major campaigns. These used resources across the world to process up to 100 million simulated events, requiring several hundred thousand processing jobs and generating several PetaBytes of data across up to half a million files. This scale was already significantly above anything achieved during the LEP running. While it proved difficult to engage large numbers of physicists in these system tests, it was invaluable for the computing teams within the experiments to gain experience in the large-scale data-processing activities. The early data challenges typically just about met their technical goals, but not without important lessons being learned about the fault tolerance required in the distributed computing models, the need for automation, and the book-keeping and monitoring challenges at the scale of the LHC data processing. It was also clear that the data challenges succeeded in large part because of ‘Heroes’ prepared to work above and beyond the call of duty to make sure things worked. Not a model that could be sustained during an LHC run of many months or years, with the computing systems required to run at full scale essentially all of the time.

The early data challenges also made some of the technical requirements of the computing systems much clearer. They identified a number of failure points, and pointed to a number of gaps in the Grid software deployed at that stage. Some of these latter issues, e.g. the ability to manage user authorization based on roles, were already under development, illustrating another lesson from this period: that the length of time to develop a reliable software service from a tested idea into a deployed and operational service at large scale was typically 2 years. The large-scale data challenges also contributed to the LHC experiments aborting their attempts to use full commercial object-oriented databases for the primary data storage. The wLCG initiated the collaborative development of a bespoke object relational database solution (the solution, known as POOL, exploits the ROOT analysis tool-kit used for most data analysis at LHC). These and other lessons were captured in a technical design report (TDR) [2] for the wLCG system published in 2005. The overall structure detailed in this report was largely unchanged from that proposed at the beginning of the decade, but the TDR was able to specify the services and software to be deployed. The TDR also
gave improved estimates of the scale of LHC computing requirements. The most notable change being the significant increase in disk space requirements compared with earlier estimates.

In 2004, it was hoped and expected that the LHC would begin operation in 2007. Given the experience of the first half of the decade, the emphasis therefore shifted to focusing on stability in the widely deployed tools and interfaces and to tackling the performance bottlenecks uncovered in the workflow management systems and databases. The power of having simple interfaces to large scalable services was a lesson from the development of the World Wide Web and one that was being picked up effectively by the commercial world with a growing interest in service-oriented architectures, which would lead eventually to the Cloud computing services that are widely available today.

A major advantage of the simple interface is that the service provider can optimize the systems behind it in order to deliver them quickly and efficiently. The tension for high-energy physics is the need to balance the desire for a simple interface against the need to control the behaviour of the underlying system to a much more fine-grained extent than would normally be the case for commercial web-based services. When simulating or reconstructing several hundred million events, the details of the underlying operating systems and maths libraries become important.

(c) Making it all work

Through the first half of the last decade, the various Grid projects contributed significantly to worldwide developments in standards and tools for managing and using large-scale ‘data centres’, and distributed infrastructures. The DataTAG project, a partner project to the DataGrid, contributed significantly to the development of high-speed data links between Grid centres, particularly across the Atlantic. This work contributed to the development of the LHC optical private network (OPN), which today connects all of the Tier-1 centres together over the Internet with, what are effectively, dedicated 10 Gb links. This allows the reliable transfer of the huge volumes of data produced by the LHC. The international standards for Grid information systems also grew out of the DataTAG project. The wLCG Grid developments also drove the standardization of access to mass storage (disk and tape) systems. The storage resource management (SRM) interface became an approved standard of the Global Grid Forum in 2002.

The period from 2004 to 2006 saw a huge amount of work from a large number of dedicated people to understand exactly how to make the wLCG work and how to use it. Just a few of the challenges tackled were:

— to understand and improve the networking to the 100+ Tier-2 sites;
— to cope with the variable quality of individual sites;
— to cope with failures of everything; and
— to automate everything, including the error handling.

Further data challenges by the LHC experiments continued improvements in each of these areas. An excellent example is the improvements in data transfer achieved through detailed tuning of the Internet protocol and improvements in the data servers and mass storage systems. Between 2004 and 2006, the amount of
data successfully moved around the world by the Compact Muon Solenoid (CMS) experiment alone increased gradually from 0.5 to 50 PetaBytes per day; a factor 100 improvement in slightly over 2 years. One lesson learned repeatedly through the data challenges was that, if a system or procedure had not been tested at large scale, it would not work first time at large scale. All dimensions are relevant here: number of systems, sites or users, length of time for an operation, numbers of files or volume of data. Many of the problems encountered are obvious after the fact, but limits and thresholds, particularly in distributed systems, are not always obvious until they are encountered.

Through this period, there were a number of experiment-based initiatives to improve the performance and reliability of Tier-1 and Tier-2 sites. In CMS, specific task forces were launched to help debug and monitor networking connections and to set up physics analysis processing at Tier-2 sites. Sites were required to pass, and maintain, specific automated quality tests to be considered as certified sites. If not certified, sites would not be (and are not) used for CMS production work. A Toroidal LHC Apparatus (ATLAS) experiment created the ‘hammercloud’ test suite. This set of tools moves and processes increasingly large amounts of data repeatedly over a specified set of sites until reaching saturation, or something breaks. Although not completely reproducing chaotic workflows seen during data analysis, these automated tools can quickly be used to stress test and debug ATLAS Grid sites.

Aside from the direct Grid developments, there were notable improvements in our ability to deploy cost effectively and to operate larger and larger scale computing installations. Increased automation and improved monitoring saw installations increase from a few tens of CPUs and disk servers to many hundreds of systems; thousands of systems at the largest of the Tier-1 centres. To cope with this step change in scale, it quickly became clear that the power and cooling systems in the traditional computer centres must be upgraded. The period 2003–2007 saw major refurbishments at all Tier-1, and most Tier-2, centres, with several sites building completely new machine rooms to cope with the power density and cooling demands of modern computer equipment.

If the LHC had indeed started in 2007, the major centres on the wLCG would have been ready to process the data, albeit not with the efficiency hoped for. However, the variation in the infrastructure and the quality of service at different Tier-1 centres and, even more so, at the Tier-2 centres, made it extremely difficult for users to exploit the full power of the Grid. Frequent failures would occur because of small differences between different sites, or incomplete installations. Apart from being frustrating for individual users, this made large-scale processing exercises extremely labour intensive and time consuming. Improving this situation, focusing on stability, fault tolerance and error recovery, became the focus on the period from 2007 onwards.


(a) Ready for data?

With the LHC start-up imminent, the focus of the wLCG was on stability and incremental improvements to the operating procedures and functionality. At the end of 2006, the bulk-computing problem of processing the LHC data prior to
Opinion piece. The LHC and Grid computing

user analysis was believed to be under control, although the performance required for the LHC running at design luminosity was only really achievable for limited periods. Reliability and efficiency of sites was a major issue, with few Tier-1 or Tier-2 sites maintaining the levels specified in the wLCG MoU. It quickly became apparent that many features of the experiment computing models assumed that the worldwide infrastructure would work at close to 100 per cent all of the time. This was simply not the case. Tier-2 centres had very variable performance, often reflecting the very limited manpower available at sites to support the local infrastructure. Networking could not always be relied upon, with a multitude of ‘last-mile’ problems restricting our ability to exploit the full potential of the wide area network links. Even in the Tier-1s it quickly became apparent that there was a level of technical and natural disasters that it was not possible to mitigate cost effectively. It was evident that there would be at least one major Tier-1 outage every few months for reasons out of the control of the wLCG, including, for example: regional power outages, faulty hardware deliveries, network breaks, fires and tornados.

In 2007, the UK GridPP collaboration was successfully funded for a third phase, to take the UK particle physics Grid into full operation for LHC data. Reflecting the growing maturity in some areas of Grid computing, GridPP3 focused tightly on the functionality required for the LHC, increasingly relying for some of the more generic services, such as the authentication infrastructure and accounting, on the emerging UK National Grid Service (NGS), which was providing a collaborative computing infrastructure for non-high-energy physics researchers in the UK. The EGEE project also began to look at how the production and support infrastructure across Europe might be sustained beyond the end of the EGEE project. A design study, funded by the European Commission, was launched in 2008 to develop a blueprint for a sustainable European Grid Infrastructure (EGI) that could be exploited by the LHC and other large-scale research endeavours in Europe. In early 2009, the EGI design study proposed a European structure based on National Grid Initiatives (NGIs), such as the NGS and GridPP in the UK, with pan-European coordination provided through a new organization wholly owned and funded by the NGIs themselves. This structure was successfully put in place through 2009, with the coordinating office, EGI.eu, founded in Amsterdam in February 2010 with over 30 countries as members. This major reorganization of the Grid infrastructure in Europe occurred during the period of the LHC startup and it is a credit to the wLCG centres, which form the major resource centres of the EGI infrastructure, that the European infrastructure continued to operate effectively throughout this period.

(b) Enter the Cloud

In 2007, just as the wLCG, focused on stability, became averse to major changes ahead of the first LHC data, Cloud computing arrived in the public consciousness. Although perhaps primarily a marketing strategy for information technology infrastructure providers, the Cloud has doubtless benefited from the experience of the Grid developments from the last decade, not least in the cost-effective management of large-scale data centres. The Cloud providers focus on relatively simple, though still rather technical, interfaces. The services they provide are attractive to end-users as they put the resource control in the hands
of the users in exchange for their credit card details. The latter is effectively the mechanism used to manage security and scheduling risks. This can give innovative developers (and companies) access to rather powerful resources on demand, without the need to go via a system administrator or allocation committee. Of course, this freedom comes at a price. Cloud computing does not solve the problems tackled by the Grid developments, namely how, cost effectively, to coordinate widely distributed resources owned by a number of partners, how to move and process very large volumes of data, and how to share resources with large numbers of distributed users. The model does, however, provide some interesting opportunities for provision of the underlying infrastructure. Also, many of the technical developments driven and exploited by the Cloud providers, such as fault-tolerant file systems, high-throughput alternatives to databases and massive use of virtualization, look extremely interesting in the context of particle physics computing infrastructure. Many of the tools and techniques deployed by Cloud providers offering software as a service (SaaS) or platform as a service (PaaS) are already exploited by the wLCG centres to provide scalable information and monitoring services such as data catalogues and interfaces to the conditions databases.

(c) A false start

On 10 September 2008, the proton beams were successfully circulated in the main ring of the LHC for the first time. The excitement was palpable and the wLCG was ready for the data. The first collision events were quickly distributed around the world, but this did not represent any real test of the LHC computing systems. Google caught the excitement, with a special logo for 9 September. Nine days later operations were halted owing to a serious fault. A year later, on 20 November 2009, beams were successfully circulated again, with the first recorded proton–proton collisions occurring 3 days later at the injection energy of 450 GeV per beam.

5. First data

Following the 2009 winter shutdown, the LHC was restarted and the beam was ramped up to 3.5 TeV per beam (i.e. half its design energy). The first significant LHC physics run with colliding beams began on 30 March 2010. Coincidentally, this was also the day of the formal opening of the GridPP Tier-1 centre at the Rutherford Appleton Laboratory (RAL). First collisions were observed almost exactly as the official opening started. Triggered by this coincidence, a study of exactly what happened to some of the first data recorded at the LHC was carried out by Brian Davies at RAL and reported in an article in Nature [3]. This followed data recorded by the ATLAS detector on 30 March 2010, during the first physics run. Events recorded by ATLAS were transferred, via the CERN computer centre, to the Tier-1 centre at RAL, for processing and storage. Later, in May, in response to an analysis request from a group in Chicago, several subsets of the collisions from 30 March travelled from Oxfordshire in the UK, via New York, to the University of Chicago in the USA. Here, post-doctoral researcher Antonio Boveia analysed this and similar data stored at other sites, resulting ultimately in a paper on W production, published onto the pre-print server arXiv.org [4] on
Opinion piece. The LHC and Grid computing

24 December 2010. This publication includes collisions from the first day’s run, along with many others. This was not the first, or only, use of the data recorded on that first day, but illustrates the scope of the computing infrastructure supporting the LHC endeavours.

Throughout 2010, the full computing system functioned without major breaks. There were clearly a number of problems and new lessons to be learned, for example, major hardware or software problems at some sites, including data loss and learning when and how to delete data in such a large distributed system. Overall, the system did not impose significant restrictions upon the LHC exploitation beyond those arising from the scale of the operation. Indeed, the wLCG was widely acknowledged to have performed well beyond design specifications in key areas of data transport and processing. A recent survey of the problems and incidents that caused measurable outages revealed that software was not a significant contributor. Rather, there remains a steady level of problems related to power supply, cooling, installation of new equipment and obscure bugs in commercial software. To this list, we can add a number of natural disasters.

During 2010, the general purpose detectors, ATLAS and CMS, each collected samples of proton–proton collision data, corresponding to approximately 1 PetaByte of raw data. Processing this data saw an average of approximately 1 million jobs per day running across all the wLCG sites, using 100 000 CPU-days per day of computer processing. The OPN saw sustained transfer rates of over 1 Gb s$^{-1}$, with peaks of up to 80 Gb s$^{-1}$ during ATLAS reprocessing campaigns. Early data were typically available for analysis at Tier-2 centres within hours of being recorded, and the Tier-2 resources quickly became the largest contributor to data analysis and Monte Carlo simulation.

Throughout the data taking, the Tier-0 and Tier-1 sites, and most of the Tier-2 sites, achieved the level of resources and service agreed in the wLCG MoU.

6. The future

In 2002, the Tier-1 prototype at RAL consisted of 250 CPUs, 10 TB of disk storage and 35 TB of tape storage. In 2010, the typical Tier-1 centre consists of between 10 000 and 20 000 computer cores (CPUs), and between 5 and 10 PetaBytes of disk and tape storage. In total, these systems cover several hundred square metres and require up to 4 MW of electricity to power and cool them. The Tier-1 centres are connected together with a virtual private network of redundant 10 Gb links. This scale of data centre (10 MW/10 PB/20 Gb s$^{-1}$) is very common across a number of domains, probably representing a natural size through constraints and costs associated with the power and cooling requirements. Google and a few other large Internet service providers have now constructed extremely large data centres requiring 50–100 MW of total power. The capital investment required for these centres means that we are not likely to see them within the wLCG any time soon. However, the lessons learned in operating them may well influence future developments.

Looking forward, what changes do we expect? We know that the systems are too complex as viewed by the users, and the interfaces are not yet simple enough. There has been an overdue move away from centralized workload management.
to a more robust distributed model with central planning and monitoring. Additionally, the world has moved on. One thing that is already having an effect is that the network is generally much better than was feared at the beginning of the last decade. It has proved possible to move extremely large volumes of data reliably around the world. This is changing the way in which the LHC experiments are using the resources at their disposal. The ATLAS and CMS experiments are moving towards a model where they place less reliance on carefully managed data placement, to a more dynamic ‘cache’-like use of the distributed storage resources. This automatically ensures that popular or ‘hot’ datasets are widely copied, avoids moving infrequently used datasets and makes more effective use of the distributed resources available.

With the arrival of data and the start of large-scale processing, the critical functionality becomes much clearer. This allows simplification of the various interfaces and optimization of the underlying systems. This should improve both usability and robustness of the overall service. The power of appropriate simple interfaces has been proved many times over in the commercial sector.

The wider computing world has also evolved over the last decade. Cloud computing arrived in 2006, just as the wLCG became focused on stability! The growth in Cloud computing has led to a number of technical innovations in the operation of very large data centres, which are now looking very attractive. Examples include large-scale file systems, tools for parallelizing calculations, alternative databases and improved virtualization. Virtualization has been around for much more than a decade, and underpins the Cloud computing systems. The Cloud boom has driven significant technical innovation in virtualization, from which the LHC is already benefitting. Virtualization offers a means to further insulate the LHC production and analysis software from the variations of the underlying systems. Such variation is inevitable in a distributed system of the scale of the wLCG. Developments are currently underway to exploit virtualization more widely in the wLCG. The ability to deploy bespoke virtual systems is also very attractive to other user communities who want to exploit the organizational and operational experience of the currently deployed Grids, but want to operate their own software and systems wherever possible.

Historically, one of the major problems has been the additional software overhead that virtualization introduces. This has improved significantly through the Cloud evolution, but for computing tasks of the scale of the LHC, another current development potentially makes this virtualization overhead much worse: the trend towards including more and more computing cores on each chip. This is the current mechanism for maintaining the Moore’s Law growth in CPU power. However, as the number of cores grows, the input/output bandwidth and memory per core do not keep pace. Hence making effective and efficient use of all of the cores is increasingly challenging, and often made even worse with virtualization added into the mix. Increasing the number of cores per CPU is clearly the direction of travel for today’s computing industry. It will be essential that LHC software evolves to take account of this trend in the future, although the challenge is much broader than just high-energy physics.

With these developments, and with whatever other innovations emerge, it will be important to retain experience and features central to the operation of the wLCG. These are unlikely to be provided by anyone else. These include the international trust and supporting systems used to authenticate and authorize
users, data and software, the virtual organization infrastructure that allows us to organize users and communities on an international scale, and the operational systems and processes that allow us to maintain and operate the global computing system upon which LHC physicists rely.

I would like to thank all the members of the wLCG collaboration for their hard work, which made the 10 year programme (to date) the success it undoubtedly has been.

References