How to deal with petabytes of data: the LHC Grid project

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2014 Rep. Prog. Phys. 77 065902
(http://iopscience.iop.org/0034-4885/77/6/065902)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 143.106.11.171
This content was downloaded on 06/11/2014 at 19:50

Please note that terms and conditions apply.
How to deal with petabytes of data: the LHC Grid project

D Britton and S L Lloyd

1 School of Physics and Astronomy, University of Glasgow, Glasgow, UK
2 School of Physics and Astronomy, Queen Mary University of London, London, UK

E-mail: david.britton@glasgow.ac.uk and s.l.lloyd@qmul.ac.uk

Received 8 November 2013, revised 11 April 2014
Accepted for publication 16 March 2014
Published 9 June 2014

Abstract

We review the Grid computing system developed by the international community to deal with the petabytes of data coming from the Large Hadron Collider at CERN in Geneva with particular emphasis on the ATLAS experiment and the UK Grid project, GridPP. Although these developments were started over a decade ago, this article explains their continued relevance as part of the ‘Big Data’ problem and how the Grid has been forerunner of today’s cloud computing.

Keywords: Big Data, LHC, Grid project

1. Introduction

‘Big Data’ has become one of those chameleon phrases of the moment that are slightly hard to pin down. It hints at the volume, the variety, the velocity, the veracity and the value of data that is now appearing, as well as our (in)ability to process it with current tools. More tangibly, it recognizes that the data deluge that has developed over the last decade or so is the start of something much bigger. The underlying trends are stark: mankind has produced more data in the last 5 years than in the whole of history; the world’s total data is doubling every 2 years; and projects currently under construction will soon dwarf what is done today [1]. Without a doubt, we are in a digital age, experiencing a digital deluge that demands new approaches and offers in return nothing short of a new paradigm for science and society. The Large Hadron Collider (LHC) Grid Project [2, 3], conceived and developed at the start of the Millennium, was in the vanguard of this action as the warning-shots of the digital century were fired. This article will look at how the LHC Computing Grid was developed as a bespoke solution to the challenge of handling hundreds of petabytes (PB) of data, at a time when the largest integrated datasets were a few terabytes in size. Although the technologies (and their names) will evolve, the underlying challenges and some of the component solutions, discussed below, will remain.

2. Scientific motivation

Particle physics may have been one of the first, but it is far from the only discipline that rides the digital wave. The Sloan Digital Sky Survey [4] began collecting astronomical data in 2000 at the rate of 200 gigabytes (GB) per night and within weeks had amassed more data than in the history of astronomy. Earth-observation satellites generated increasingly large amounts of data and the increasing deployment of large sensor arrays pushed GB to TB in many fields. But the data deluge is much broader than just science. The digitalization of written or pictorial data is creating a machine-accessible corpus from which secondary data can also be generated in many non-science disciplines. The advent of search-engines and social media sites are highly visible examples that hint at enormous underlying data sets, but perhaps less visibly, governments produce, gather and store even larger amounts. When the Obama administration announced the Big Data Research and Development initiative in 2012, data was defined as a ‘transformative new currency for science, engineering, education, commerce and government’ and the scale was underlined by the funding commitment of $200 million.

Much of Big Data, particularly that which does not come directly from scientific endeavours, is still to be fully
exploited. Data-driven discovery has been suggested as the fourth paradigm of science [5]; a 21st century compliment to approaches based on observation and experimentation; analytical or theoretical methods, and computational or simulation driven discovery. Data is no longer a few measurements taken to support some idea, but a vast resource to be mined or refined for content that might not even have been envisaged when the data were collected. The LHC experiments are now custodians of more than 200 PB of raw, derived, and simulated data from which they have successfully extracted a glimpse of the hyper-rare Higgs boson. This process has been likened to trying to find a piece of hay in a haystack; a much harder problem than finding a needle in a haystack because at least with a needle, one knows when one has found it. Individual Higgs events are almost indistinguishable from ordinary events and it is only by examining all the events and looking for tiny excesses at specific, but unknown, masses can the Higgs be seen. It is as if a hundred or so pieces of hay had been cut to the same length and then hidden randomly in a haystack. The only way to find them would be to sort every piece of hay in the stack by length and find the pile that statistically seemed to have a slight excess.

The actual observation of a ‘Higgs-like boson’ in 2012 by the A Toroidal LHC Apparatus (ATLAS) [6] and compact muon solenoid (CMS) [7] experiments was a milestone in a story that started about a quarter of a century earlier with the conception of the LHC and the vast general-purpose detectors. By the early 1990s the projects to build the accelerator and detectors were starting to forge ahead, but in many areas, and not least in the computing, there had to be a leap of faith that technology would develop in time. For example, the superconducting dipoles that fill the 27 km LHC tunnel needed to operate at fields never before achieved on that scale; the silicon tracking detectors at the centre of the detectors that would cover areas larger than a tennis court if laid out flat, had to be scaled up by orders of magnitude from anything that had previously been attempted and made much more resistant to radiation; and the computing requirements were beyond anything imaginable at the time (and originally vastly underestimated). Big science has a history of driving technology and spins offs have often benefited mankind in unexpected ways. The World Wide Web, arising from the desire for information sharing across geographically dispersed collaborations centred at CERN, is the highest profile example, but there are many more. The superconducting magnet technology enabled better medical imaging and the computing Grid prototyped in the early 2000s was used to search for cures for malaria and avian flu and paved the way for cloud computing.

The volume of digital data generated by the LHC is driven by the need to detect extremely rare events, but is constrained by limits on what can actually be handled. The ATLAS detector [8], shown during construction in figure 1, is the larger of the two general-purpose LHC detectors and has 100 million channels that can be read out for each collision, which can happen 40 million times per second. At a PB sec$^{-1}$ it is impossible to read this out, but fortunately, the vast majority of this data are uninteresting (already well understood) processes and can be eliminated by an online trigger system that preferentially selects potentially interesting events. The first level trigger passes only 75 000 events per second and the more thorough, second level trigger passes about 2000 per second to the event filter. Here, the total is whittled down to about 200 interesting events per second that are selected for permanent storage. This corresponds to something like 15 PB of raw data recorded per year at the design luminosity (collision rate). The trigger thresholds are tuned to keep the data rate at a manageable level, but inevitably they cut into the acceptance for some of the interesting physics. There is constant pressure within the experiments to increase the rate. This raw data is only the starting point: it must be duplicated for safekeeping and similar amounts of Monte Carlo simulated data are produced in order to understand the detectors’ detailed responses to known physics processes and new theoretical...
ideas. Many levels of derived data are also created. By the end of 2012, the ATLAS experiment had accumulated around 10 PB of raw data, but had passed the 100 PB mark for total data on disk with half as much again on tape. The CMS experiment [9] has generated a similar amount of data.

There is a further dimension to the complexity of the LHC experiment that arises from the globally distributed nature of the collaborations, which, as we show later, has influenced the technical development of the computational Grid. ATLAS, for example, is a collaboration of around 3000 scientific authors from 174 institutions in 38 countries. Funding is typically country-based.

3. Motivation for Grid computing

Experimental particle physics moved from analogue to digital in the mid 1970s when electronic data read out from spark chambers started to replace the photograph produced by bubble chamber experiments. Around the same time, the precursor to today’s internet, ARPANET that had been developed by the US military, started to be used by other researchers, mainly for electronic mail. Analysis of scientific data was carried out on large mainframe computers usually sited at major research laboratories. During the 1980s and 1990s, workstations and then Unix servers started to replace mainframes while the C++ programming language began to replace the FORTRAN code in the computer software. Data started to be routinely moved around the internet in increasing volumes. Motivated by the desire of scientists to share information in larger collaborations, the internet became a global network with the invention of the World Wide Web in 1989 by British scientist Tim Berners-Lee and Belgian engineer Robert Cailliau at CERN. By the turn of the decade, scientists had powerful desktop computers accessing various sized clusters of servers at universities and at national and international laboratories. In general the type of computing done by particle physicists is known as ‘embarrassingly parallel’ (or ‘pleasingly parallel’) because each individual particle collision (event) is independent of any other and can be processed in parallel on a separate computer if necessary. Using clusters of servers to process many events in this way is now generally known as ‘high throughput computing’ (HTC).

In principle, the analysis of LHC data could have been carried out on one gigantic computer cluster sited at or near CERN. This may well have been a technically simpler solution. However, there were several sociological and political reasons to look for a more distributed solution. Although national funding agencies could probably have been persuaded to finance such a centralized centre, as do the accelerators or satellites for example, it is unlikely that individual research institutes or universities would have contributed. Distributing the computing meant that additional resources could be added and other communities engaged, which made scaling up and replacing the resources much more manageable.

In parallel with the development of HTC was the increasing capability of supercomputers used for ‘high performance computing’ (HPC). These are used for massively parallel computations where different parts of the problem are intimately related, such as weather forecasting. Eventually, some problems, such as global climate modelling, were too large to be solved on individual supercomputers and clusters of supercomputers started to be used, linked together by state-of-the-art networks. This culminated in the emergence of the computing Grid paradigm linking together both HTC and HPC resources as espoused by Ian Foster and Carl Kesselman in their book, The Grid: Blueprint for a New Computing Infrastructure [10].

Foster and Kesselman’s vision for a computing Grid was a computing infrastructure that is dependable, consistent, pervasive and inexpensive. The Grid took its name from the electricity grid where a number of different electricity producers (gas, coal, wind, nuclear, etc) are linked together by a series of pylons, substations, transformers, etc, to provide the end user with a standard (at least within one country) voltage and frequency through a standard socket, as shown in figure 2. Similarly, the computing Grid links various distributed computer resources, data services, CPU farms, etc, through the internet, providing standard protocols for submitting the programs to be executed and receiving the output.

The Grid paradigm struck immediate resonance with the particle physics community starting to build the computing infrastructure required to analyse forthcoming data from the LHC. A number of projects were initiated that eventually led to the Grids we have today.

In Europe, the DataGrid Project [11] was funded by the European Commission from 2001 to 2004 to set up test beds and develop so-called ‘middleware’. Middleware is the ‘operating system’ of the Grid as described later. This was succeeded by the Enabling Grids for E-Science in Europe (EGEE) project from 2004 to 2010 with three phases, EGEE-I, EGEE-II and EGEE-III by which time the name had changed to Enabling Grids for E-scienceE, but the acronym remained the same. In 2009 the European Grid Initiative (now European Grid Infrastructure) (EGI) was established to connect together and build upon National Grid Initiatives (NGIs).

In addition, in 2001, five Nordic institutes set up a project called the ‘Nordic Testbed for Wide Area Computing and Data Handling’ known as NorduGrid [12] using a different set of middleware, the Advanced Resource Connector (ARC) [13].

In the USA, the Particle Physics Data Grid (PPDG) collaboration was formed in 1999, to develop, acquire and deliver Grid-enabled tools for data-intensive requirements of particle and nuclear physics. In 2004 this became part of the Open Science Grid (OSG) [14].

In 2005, CERN proposed the setting up of the LHC Computing Grid (LCG) that became the Worldwide LHC Computing Grid (WLCG) in 2006 [15]. The WLCG draws on Grid resources from the European EGI, NorduGrid and the OSG in the USA. By 2012 it had become the world’s largest computing Grid, comprising over 170 computing facilities in 36 countries.

Other related initiatives include the Partnership for Advanced Computing in Europe (PRACE) [16] and TeraGrid, now the Extreme Science and Engineering Discovery Environment (XSEDE) [17], linking supercomputer sites in the USA.
4. Challenges in the Grid era

The concept for a globally distributed computing Grid to solve the LHC data challenge was an alignment between the developing concept of Grid computing, the imminent challenge of the LHC that would be hard to meet by extrapolating conventional solutions, and the global nature of the LHC collaborations. However, the challenges were formidable. To allow users from all over the globe access to a computational resource, the issues of authentication and authorization needed to be addressed. The traditional method of issuing accounts and passwords for each user on each system would clearly not scale. Authentication means proving that a user is who they say they are; but on a Grid the user should not need to know which resources are going to be used and so it is the task to be executed or ‘job’ itself that needs to carry the authentication mechanism, or obtain it from some other service. Authorization represents a finer level of control that defines roles or actions that an authenticated user is allowed (or not allowed, to perform). Authentication is analogous to a passport that identifies you, while authorization is like the visas inside that say where you are allowed to go. The use of global resources required the more general development of international security policies to inform acceptable use policies and the responsibilities of resource providers. At the highest level, these policies need to be consistent with national policies and laws. For example, the use of computational resources in some countries is not permitted by users from certain other countries. The regulatory issues extend to issues of data protection, freedom of information, and data preservation. In short, whose laws apply? Certainly those of the country wherein the resources are located, but what about those of the owner and, possibly separate, the funder?

Once the authentication and authorization problems are addressed there is a raft of other technical challenges that need to be solved. How is resource-use accounted across a globally distributed infrastructure that is shared by different user-groups? Although no point-of-use charges are anticipated, the implementation of agreed shares require that usage be measured. How are resources added and removed from the

---

The analogy between the electricity grid (top) and the computing Grid (bottom). Image credits: (1) Deut; (2), (3), (5), (7) and (8) Ivor Branton, Brian Robert Marshall, Mat Fascione and Argonne National Laboratory, respectively (CC BY-SA 2.0); (4) Tony Boon (Gnu General Public Licence); (6) Chef Pepín; (7) Chris Dag (CC BY); (9) NASA; (10), (12) BigRiz and Stig, respectively (CC BY-SA 3.0); (11) © Cisco Systems (GFDL).
Grid? There needs to be a method by which resources can advertise their availability and capabilities. In turn, this needs to be linked to some kind of job-submission infrastructure that knows where data is located, what resources are available, what the job requirements are, and whether the data can be moved or duplicated. The job submission mechanism needs, in turn, to log where jobs are sent and where the output will be located. The Grid user should be insulated from much of the unnecessary detail. For example, the user should refer to data by a logical name and not by a physical URL. To efficiency, the Grid may contain multiple copies of popular datasets and the job submission mechanism should translate the logical file name to the most appropriate physical file name.

To a large extent these technical challenges can be addressed by middleware, but additional challenges lay in the deployment of the infrastructure, the scaling of many things up to the number of sites involved, and the network infrastructure that connects it all together. Error reporting and understanding job failures is another demanding area: if a job fails, how does the user understand whether the job was at fault or something in the infrastructure? In the end, the design of the Grid must be fault-tolerant in a way analogous to the internet itself: resources can appear or disappear without notice and simply re-trying the job, preferably automatically, may often be the best first step.

5. Building a Grid

5.1. Toolkits

To enable a computing Grid to link together distributed heterogeneous resources in a transparent manner dedicated software is required, known as ‘middleware’. The middleware acts like the operating system on a normal PC, where it sits between the applications (Web browser, text editor, spreadsheet, etc) and the physical hardware (hard drive, network interface, CPU, etc). On a PC, a user wanting to work on a particular document does not need to know which physical bits of the hard drive the file resides on; the operating system handles this for them. Similarly, on the Grid a user does not need to know where his dataset is located or what computing resource his program is executing on. The Grid middleware that hides this complexity is assembled from various toolkits into a coherent system that has to scale to the size and architecture of the physical infrastructure.

The middleware used by the WLCG is based on the Globus Toolkit [18] and was first developed by the European DataGrid project and the Virtual Data Toolkit then hardened and extended by the subsequent EGEE projects and successors. During the EGEE years, the middleware was known as gLite, but was subsequently taken over by the European Middleware Initiative (EMI) and maintained as part of the EMI software stack. The EMI software stack includes gLite, ARC and UNICORE, used on supercomputers as an alternative to Globus. The Globus Toolkit provides the underlying layer of the software stack and includes components for security, information infrastructure, resource management, data management, communication, fault detection, and portability. These components can be used either independently or together to develop applications that sit on top of them as shown in figure 3.

A key element of Globus is the use of X.509 digital certificates for authorization and authentication. The certificates are issued by a national certification authority (CA) and are used to uniquely identify all users. Additional attributes, such as the Virtual Organization Membership Service (VOMS) [19] can be added to the certificates to provide information about what the user can do on the Grid usually related to their membership of a particular virtual organization (VO). A VO is a group of users working on a common problem such as an experiment like ATLAS or CMS or a community such as Bioinformatics. Membership of a VO gives users access to its software repositories and the authorization to use that VO’s share of the global Grid resources. A number of services are built on top of the underlying toolkits that can be used by the clients and must interface with the user applications.
Figure 4. A schematic view of the WLCG Tier structure. Raw data from the experiments is selected and processed by online triggers and offline computing farms and transferred to the CERN Tier-0 Centre for archival storage and further processing. The data is then transferred to Tier-1s, Tier-2s, Tier-3s and individual workstations for further processing storage and analysis.

5.2. Architecture

The LHC Computing Grid was originally conceived as a hierarchical structure of tier centres based around a single Tier-0 at CERN, a dozen or so large national Tier-1 Centres in different countries, and many 10s of regional Tier-2 facilities, as shown in figure 4. This architecture was driven by the assumption that network bandwidths of \(622 \text{ Mbits s}^{-1}\) when the LHC started would prove to be the bottleneck. In practice, network bandwidths grew much faster than assumed and are now typically \(10 \text{ Gbits s}^{-1}\). Thus, the strict hierarchy, which reflects the various steps of data-reduction, was unnecessary. Currently, as discussed below, the computing models of the LHC experiments are being modified to make use of networks as the third resource, alongside computational and storage resources: allowing duplication or migration of data across the network enables more efficient use of computational resources.

This hierarchical model became known as MONARC (Models of Networked Analysis at Regional Centres for LHC Experiments) [20] and later became known as the cloud model (not to be confused with today’s cloud computing that is similar but not the same) whereby all the resources in a geographical area, such as the UK or France belong to the UK or French ‘Regional Cloud’. A regional cloud normally constitutes a Tier-1 and several associated Tier-2s and Tier-3s, etc.

5.3. Computing services

A Grid ‘job’ consists of the program to be executed, the necessary input files and a script specified using a formal language, such as the job definition language (JDL), based on the Condor ClassAd language [21], that specifies input and output files and the requirements of the job in terms of memory, software, execution time, etc. A job might be anything from the simulation of molecular docking, the calculation of a bank’s risk position, or the reconstruction of a number of LHC collision events. A workload management system (WMS) [22] accepts each job and matches it to suitable site for execution where the necessary resources are available. The job is transmitted to the site and, once the job has finished, any output is saved for retrieval by the user. A gatekeeper machine known as a computing element (CE) at the site accepts the incoming jobs and schedules them to run on local ‘worker nodes’, the individual hardware components, using the local job scheduling system. The CE is a core service, whose presence distinguishes a Grid enabled cluster from a conventional cluster. Small output files are transmitted back through the WMS, while larger data files may be written to a storage element (SE) and catalogued. The WMS is an example of a ‘push’ system whereby users submit all the jobs to the Grid. Although the large LHC experiments such as ATLAS and CMS used the WMS in the early days they have moved towards their own ‘pull’ systems where the jobs to be submitted are held in a queue locally and small ‘pilot’ jobs are submitted to the Grid. Once a pilot job starts to execute it pulls the next real job off the local queue and runs that. This arrangement gives the experiments much more control over the scheduling of the jobs and their relative priorities, but leads to additional challenges, such as the issues of respecting local scheduling policies, traceability, and potentially authorization, since it circumvents local job scheduling systems.

---

7 Although the word cloud is widely used in both contexts we use ‘computing clouds’ and ‘regional clouds’ to distinguish them.
5.4. Data management

Storage on the Grid is provided by SEs that provide an interface to back-end disk or tape pools. The SEs provide a standard storage resource management (SRM) [23] interface to the storage that allows many different types of storage to be utilized in a transparent way. A number of different SEs are in use, from the small scale disk pool manager (DPM), through the medium scale dCache [24], to large-scale systems such as CASTOR [25]. The Tier-2s mostly use DPM [26] or dCache while the Tier-1s have mostly dCache or CASTOR. The SEs allow long-term storage of large files for users. The original specification has been extended by the use of storage classes and space tokens that allow the experiments to dynamically allocate space across multiple sites. Data distribution between the different sites is performed using the File Transfer Service (FTS) [27]. This service allows file transfers to be queued and provides prioritization and retry mechanisms in the case of failure.

5.5. Distributed databases and catalogues

In order to keep track of the files stored on the SEs, they have to be catalogued. Important data is usually replicated at least at two separate sites for safety and this needs to be recorded. This is achieved by the LCG File Catalogue (LFC) [28] that maintains a mapping between logical file names and physical replicas on the SEs. There used to be a number of separate file catalogue services, but the experiments have been consolidating them for ease of maintenance albeit at the risk of creating a single point of failure. In addition, the experiments maintain a number of other databases including conditions data from the LHC and detector run parameters. There are also metadata catalogues that provide data about the contents of the logical files in the LFC, such as important parameters, timestamps, event formats, etc.

5.6. Other services

In addition to those described above there are other important services such as information services that advertise what resources and services are available where. Although use of the WLCG is free at the point of delivery there are accounting services that keep track of who has used what resources so that for instance, fair share allocations can be respected over time.

5.7. The interface between the experiments and the Grid

One of the great challenges has been the interface between the experiment software and the Grid middleware. This, in retrospect, is a typical ‘boundary problem’ where aspirations and expectations were too high and funding for adequate effort was difficult to secure. As a consequence, there is an uneven interface between the experiment software and the Grid middleware across the experiments, with some duplication of work where common solutions would have been more appropriate. On the one hand, this was a pragmatic response to the need to have a full-scale functional infrastructure in time to handle the LHC data and is a natural and inevitable part of the process of building up something completely new. Over time, evolution ensures the survival of best of breed and there is now some evidence for convergence. On the other hand, there were inflated expectations as to what the Grid middleware would deliver and there was not enough collaboration between the large experiments earlier enough in the process to develop common solutions. These ‘early-adopter’ problems have become easier for the later users, who have typically been able to pick and choose parts of the solutions developed by the LHC experiments.

6. Operating the Grid

The various Grid sites (Tier-0, Tier-1 and Tier-2) are operated semi-autonomously within the framework of the WLCG and national infrastructures. The Tier-0 is operated by CERN, the Tier-1s by National Laboratories and Tier-2s mostly by universities. Some Tier-2s, such as the UK’s, are federated in the sense that they involve several separate sites. In the UK there is a Tier-1 at RAL and 20 institutes contribute to four regional Tier-2s, LondonGrid, NorthGrid, ScotGrid and SouthGrid, most of which support all the major VOs. Similarly, for example, in Spain there is one Tier-1 providing for ATLAS, CMS and LHCb plus eight sites contributing to three separate federated Tier-2s for ATLAS, CMS and LHCb. There is a distributed Nordic Tier-1 and associated Tier-2s in four countries serving ALICE, ATLAS and CMS. In the USA there are separate Tier-1s for ATLAS and CMS and many Tier-2s supporting one or more experiments. In fact the Tier-0 is also federated in some sense following the recent extension to the Wigner Research Centre for Physics in Budapest.

6.1. Fabric

Each Grid site manages a large collection of computers and storage systems, typically rack mounted multicore CPUs and RAID disk servers plus associated switches, power supplies, etc. The Tier-1 centres also have robotic tape servers that require specialized storage tools such as dCache or CASTOR. Installing and regularly upgrading the necessary software manually is labour intensive, so configuration management tools such as the CERN Quattor system [29] or Puppet [30] automate these services. This ensures that everything from the operating system to the middleware and experiment-specific physics libraries are correctly installed and published to the overall Grid information systems.

6.2. Networking

A dedicated high-bandwidth optical fibre network called the LHOCOPN (LHC Optical Private Network) [31] working at 10 gigabits per second connects CERN to each of the 11 major Tier-1 centres around the world, as shown in figure 5. Almost all of the centres have at least two networks path to allow operation to continue in the event of a fibre breakage. USLHCNet [32] provides connections between the US Tier-1s and transatlantic links to CERN as well as Tier-1s elsewhere in Europe and Asia. Links between other Tier-1s and Tier-2s use the normal research networks such as GÉANT [33] in Europe and ESnet [34] and Internet2 [35] in the USA.
Figure 5. A topological view of the LHC Optical Private Network, LHCOPN showing the logical connections between CERN and the Tier-1 Centres. The coloured boxes indicate which LHC experiments are supported at each Tier-1 centre.

6.3. Monitoring

A very important aspect of day-to-day Grid operations is monitoring, which is carried out at several levels. Local site monitoring covers cluster load, resource utilization, network bandwidth monitoring and fault condition monitoring. The most common software used to monitor a cluster is Ganglia [36], which allows data from clients to be collected on a master node and displayed via a web server. Sites also use monitoring specific to their own particular batch system that provides a graphical way to monitor the occupancy of the cluster, the job shares, and the efficiencies for each user. The state of the network is obviously very important and various tools are used to monitor the switches in and around their clusters.

As well as local site monitoring, several countries such as the UK carry out national monitoring and testing. At a national level, the network monitoring is of particular interest to measure the matrix of bandwidths and latencies between sites to and from CERN and other international sites using tools such as perfSONAR [37]. In the UK, relevant WLCG tests coupled with dedicated UK tests are collated and made available to all sites. Nagios [38] is used as a basis for a service availability monitoring service that queries a central database and Grid information services to create a list of sites and systems to be tested. The services offered are tested and the results of the tests are sent to the EGI Regional Operations Dashboard where they are collated.

Globally, there are many such monitoring pages that look at all the EGI, or focus particularly on the WLCG Grid. The central database registers details of all sites, their production and certification status and the contacts and services hosted there. From the various monitoring statistics EGI and WLCG provide summaries of availability (fractional uptime) and reliability (fractional uptime excluding scheduled downtimes) figures for all sites compared with agreed metrics, such as 97% availability and reliability for Tier-1s.

In addition to the Grid monitoring, the major VOs provide their own experiment monitoring that enables them to track their own jobs and provide summaries of CPU usage, failure rates, disk and tape storage utilization, etc. Examples are the LHC experiment Dashboards, ATLAS Distributed Data Management accounting, the Panda Analysis Dashboard, etc.

In fact, there is so much monitoring available it is sometimes difficult for System Administrators to see the wood for the trees and many sites have set up their own ‘Dashboards’ to federate information from multiple sources and display an overall status of their site. An example is the RAL Tier-1 Dashboard that provides an overview of the UK Tier-1 status [39].

6.4. Grid operations

As a concrete example of how the Grid is operated, we take GridPP [40, 41] in the UK. The 20 individual Tier-2 sites each have 1–2 system administrators. Most of these plus some of the Tier-1 staff form a so-called ‘Operations Team’ to ‘run’ the UK Grid under the leadership of a full-time Operations Manager. In addition to day-to-day activities, such as maintaining the hardware and software, the Operations Team has defined a number of ‘core tasks’ that are undertaken on behalf of all UK sites.

The Operations Team is responsible for the deployment and operation of core services such as the WMS and top-level Grid information servers. These are essential for the
Grid to operate and for improved resilience they are hosted at several sites in addition to the Tier-1. It also provides deputy security coordinators to support, and provide cover for, the GridPP Operational Security Officer. This helps ensure a well-coordinated and rapid response to security incidents, minimizing exposure.

The Grid is based on well-defined middleware components that are periodically updated with new functionality, security patches and bug fixes. Although verified at the developer-level, the UK contributes to the worldwide rollout process that checks interoperability of components and requires live site testing and feedback of problem to the developers. The Operations Team coordinates the certification of all releases in the UK before rolling them out to all UK sites. They need to remain aware of, and contribute to, developments in other national and international Grid projects and coordinate how they work with those other Grids knowing what is required in the wider context.

To ensure the high quality of UK resources, the Grid is instrumented with extensive monitoring and alarm systems as described above. An on-duty team responds to incidents by raising problem-tickets that inform sites of issues and track their resolution. The Operations Team coordinates this activity and helps site system administrators deploy these tools, customize them and interpret information relevant to their own sites. Problem or help tickets may be raised by users via their VO(s), from the on-duty team as described above, or from site-administrators experiencing problems with the middleware or infrastructure. Tickets may need to be forwarded through several interfaced ticketing systems to arrive at the appropriate expert and, once there, additional information is often required. The Operations Team ensures that tickets are followed-up by facilitating communication; providing advice on the next step; and ensuring that tickets do not stagnate in the system. In addition, the Operations Team helps the UK sites to publish the correct accounting information by leading benchmarking activities and performing checks on the published data.

GridPP’s infrastructure is used by many non-LHC (smaller) VOs. The Operations Team is responsible for tracking usage requests and ensuring that utilization is as smooth as possible. Where problems are encountered by the VO or its users, the Operations Team will provide support and guidance. Good documentation improves the efficiency with which day-to-day tasks are undertaken and is essential for long-term stability of the project. The Operations Team ensures that installation and deployment instructions applicable to the UK are well maintained.

In addition to the core tasks there is a GridPP Storage Group that investigates and advises on the evolution of Grid storage technologies.

7. Using the Grid

The experiment collaborations use the Grid for a wide range of tasks, but the use-cases can be categorized, at a high level, into two types: structured and chaotic. Structured work refers to well-defined processes that are run, typically, by a small number of experts. These include processing the real data from its raw form to a form that is used for physics analysis, and the generation of Monte Carlo simulated data that is used to understand the detector performance. These tasks are often known as ‘production’. Chaotic workflows are typically bespoke analysis jobs, run by a large number of different users. The tag ‘chaotic’ reflects the fact that these are not centrally managed in the same way as the structured ‘production’ work.

As discussed above, users wishing to use Grid resources require both authentication and authorization. Technically these are provided by a Grid Certificate and membership of the appropriate VO. The former is an X509 digital identity certificate issued by an appropriate CA providing authentication of the user. However, somebody else must decide whether to accept the certificate or not, which implies that there must be a ‘chain of trust’, such that the acceptor trusts the CA who issued it. A CA, therefore, will only issue a certificate upon verification of the identity of the applicant (usually by somebody it already knows and trusts). The certificate acceptors (e.g., Grid services), in turn, maintain lists of CAs that they trust. Certificates normally need to be renewed annually. The second requirement, membership of the appropriate VO, provides an authorization step, defining what the authenticated user is actually allowed to do on the Grid. Technically, individual users are assigned roles within the VO and when a job is submitted to the Grid, the Virtual Organization Management System (VOMS) appends an authorization field to a short-lived proxy certificate, which defines the actions the job is able to perform. Typical roles are the authorization to submit production jobs, which usually have higher priority than normal user jobs, or to access particular storage resources.

As discussed above, there is both a push and a pull model for Grid job submission, with the larger LHC experiments moving away from the push model based on the WMS in recent years. The pull model has proved to be more performant, but the push model is still typically used by small non-LHC VOs and is illustrated in figure 6. Here, a Grid job consisting typically of a bit of code that describes the name of the executable, the logical name of the input data, any specific job requirements, and the location for output, is submitted to the WMS along with the user’s proxy certificate. The WMS locates input data by translating the logical file name via an appropriate service such as the LFC and identifies available computational resources. The job is then queued and eventually submitted and run, with the status at each step updated in a so-called Logging and Bookkeeping database. The user consults the latter database to find out when and where the output is available. The input files are bundled together in the so-called ‘Input Sandbox’ while the output files are returned in the ‘Output Sandbox’.

As discussed earlier, the boundary between the experiment applications and the Grid middleware is a difficult and complex area. Software tools are required that handle the job submission, monitoring, data-movement and bookkeeping. These utility programmes need to speak both the Grid language and interface with the experiment-specific computing environment. Some of those used by the ATLAS experiment are described below.
7.1. Computing model evolution

The original MONARC computing model was a push model, designed for the network capabilities available at the time with central organization of systems and services while deployments, optimizations and complex trouble shooting were handled at the regional cloud level. In most cases a regional cloud is a country or a collection of nearby sites. Driven by the higher network bandwidths available, by the middle of 2011 a new model had evolved with much less hierarchy, as shown in figure 7. Any site can use any other site as a source of data. There is dynamic data caching whereby analysis sites can pull datasets from other sites ‘on demand’. Production activity utilizes multi-cloud Tier-2s and user analysis jobs can be split beyond regional cloud boundaries. The network has become a resource like CPU and disk and data is no longer necessarily considered to be at a particular site, but is becoming ‘federated’ across many sites. Jobs are able to seamlessly access data that is held locally or remotely using direct access protocols. There is no longer geographical isolation at the expense of increased reliance on network performance. Optimization of resources is becoming more complex as there is a trade off between transferring a whole file that might subsequently be read several times and just reading part of the file remotely. The former requires good network bandwidth (throughput) whilst the latter requires good latency (response time).
7.2. The ATLAS computing model

The current ATLAS distributed computing model is shown in figure 8. At the heart of this is the PanDA (Production ANd Distributed Analysis) production and distributed analysis system that controls the flow of jobs to and from the Grid, and the Distributed Data Management (DDM) that handles the replication and placement of the data on the Grid. This is augmented by the ATLAS specific front-end interfaces, Ganga and pAthena that allow users and production teams to easily interact with the system. The major components are described in more detail below.

7.2.1. PanDA. The Production ANd Distributed Analysis (PanDA) system [42] was developed by ATLAS, on top of the middleware tools described above, to handle the distributed processing of ATLAS production (generation of simulated Monte Carlo and reconstruction of real and simulated events) and analysis jobs. PanDA provides tight integration of data management with the processing workflow requiring a few experts to operate it, rather than many users moving their own data. It proved itself as a scalable and reliable system capable of handling very large workflows. PanDA throughput has been rising continuously over the years. In 2009, the PanDA processing rate was typically about 50 k jobs/day and 14k CPU wall-time h/day for production at 100 sites around the world, and 3–5 k jobs/day for analysis. Three years later, the number grew by at least an order of magnitude and PanDA is processing up to a million jobs per day with 100k to 200k jobs active at any given time. The PanDA analysis user community numbers over 800, about 300 of whom are heavy/frequent users. PanDA uses a push model for data placement combined with a pilot based pull job submission model.

7.2.2. Distributed data management. DDM [43] manages the accounting of all ATLAS file-based experiment and user data. It manages data distribution across sites, ensuring the distribution is according to the computing model. DDM was designed to meet the scalability, robustness and flexibility needed by ATLAS to manage the complete data flow. While the underlying middleware operates at the file level, DDM revolves around the concept of a ‘dataset’, which is a collection of files that share some common attributes. Each file has a logical file name, for easy human use, and a Grid unique identifier, or GUID in short, that is used internally to identify the file. The GUID of each file is assigned when it is stored for the first time on the Grid. There may be multiple replicas of the same file stored at different SEs. As well as providing redundancy in case of file loss this enables popular datasets to be accessed more efficiently by removing bottlenecks.

7.2.3. DQ2. DDM is implemented through a sophisticated set of tools known as Don Quijote2 (DQ2) [44] to handle the tens of PB of experiment data per year, distributed globally via the WLCG and its replication between the Tier-0, Tier-1 and Tier-2s. Sites can subscribe to receive a replica of a particular dataset they require or replicas can be centrally distributed. Management of these subscriptions is handled centrally by ATLAS computing operations and data management experts. There also exists a separate suite of DQ2 ‘end-user’ tools that allow users to inquire which datasets exists and to copy data files to or from one of the Grid SEs to their local computers. These tools are being replaced with more sophisticated ones, such as Rucio [43], as part of the preparations for the next LHC runs.

7.2.4. PanDA dynamic data placement (PD2P). For user analysis the initial computing model for data distribution was rather static with data being ‘placed’ at sites and PanDA jobs being pulled to where the data were located. This proved to be rather inefficient as some data were more popular than
others and hence the distribution of jobs over the Grid was non-optimal. After a few months of LHC data taking, an intelligent subsystem of PanDA was developed and deployed to interact with DDM and automatically replicate data to reduce job waiting times, saving storage space and making better use of replicated data. In addition the regional cloud model where data flowed from CERN to a Tier-1 to its associated Tier-2s was relaxed to support copying from/to Tier-2s across regional clouds in the system, as shown in figure 7 (right). PD2P is still being optimized in light of experience during recent LHC data taking.

7.2.5. CERN virtual machine file system. The CERN Virtual Machine File System (CVMFS) [45] was originally developed for use within a packaged virtual machine capable of being installed on Tier-3s, desktops and user laptops etc. It has evolved into the favoured mechanism for providing worker node access to experiment (and other) software. Rather than the experiments having to install their software at all participating Grid sites, CVMFS with its hierarchical caching, allows the sites to download and maintain the software from a central repository. A job trying to use a different version of the software from normal will automatically trigger CVMFS to install that version from the nearest cached version.

7.2.6. Ganga. Ganga [46] was developed by the ATLAS and LHCb experiments to facilitate user job submission to the Grid. It allows users to configure their jobs and to test them on local batch systems before using them in large-scale processing on Grid resources. Although Ganga has built in knowledge of the ATLAS and LHCb software frameworks it can be used to submit any type of job and is increasingly used by smaller VOs who do not have the manpower to develop their own bespoke solutions.

7.2.7. pAthena. pAthena [47] allows users to submit ATLAS jobs to the Grid through PanDA. It provides a simple and straightforward command line interface that has almost the same syntax as the ATLAS analysis software, Athena, and it basically wraps it up adding some more information in order to prepare the job for a Grid runtime environment. Once the jobs have been created by pAthena, they are sent to the PanDA server and they will end up in a task queue and will be eventually picked up when a pilot asks for new jobs.

7.2.8. ATLAS metadata interface. The ATLAS Metadata Interface (AMI) [48] provides a catalogue of all ATLAS datasets and their stored metadata. This allows users to find out what datasets match certain criteria for instance, all datasets corresponding to a particular trigger stream, in a particular time period in a certain event format. The list of dataset names can then be given to Ganga or the DDM tools to access the data itself.

7.3. Other experiments

The CMS computing model [49] has similarly evolved from push to pull with a global task queue and a local task queue from which jobs are pulled by worker nodes. The hierarchical data model has been replaced by a full mesh connecting every Tier-2 to every Tier-1 and to other Tier-2s as well. CMS use PhEDEx [50] as their data transport mechanism.

LHCb have a somewhat simpler computing model than ATLAS and CMS whereby nearly all data analysis takes place at Tier-1s and Tier-2s are mainly used for Monte Carlo simulated data production. This is expected to change in the future with more use being made of Tier-2s for analysis. LHCb use DIRAC (Distributed Infrastructure with Remote Agent Control) for their distributed computing activities [51].

ALICE has used a pull mechanism for their distributed computing from the outset. ALICE uses an architecturally different middleware known as the `Alice Environment’ or AliEn [52] to present the user with a seamless interface that can join together the different Grid systems.

8. Current status

The WLCG has successfully transitioned from a test bed in 2004, through pre-production around 2008 to a fully functioning Grid operating routinely at a scale and complexity well beyond that imagined at outset. There are currently around 170 sites in 36 countries providing WLCG with a total of about 400 000 CPU cores, 260 PB of disk and 217 PB of tape. Figure 9 shows the number of jobs run per month since 2004 and peaks at about 36 million corresponding to just over a million jobs per day. The corresponding CPU time is shown in figure 10 and saturates at about 60 million hours a month. The CPU efficiency, defined as CPU time divided by real (wall clock) time is shown in figure 11. The overall efficiency showed wild fluctuations in the early days as different versions of the middleware and experiments’ software were deployed and disk access parameters were optimized, but has recently settled down around 90%. Figure 12 shows the growth in ATLAS storage usage since 2011 and has now reached 140 PB.

This is split approximately 20%/50%/30% across Tier-0, Tier-1s and Tier-2s. The throughput is currently between 5 and 10 GB s⁻¹ across the whole Grid.

Figures 9 and 10 show that ATLAS submits more jobs and uses more CPU than the other experiments. ATLAS has almost twice as many channels in the detector as CMS, and so need more CPU for simulation and reconstruction. They also need more detailed simulation/modelling of hadronic shower models in the forward calorimeters because of a finer grained detector. In reconstruction, the Inner Detector tracking dominates, and again ATLAS has more digits.

9. The future

The LCG is a bespoke solution to the first Big Data challenge and has handled hundreds of PB of data at rates far beyond the design goals. By any measure it has been spectacularly successful. When the discovery of a Higgs Boson was announced by CERN on 4 July 2012, some of the data included in that result had only been collected two weeks earlier and had been reconstructed, verified, filtered, and analysed within a week. This is far quicker than, for example, the data processing
for the previous generation of experiments that only had to handle miniscule amounts of data by comparison. However, the LCG is a complex infrastructure that has evolved over the last decade, shaped by rapidly developing technology, changing requirements, and the accumulation of experience. It will continue to evolve adiabatically for the same reasons as the LHC ramps up in energy and luminosity.

The LHC has been in the vanguard of the Big Data movement and the staged LHC upgrades proposed over the next decade will continue to ramp up the data rate and challenge the computing infrastructure. However, other new scientific endeavours will come on-line in the next few years with more ambitious goals. The Square Kilometre Array [53], an enormous array of radio telescopes, will feed up to 10 PB of data per day as input to a 100 Petaflop/sec supercomputer for real-time processing. This is larger than the combination of the top 10 supercomputers in existence today. Many other fields are starting to generate, or are planning to generate, vast quantities of data through arrays of sensors that monitor anything from the oceans to the road networks, or through the digitization of existing material. Big instruments such as Earth observations satellites are another source, as are supercomputers themselves when used to generate data via simulations rather than simply analyse data. The biomedical fields such as genomics and proteomics are already tackling Big Data problems and the Web itself is an enormous source of data through, for example, social media sites such as Facebook that ingests 500 TB of data per day.

The facilities required for processing large datasets depends on the nature of the data and the nature of the analysis. The LHC data are a series of event collisions that are largely independent from one another and thus can be processed in
parallel with no interconnection between processors, but this not true of other data such as that from the Square Kilometre Array where observations from the same source in the sky have to be correlated across several instruments.

For many years, the Moore’s law increase in computing power was achieved by increasing the clock-speed so that more operations were preformed per second. However, limitations on power and cooling have halted this progress and over the last 5 years increases in computing power have come from the move to multi- and then many-core processors. Most recently, the use of co-processors such as graphical processor units (GPUs) or the Intel Xeon Phi, is being offered as the way forward. Unfortunately, the advantages offered by these more recent trends do not come for free because the application code must be modified to make efficient use of the new hardware architectures. This is a challenging problem for legacy code: For example, the ATLAS collaboration has 5 million lines of C++ that is used to process the data. The experiments are starting to adapt their code to use multi-core processors more efficiently and have started to make, mostly opportunistic, use of HPC resources for special, CPU intensive, short jobs that can be started or stopped at will.

So for handling Big Data sets, one size does not fit all. It depends on the nature of the data, the nature of the analysis and the details of the hardware. The explosive growth in mobile devices based on the ARM (Acorn RISC Machine) processor suggests that this technology might become a competitor for providing low-cost and low-power computing, challenging the X86 architecture used in much of the existing distributed computing. Energy and performance considerations are also driving storage technology towards more complex architectures. Hierarchical storage arrays with fast, but expensive, solid-state devices (SSDs) at the top level and zero-energy disks, which spin down when not in use, at the bottom level are being developed. The migration of data between the top cache-like layer and the bottom archive-like

Figure 11. The CPU efficiency defined as CPU time divided by real elapsed (wall-clock) time in hours for jobs run on the Grid per month by the four LHC experiments, ALICE, ATLAS, CMS and LHCb from 2004 to 2013.

Figure 12. The amount of storage (disk plus tape) used on the Grid by ATLAS in TB from 2011 to 2013.
layer must be automated, but matched to the problem in hand.

The future of the LHC Computing Grid is likely to embrace academic, research and possibly commercial computing cloud offerings. Cloud computing has evolved, or perhaps devolved, from Grid computing and allows on-demand provision of utility resources (i.e. the user pays for what is used). Cloud computing works by having a large pool of shared resources, a charging mechanism, and a user interface. However, it is extremely unlikely that the majority of LHC computing would all be done on computing cloud resources for the same reason: one does not rent a car if one wants to use it everyday: it is cheaper to buy one’s own car. What is more likely is that academic computing clouds will be offered by, for example, universities, based on shared institute resources. Commercial cloud offerings could be used to handle peak loads (called ‘bursting to the cloud’) where the economic arguments are sounder. Various computing clouds are being tested by the LHC experiments, mostly for CPU intensive work that does not have large storage or input/output requirements. The use of virtualization is now becoming widespread on the LHC Computing Grid and this will facilitate the use of computing clouds. The LHC computing Grid, therefore, is likely to evolve into a Grid containing computing clouds.

10. Conclusion

The discovery of a Higgs particle in 2012 marked the culmination of a quarter of a century of work to conceive, design, fund, build, commission and operate perhaps the greatest scientific experiment ever performed. This audacious undertaking required many leaps of faith along the way, assumptions that, with investment and hard work, technology could be pushed far beyond its existing limits at the time. Even as construction of the Large Hadron Collider was underway, no real technical solution to the computing challenge existed and no realistic financial provision had been established. However, over the last 12 years an infrastructure was developed based on the concept of Grid computing, enabled by the dramatically falling costs of hardware. The fact that the LHC Computing Grid was created by a loosely coordinated global effort built on the strength of a common goal, is testament to the success of scientific collaboration across national boundaries and political ideologies, and to the cooperation of funding agencies. The end result is the largest scientific computing Grid in the world, which was delivered on time and which outperforms all expectations. Physicists working on the LHC are able to access and analyse the huge amounts of data in a timely manner and scientific results are generated much faster than previously, an almost inconceivable situation given the step-change in the volume and complexity of the data. In July 2012, Rolf Heuer, Director General of CERN, commenting on the discovery of a particle consistent with the Higgs boson announced ‘It’s been a global effort, a global success. It has only been possible because of the extraordinary achievements of the accelerators, experiments and the Grid computing’.

Acknowledgments

We wish to thank all our collaborators in GridPP and ATLAS and anyone else who knowingly or unknowingly helped us, especially Pete Gronbech.

References

[17] XSEDE The transition from TeraGrid to XSEDE. www.xsede.org/
documents/10157/169907/TG-XSEDE-transition.pdf
[27] Stewart G A and McCance G 2006 Grid Data Management: Reliable File Transfer Services’ Performance (Mumbai, India: CHEP06)
[33] GÉANT website. www.geant.net
[34] ESnet website. www.es.net
[35] Internet2 website. www.internet2.edu/