Current Grid operation and future role of the Grid

Smirnova, Oxana

Published in:
[Host publication title missing]

DOI:
10.1088/1742-6596/396/4/042055

2012

Current Grid operation and future role of the Grid

2012 J. Phys.: Conf. Ser. 396 042055

(http://iopscience.iop.org/1742-6596/396/4/042055)

Current Grid operation and future role of the Grid

O Smirnova[1,2]

[1] Lund University, Particle Physics, Department of Physics, Box 118, SE 22100 Lund, Sweden
[2] NeIC, NordForsk, Stensberggata 25, 0170 Oslo, Norway

E-mail: oxana.smirnova@hep.lu.se

Abstract. Grid-like technologies and approaches became an integral part of HEP experiments. Some other scientific communities also use similar technologies for data-intensive computations. The distinct feature of Grid computing is the ability to federate heterogeneous resources of different ownership into a seamless infrastructure, accessible via a single log-on. Like other infrastructures of similar nature, Grid functioning requires not only technologically sound basis, but also reliable operation procedures, monitoring and accounting. The two aspects, technological and operational, are closely related: weaker is the technology, more burden is on operations, and other way around. As of today, Grid technologies are still evolving: at CERN alone, every LHC experiment uses an own Grid-like system. This inevitably creates a heavy load on operations. Infrastructure maintenance, monitoring and incident response are done on several levels, from local system administrators to large international organisations, involving massive human effort worldwide. The necessity to commit substantial resources is one of the obstacles faced by smaller research communities when moving computing to the Grid. Moreover, most current Grid solutions were developed under significant influence of HEP use cases, and thus need additional effort to adapt them to other applications. Reluctance of many non-HEP researchers to use Grid negatively affects the outlook for national Grid organisations, which strive to provide multi-science services. We started from the situation where Grid organisations were fused with HEP laboratories and national HEP research programmes; we hope to move towards the world where Grid will ultimately reach the status of generic public computing and storage service provider and permanent national and international Grid infrastructures will be established. How far will we be able to advance along this path, depends on us. If no standardisation and convergence efforts will take place, Grid will become limited to HEP; if however the current multitude of Grid-like systems will converge to a generic, modular and extensible solution, Grid will become true to its name.

1. Introduction

This paper addresses present and future status of Grid infrastructures in a very specific context of High Energy Physics (HEP) research. HEP scientific communities in general and LHC experiments [1] in particular are rightfully credited as active promoters, developers and key users of Grid technologies. Grid infrastructures of today are integral parts of many HEP experiments, and they are often fused with generic research infrastructures, which is readily seen in, for example, the European Grid Infrastructure, EGI [2] – the largest single Grid infrastructure as of today.

The main reason for using Grid approaches is the necessity to federate disparate computing and storage resources from a large number of administrative and geopolitical domains. Such resources are typically funded through a variety of national projects, which do not necessarily

Journal of Physics: Conference Series **396** (2012) 042055 doi:10.1088/1742-6596/396/4/042055

prioritise cross-border resource sharing and Grid technology requirements. Local policies always have precedence over global ones.

Still, a characteristic feature of modern Grid infrastructures is large degree of homogeneity in terms of underlying system software and middleware. A quick glance at EGI resources reveals that by far most contributing facilities use Scientific Linux [3] distribution, since it is tailored for HEP communities needs and large number of HEP applications are supported only on this distribution. Likewise, by far dominant computing middleware services are of gLite [4] flavour. This homogeneity, while simplifying maintenance, may obscure technical deficiencies and result in a customer lock-in.

It is important to note that not everything called Grid actually uses Grid technologies. Until appearance of Cloud computing, every distributed computing or data solution was referred to as Grid. For the purposes of this paper, Grid is defined as a federation of heterogeneous conventional systems, enabled by fast networks and a middleware layer that provides single sign-on and delegation of access rights through common interfaces for basic services.

A distributed computing and storage infrastructure built following this definition is not necessarily optimal for all research communities, but it certainly proved to be vital for HEP.

2. Grid operation

2.1. Infrastructures for HEP

In the HEP context, Worldwide LHC Computing Grid (WLCG) [5] is the most relevant computing and storage infrastructure. Resources contributed to WLCG are provided by a very large number of agencies, and offer quite disparate service levels.

Historically, network bandwidth was considered the key service for data-intensive computing, and the MONARC [6] model called for a hierarchical organisation of the infrastructure. In this model, well-connected nodes serve as first tier (Tier-1), and nodes with higher latency constitute second tier (Tier-2). Data to be processed were expected to be transferred in a cascade manner from higher to lower hierarchy sites. A similar model is adopted by other research communities.

While MONARC model strives to cope with slow networks, ironically, the original Grid concept [7] was based on the assumption that wide area networks are becoming as fast as local connectivity, allowing for a truly distributed system.

After years of rapid development of network technologies, MONARC-style hierarchy is less pronounced, and at times becomes an obstacle, when a transfer between e.g. two Tier-2 sites is faster than a transfer via Tier-1, but is forbidden operationally. Currently, the main difference between Tier-1 and Tier-2 sites is not capacity, connectivity or technology, but the agreed service levels, where Tier-1 is expected to be operational essentially round-the-clock. Tier-1 centres still tend to enjoy better network connectivity, as really fast network can not be afforded by most Tier-2s.

A parallel hierarchy is imposed by the fact that most resources are provided by national Grid infrastructures. This results in a network of National Grid Initiatives (NGI) and Regional Operation Centres (ROC), typical of EGI. Coupling between NGIs and ROCs on one side and Tier-1s and Tier-2s on another is rather lose, as technical choices do not always coincide with geopolitical ones: Tier-2 centres sometimes prefer to be affiliated with Tier-1 centres belonging to a different NGI or ROC.

As a result, a single traditional computing centre committing its resources to WLCG becomes entangled in a rather complex set of infrastructures. A modest Linux cluster can be at the same time a member of a generic national computing infrastructure, national Grid infrastructure, regional e-Science infrastructure, continental Grid infrastructure and a worldwide Grid infrastructure. Each such infrastructure is not fully uniform: it brings together different contributors, different users and applications, different middlewares, different policies and priorities and even different funding cycles. These infrastructures always overlap, to a different extent, complicating service levels.

While in a conventional world a routine hardware failure at a University cluster only affects users at this University, in a Grid world it will affect a large number of infrastructures and their users. These users are oblivious to all the Grid complexity, and always expect top service levels.

Another challenge faced by such a resource centre is the necessity to deploy a large variety of Grid services. These include a set of generic services, such as compute and storage elements, file transfer and indexing services, user databases and authorisation services and information services. Moreover, there can be a large number of community-specific services, depending on how many HEP experiments a site has pledged to support. There is very little commonality between all these services, and operating them requires very substantial expertise and human effort, which may not have been available at the site prior to joining the Grid.

Despite all such challenges, as of today, WLCG successfully federated more than 140 resource centres worldwide, which provide HEP scientists with more than 250 000 CPU cores and more than 150 PB of disk storage. This infrastructure can process 1.5 million computational jobs a day, and while resource usage pattern is not quite uniform, it is equivalent to 150 000 CPU cores being used continuously. In general, WLCG usage can be characterised as continuous and steadily growing, thanks to the excellent performance of the LHC. Facing further increase of data, WLCG recently launched re-assessment of technological approaches, including operational aspects. Results of this assessment are likely to affect not just WLCG, but many other multi-science Grid infrastructures.

2.2. Multi-science infrastructures

Grid user communities are very many, but few are as large as HEP to afford large amount of dedicated resources. Smaller communities have to make use of existing multi-science infrastructures.

In Europe, EGI is the key e-Science infrastructure. It organises all Grid-enabled resources in a centralised hierarchy, rooted at the EGI.eu foundation, which governs operations relying on ROCs and NGIs. EGI has a major worldwide impact, and is actively engaged in collaboration with Asian, African and Latin American Grid initiatives. As of April 2012, EGI federated 352 resource centres providing almost 400 000 CPU cores and 140 PB of disk storage. One can see that, compared to WLCG, EGI has on average smaller capacity per site; this may result in larger operational overheads, but it is difficult, if at all possible, to quantify. EGI currently supports 226 Virtual Organisations from a variety of disciplines, with top 9 disciplines constituting 80% of users; 40% of users are affiliated with HEP – the largest single discipline [2].

Asia-Pacific Grid Initiative (APGI) is essentially an integral part of EGI, governed by a dedicated ROC (APROC). The NGI-based hierarchy is not always feasible in this region, but the model based on APGI and APROC affiliated with EGI proved to be successful.

Latin American Grid Initiative (LGI), grouping 18 resource centres, is also born out of collaboration with European Grid projects, most notably, GISELA [8]. LGI is rooted in NGIs and Equivalent Domestic Grid Structures (EDGS), and maintains strong cooperation with EGI.

Researchers in Africa are getting engaged in Grid infrastructures through a variety of specialised projects. Recently, South Africa established a National Grid infrastructure [9], also in close collaboration with EGI.

Grid operation in USA is less centralised when compared to Europe, and involves a variety of federations and autonomous contributors. Grid infrastructures in the USA and North America in general are largely defined by underlying technologies, but they are flexible enough to contribute to each other. The Open Science Grid (OSG) [10] federates more than 100 resource centres that provide more than 70 000 CPU cores, – the largest average capacity per centre. OSG serves more than 30 user communities, with HEP still being the largest single community. Some OSG resources are contributed to another infrastructure, XSEDE [11], which is built upon traditional HPC resources and includes 16 supercomputers and high-end visualization and data analysis resources; HTC resources are contributed via OSG and through the Condor [12] pool.

Grid resources in Canada and Australia do not form dedicated national Grid infrastructures, and are either a part of a generic national e-Science consortia, or enter the Grid via WLCG.

3. Observations and analysis

Despite the large variety of Grid infrastructures, they all face a number of similar challenges.

Current Grid operation involves a complex set of interacting domains: cross-national application-specific domains (like WLCG), generic national domains (like OSG), and global coordinating bodies (like EGI). Levels of expertise and quality of resource differ strongly within each domain and between domains; this is reflected in different service levels. Operation must also take into account different deployed middlewares and other technologies, especially between domains. Convergence to a smoothly working infrastructure that provides a user experience comparable to a single supercomputer, even within a domain, is not always an easy task.

From the user experience perspective, a failed job or an interrupted file transfer is unacceptable. When such failures occur, users seek assistance within the community, rarely reaching the actual experts who would provide a proper assistance. As a result, some failure modes are alleviated through custom solutions developed inside communities, leaving experts often unaware of issues. Unfortunately, many such custom solutions are made to circumvent tight Grid security requirements, resulting in vulnerable systems.

From the resource provider perspective, attending to single failures is not feasible, rather, overall performance figures matter. Grid infrastructure operators prefer centralized operations, real-time monitoring of generic system parameters, coordinated resource downtime handling, synchronised software upgrades, centralised issue tracking and similar streamlined procedures. And unlike end-users, resource providers are committed to strict security.

There are still some commonalities between resource providers and users. Neither community likes faulty or poorly tested Grid middleware, and yet neither community is flexible enough to engage in profound prototype testing or to install frequent upgrades. This puts heavy burden on technology providers, who do not necessarily have testbeds of complexity sufficient to reproduce all possible Grid infrastructure deployment and usage modes. Even with the Early Adopters program of EGI, the issue of middleware testing is not yet solved.

Middleware design problems are often solved by operational means. It is not uncommon to deploy large number of human operators monitoring Grid services and restarting them manually in case of failures. Not all such failures are caused by middleware: hardware and system software failures are still unavoidable, and have to be closely monitored as well, since any minor event on a local scale can become a major global issue in the Grid world.

Another challenge faced by Grid infrastructures is their non-uniformity and non-predictability. Resource owners prefer standard services and predictable usage patterns, while users prefer tailored solutions and can not predict their activities with great accuracy. It is therefore difficult to coordinate maintenance slots and minimise their impact; moreover, there is neither technology nor policy that would prevent accidental simultaneous shut down of a large number of infrastructure resources. There is no common worldwide policy authority, and although each individual infrastructure has such, grey areas exist where the infrastructures overlap.

While a significant effort has been done agreeing on service levels defining critical infrastructure services, there is still significant complexity involved. A service critical for one user community can be non-important for others. Different infrastructure elements come with different monitoring levels, and different resource centers have different priorities for incident response. Users worldwide monitor performance of their tasks on a round-the-clock basis, while operators monitor resources only during working hours, and use different tools.

In the end of the day, current Grid operation is a rather expensive task, requiring exceptional commitment and expertise from a large number of people.

With Grid being by nature a technology to federate resources, viable Grid organizations must be based on federation principles. This implies different owners, different consumers and different underlying technologies. Being aware of these differences, Grid community established Open Grid Forum (OGF) [13], through which technology providers and consumers can converge on common standards and best practices. Interestingly, OGF nowadays encompasses mostly computing-related aspects, as data storage and handling is considered less relevant for Grids.

In practice, however, many Grid organisations are not quite federated yet, and are still heavily rooted in HEP institutes. Grid technologies and policies are largely driven by WLCG needs, sometimes resulting in *de facto* standards which never reach OGF. Global Grid standardisation pace has slowed down, lacking support from user communities. Moreover, it becomes increasingly clear, at least in the HEP context, that data storage and handling is the source of most problems encountered by Grid users.

If we as a community strive to build a real Grid that is attractive for other user communities, we have to follow a set of simple rules:

- Build federated collaborative infrastructures, and not single-rooted centralized hierarchies
- Engage in common standardisation and convergence effort whenever possible, and avoid tailored core solutions
- Follow common practices in Open Source development and software distribution
- Be truly open to new user communities, offering same support levels throughout different Virtual Organisations
- Take good care of data storage, access and transfer services

4. Business models and vision

One may argue that Grid technologies exhausted their possibilities, being too rigid to serve dynamically changing user base. Need for computing and storage capacity for academic research will keep increasing in foreseeable future, and this leads to a dilemma: shall public research keep relying on public e-infrastructures, or shall researchers purchase computing and storage services from commercial structures?

Grid "business model" so far was fully based on availability of public funding to purchase commodity hardware and hire developers and operators. Research funding bodies worldwide were investing in Grid technologies as much as they could, and scientists used as much as they needed, without necessity to pay for resource usage, even if this resource was abroad. This model is satisfactory as long as public computing and storage resources are cheaper than those offered commercially; it is also appealing from the national funding point of view, being an investment into national e-infrastructures.

Cloud computing, while sharing many technology aspects with Grid, has a profit-oriented business model. Profit is achieved by reducing resource and operational costs by a variety of means. An important difference is that while Grid resources are on average rather small but tailor-made clusters, Cloud service providers rely on huge centres relying on comparatively cheap hardware. Still, costs incurred by researchers can be very substantial: the recent success by CycleComputing [14] involved costs approaching USD 4900 per hour. While this is an attractive solution for a quick one-off computational task, it is very unlikely that communities that require continuous massive computations over a period of many years (like HEP) will be prepared to pay this amount. Another issue with commercial Clouds is security, as terms of service are not always compatible with national regulations regarding data storage and access. Public Clouds may well solve some of these problems, but for this they will have to be able to provide services comparable to commercial ones, and for lower costs.

In retrospective, new technologies do not always make old ones obsolete. Just like personal workstations and clusters did not lead to demise of supercomputers, Grid technologies did not cause users abandoning clusters, and Cloud technologies did not prevent IBM from delivering a new Blue Gene/Q machine. Grid is likely to stay around as long as diverse specialized facilities exist, and their users are engaged in common research programs and are willing to federate resources.

HEP community embraced Grid because it offered a solution for such federation of resources. The resulting infrastructures appeared to be attractive for some other research communities. If the trend will continue, Grid may mature into permanent national and international infrastructures, independent of irregular project-based funding. If however the current Grid communities will keep giving low priority to standardisation and common practices adoption, Grid will gradually become limited to HEP and few other communities, the rest moving on to more attractive technologies with less overhead. If a serious investment will be made in standard approaches, Grid will become useful to every researcher. Standardisation and streamlining here should not be confused with defining a limited set of approved services or tools. Standardisation must be done on basic infrastructure levels, such as protocols, interfaces, schemas, packaging, configuration, monitoring and so on. When such basic aspects will be well defined, technology providers will be able to offer a choice of interchangeable services and tools offering different functionalities, user communities will be able to chose those which suit them best, and operators will be able to handle Grid services in a uniform manner, reducing operation costs.

5. Conclusion

Major Grid infrastructures of today were (and still are) largely driven by the needs of LHC computing and data storage. The WLCG infrastructure allows LHC to achieve scientific results almost instantaneously – the feat previously unseen in experiments of this scale. Impressive success of the LHC research program clearly shows that Grid computing is not just a viable paradigm, but an essential tool for scientists.

Still, researchers that benefit from Grid solutions for computing and data handling constitute, by different estimates, just about 1% of all scientists [15]. Benefits of Grid computing do not always outweigh overheads it introduces. One of largest overheads comes from complexity of Grid infrastructures operations.

Even after a decade of intensive development, Grid operations are extremely complex and costly, especially in terms of human resources. Commodity resources which constitute Grids do experience system failures; the sheer number of such resources means that failure frequency is rather high. Meanwhile, Grid middleware is still not sufficiently fault-tolerant. This requires permanent monitoring and frequent manual interventions. Huge LHC communities can handle these tasks, but smaller scientific groups can hardly afford it.

More complexity and thus additional failure modes are introduced through large number of highly customized application frameworks, developed as a higher abstraction layer on top of middleware. In the current deployment model, each Grid resource center needs a number of applications experts, familiar with large variety of scientific tools; this is hardly a scalable approach.

Different resource ownerships add another layer of complexity, which has implications on the choice of technologies and operations models. Consequently, different resource centres provide different service levels, making the overall infrastructure rather intricate.

Despite all such shortcomings, Grid is here to stay in foreseeable future. The main reason is that data collected by scientists will always be as distributed as scientific communities are. Modern science is more than ever a global collaborative effort, and this is exactly what Grid is:

a technology that enables collaboration of diverse computing and storage resource providers.

Cloud models, especially when used within a single administrative domain, proved successful in overcoming many of Grid complications. Still, as creation of one uniform single-domain worldwide scientific Cloud is extremely unlikely, federations of Clouds are inevitable, and one may argue that federated Clouds will face many of the traditional Grid problems. Moreover, traditional small-scale computational and storage resources are not likely to be completely abandoned as a scientific tool. Most likely scenario nowadays appears to be that of hybrid infrastructures consisting of Grid-like resources with a spill-over to Clouds. This outcome will not make Grids cheaper or redundant; rather, it will introduce more layers of complexity to existing e-Science infrastructures.

The way to increase efficiency and decrease operational costs of such infrastructures is to reduce heterogeneity and complexity as much as reasonably possible, while maintaining the desired functionality. It is neither feasible nor reasonable to force resource providers to deploy e.g. identical custom operating systems, or to limit the choice of services offered by infrastructures. Instead, Grid communities must focus on standard approaches in all areas, from software development to operations. Commitment to use and contribute to open standards, and active engagement in common approaches is absolutely necessary. Without such commitments, Grid will remain a custom tool relevant only to few selected applications that use it today, like those in High Energy Physics.

Looking further into future is an unrewarding task. One thing is certain: many new technologies and models will come, probably using completely different approaches. We must stay vigilant, and use any opportunity to get most efficient tools for our research.

Acknowledgments

The author would like to thank Lothar Bauerdick, Ian Bird, Simon Lin, Miron Livny, Steven Newhouse and Ruth Pordes for valuable contributions.

References

[1] The Large Hadron Collider Web site URL http://public.web.cern.ch/public/en/LHC/LHC-en.html
[2] European Grid Infrastructure Web site URL http://www.egi.eu/
[3] Scientific Linux Web site URL http://www.scientificlinux.org/
[4] gLite, Lightweight Middleware for Grid Computing Web site URL http://glite.web.cern.ch/glite/
[5] Worldwide LHC Computing Grid Web site URL http://cern.ch/lcg
[6] Aderholz M *et al.* 2000 Models of Networked Analysis at Regional Centres For LHC Experiments (MONARC), Phase 2 Report Tech. Rep. CERN-LCB-2000-001. KEK-2000-8 CERN Geneva
[7] Foster I and Kesselman C 1999 *The Grid: Blueprint for a New Computing Infrastructure* (Morgan Kaufmann)
[8] Grid Initiatives for e-Science virtual communities in Europe and Latin America (GISELA) Web site URL http://www.gisela-grid.eu/
[9] CSIR Meraka Institute South African National Grid (SAGrid) Web site URL http://www.csir.co.za/meraka/
[10] Open Science Grid Web site URL http://www.opensciencegrid.org
[11] Extreme Science and Engineering Discovery Environment (XSEDE) Web site URL http://www.xsede.org
[12] Litzkow M, Livny M and Mutka M 1998 *Proc. of the 8th International Conference of Distributed Computing Systems, pp. 104–111*
[13] Open Grid Forum Web site URL http://www.ogf.org/
[14] CycleComputing Ramps Global 50000-Core Cluster for Schrodinger Molecular Research Press release URL http://cyclecomputing.com/news/news?start=6
[15] Foster I 2012 *Keynote talk at Grid Computing: the Next Decade, 4th – 6th January 2012, Zakopane, Poland*