Evolution of Scientific Computing in the next decade: HEP and beyond

WLCG Overview Board

17th December 2018

Contact: Ian Bird (Ian.Bird@cern.ch),
Simone Campana (Simone.Campana@cern.ch)

Abstract

High Energy Physics (HEP) has demonstrated a unique capability with the global computing infrastructure for LHC, achieving the management of data at the many-hundred-Petabyte scale, and providing access to the entire community in a manner that is largely transparent to the end user. Other HEP experiments have expressed a desire to benefit from this infrastructure and organization, and recent interactions with other related science communities that have similar needs on the HL-LHC timescale show interest in collaboration. In this paper we outline a proposed strategy by which the computing infrastructure developed for LHC could be broadened in collaboration with other interested HEP experiments, while retaining the LHC-specific needs within the WLCG collaboration. This proposal is in line with the WLCG strategy for addressing computing for HL-LHC, and is aligned with European and other international strategies in computing for large scale science.

Introduction

High Energy Physics (HEP) has demonstrated a unique capability with the global computing infrastructure for LHC, achieving the management of data at the many-hundred-Petabyte scale, and providing access to the entire community in a manner that is largely transparent to the end user. This is still a rather unique facility in science, but as other communities' needs grow beyond what can be provided by individual facilities, they too are starting to tackle similar issues. Seventeen years ago, when the work started, there were no examples of how to build such a system, and no experience from industry or others. If we were to design the system today, of course, we would benefit from the tools and expertise of the global internet companies. HEP has a challenge for the foreseeable future – which is how to achieve a scale of computing and data management that is orders of magnitude greater than that of today, while maintaining a reasonable cost envelope. HL-LHC is the most immediate such challenge, but we also have other high data-rate experiments, and future potential facilities that must be considered. Some HEP projects have already expressed their desire to build their computing system leveraging aspects of the infrastructure deployed for LHC, while possibly contributing with new solutions. We believe we should facilitate this process and evolve the current Worldwide LHC Computing Grid (WLCG) infrastructure into a HEP-wide scientific data and computing environment available for the future to interested parties in our field. Importantly, in addition, we observe similar needs arising in related fields (astronomy, astro-particle) with many of the HEP facilities often directly involved. In planning for the future, we must consider compatibility and synergy at the facility level. Taking the success of the WLCG as a starting point it could be envisaged to evolve the infrastructure and tools as a basis for computing for HEP for the coming years, while challenging the concerns of cost (both in terms of equipment and operationally), organization, and community needs.

Current situation and opportunities of the WLCG Organization

The WLCG collaboration has demonstrated and implemented a distributed computing model for the LHC experiments, which has played a crucial role in their scientific mission. The WLCG community is strongly established and a network of trust exists between stakeholders. This allows the implementation of a very lightweight decision-making process, based on consensus at various levels of the organization; particularly at the Overview Board, the Management Board and the Grid Deployment Board. The non-contractual nature of the agreements WLCG is built upon are in sharp contrast with the commercial computing infrastructures that now exist at similar or much larger scale. The infrastructure, while heterogeneous, builds upon a set of agreed interfaces to compute and storage, common middleware tools and policies in matters of security, identity management, resource sharing, monitoring and accounting.

The WLCG collaboration has a strong focus on computing services, infrastructure and policies. The importance of application software was recognized from the very beginning and an LCG application area was established in the WLCG organization. The Architects Forum holds the responsibility to coordinate the development of common application software and reports to the WLCG Management Board. The activity of the Architects Forum has however decreased in the last years, as the software stack consolidated around well-established components. Preparing for scientific computing in the 2020s, application software will play a critical role. In fact, innovative solutions will need to be considered and implemented: experts with different skill-sets, such as parallel programming and machine learning, will need to complement the more physics-oriented expertise today available in our community. In addition, improving software performance will be a critical aspect of the way to reduce the cost of computing in the future. Existing initiatives and projects such as the HEP Software Foundation (HSF)[1] and IRIS-HEP[2] should be leveraged and form part of the strategy.

From the point of view of computing services, WLCG, since inception, has been a driver for middleware initiatives sponsored by the Funding Agencies and the European Union. As a result, it was able to focus its own effort on delivering compute and storage services to its scientific community. WLCG has benefited significantly from solutions coming from diverse sources and built a coherent infrastructure. In recent years, the mechanism to identify the future needs and influence the directions of the various development efforts has not been as effective as in the early phase, with the consequence that some aspects of the WLCG needs have been neglected.

The demand for scientific computing capacity will certainly increase enormously, while the funding is almost certain not to increase at the same rate. Already during LHC Run-2 (2015-2018) the funding agencies of the LHC experiments made clear that funding beyond a constant budget should not be expected during Run-2 and beyond. This constraint is purely economical and has no basis in terms of computing needs for science. Such a level of flat funding enabled the provisioning of enough resources for physics at Run-2 and we expect it to be also adequate for Run-3, but will unlikely be sufficient for HL-LHC unless a major evolution in the application software, the computing services and the infrastructure happens. Without such an evolution and because of the flat funding constraints, WLCG would not be able to cope with the future needs of the LHC experiments.

WLCG has so far been the major player among high energy physics and many other sciences in terms of data volume and compute capacity. This allowed it to steer the evolution of the infrastructure and services in the direction of the LHC experiments' needs. WLCG has been characterized by a pragmatic ability to accommodate the different and changing requirements of the experiments with time and the evolution of their computing systems. In the 2020s we expect other high energy physics experiments, such as DUNE, and other sciences, such as astronomy (e.g. the SKA organization) to require a similar level of resources to LHC. In order to maximise the return on investment of the Funding Agencies it would be advantageous to foresee, where appropriate, a common infrastructure and set of tools serving the needs of the set of sciences they support. A close collaboration between WLCG and other communities will be required in order to maintain influence and help ensure the infrastructure is able to effectively meet the requirements of all stakeholders.

The evolution of the WLCG collaboration, bringing in new communities and new ideas, naturally provides an opportunity to revisit the relationships between stakeholders. Ambitious initiatives such as the European Open Science Cloud (EOSC) provide a timely opportunity to rethink how an evolving WLCG can most effectively collaborate with other sciences in HEP and beyond.

[1] https://hepsoftwarefoundation.org
[2] https://iris-hep.org

Finally, some countries have invested and intend to continue to invest in capacity at computing facilities that historically were not particularly suited for HEP, such as High Performance Computing (HPC). Whilst this is not economical in its own right for HEP, the additional funding streams for HPC in such places mean HEP must take advantage of the opportunity to access those spare resources, at least for certain computing tasks. The utilisation of resources at HPC facilities comes with challenges at the level of application software, data access, as well as access policies and resource scheduling.

One of the major issues we identify below is software. Historically software engineering posts have not been given adequate priority, and consequently funding lines have not supported these in a systematic way. This is not tenable for the scale and technical challenges of the future. A major change is needed in the particle physics funding mind set. Adequate support for training initiatives is an important aspect of this: there is a growing consensus that the kind of professionals needed in the field of scientific computing are not just the traditional computer scientists, but rather a profile of researcher with competences crossing the domains of computing as well as physics. This kind of professional, in a small part, already exist as young researchers with a particular attitude towards computing, but their number needs to increase with a new generation, nurtured from the beginning with a specific training in the required disciplines.

WLCG has the opportunity to leverage its strengths and play a central role in the evolution of scientific computing, as it is the community with the largest experience in large scale distributed scientific computing. As we evolve the infrastructure for LHC computing, there is an opportunity to see what parts of the infrastructure are of interest to other communities; how the infrastructure can be evolved to better match the needs of others; and thus achieve a more cost effective overall solution.

Changing the governance model and splitting the roles of WLCG as a collaboration from the role of WLCG as an infrastructure will allow to open up the latter to more communities in and outside HEP, helping to create consensus among such sciences. It will also allow WLCG to retain the current aspects of its governance which have so far demonstrated their effectiveness and might be models for other sciences if they wish.

The scientific computing evolution strategy is based on three pillars: 1. leveraging the existing HEP computing infrastructure and evolving it to serve as a common computing system for HEP and sciences beyond HEP; 2. evolving the facilities and services to build a HEP Data Cloud[3]; 3. investing in common software and software techniques, including training, dissemination and recognition. As part of the process to achieve such a strategy we will propose an evolution of the current WLCG governance: splitting the current WLCG organization into firstly a project specific to LHC needs, and secondly a collaboration which serves a common infrastructure for HEP and sciences beyond HEP. The next sections will elaborate the details of these ideas.

[3] Here we refer to a “Data Cloud”, also often described as a “Data Lake”. The intention is the same - a distributed data repository serving data to compute resources and clients.

General Infrastructure

The underlying core services of the infrastructure constitute today one of the major values of WLCG. For LHC we have the computing resources at close to 200 sites using this infrastructure. Baseline grid services are supported by mature monitoring, operational and support processes and teams, including worldwide collaboration on security and incident response. The "grid" that enables the coherent use of those resources must evolve over the coming years, and be capable of supporting continually evolving computing models, and be agile to technology changes.

We have in place global networking infrastructures, not only those provided by the National Research and Education Networks (NRENs) and their coordinating bodies, but HEP-specific[4] structures such as the LHC Optical Private Network (LHCOPN), and the very successful LHC Open Network Exchange (LHCONE) overlay network, which provide the ease of management and connectivity that will be essential for the future. Today this is already used by more than the LHC experiments. While the LHCOPN is a private network, it is nevertheless a good model for specific situations in the future – as will be discussed below. Network resources, originally expected to be the main limiting factor, emerged as probably the most solid one, both in terms of capacity growth and reliability. They will play a central role in the evolution of the infrastructure.

We also have a global Authentication, Authorisation, Accounting (AAA) service, and associated identity management, trust and policy networks. This is extremely valuable, and very unique. However, it is clear that the X.509 underpinnings are too specific to our infrastructure and are not the best for the future as they diverge from the widely adopted open source trends. Federated identity mechanisms (e.g. eduGAIN) and token-based authentication are being introduced as core components of the AAA infrastructure.

The WLCG infrastructure already integrates heterogeneous resources such as commercial and academic clouds, HPC facilities and volunteer computing. Such resources will play a more significant role in the future and further harmonization in their adoption is a key element of our strategy.

The WLCG infrastructure also provides a sophisticated way of dealing with all aspects of distributed data management including the distribution, resilience, archiving and cataloguing of the huge volumes of data, and the means to match compute resources to data across globally distributed compute resources.

However, one of the concerns today for LHC is the cost of the computing resources, and in particular the cost of storage which accounts for close to 70% of the overall hardware cost. That cost is in large part driven by the need to distribute data globally, with consequently several copies of much of the data, and consequent costs both in storage and in operations. For the future, we must consider an alternative model to reduce those costs, such as the HEP DataCloud described below.

[4] https://edugain.org

We have seen other HEP experiments asking to be allowed to benefit from the WLCG infrastructure already. The needs of such experiments are in line with the priorities and the evolution strategy of the infrastructure and there is therefore a strong motivation for it to be a shared resource. We should note that this paper is not proposing to use the same resources for all experiments, but rather to try and use the same underlying infrastructure, tools, software, and support as far as possible so that new projects are easier to support on existing facilities. Of course, this helps opportunistic use and sharing, but does not impose it.

HEP DataCloud

Because the majority of the cost in WLCG, both in terms of hardware and operational staffing, is on storage and data management, these aspects deserve special attention in defining a strategy for the future. The currently envisaged model builds on the experience of large commercial cloud providers, as well as the LHC expertise in many-hundred-Petabyte scale data management. It is foreseeable on the HL-LHC timescale to connect most of the large HEP data centres with dedicated and private multi-Tb/s network links. The combined "virtual data centre" would store all of an experiment's data, and by policy replicate it between the data centres as needed. In this way, we would achieve reliability and availability of the data. Into this data cloud we would plug compute resources. These resources may be co-located at the data centres, or may be other facilities, such as commercial centres or other large-scale, HEP-owned resources. The model also allows for inclusion of commercially procured storage. Policy and practical reasons would prevent reliance on the latter for non-reproducible data sets, and such storage should be redundant enough that a commercial centre could "unplug" without loss of vital data. The architecture clearly relies on a very strong collaboration with the networking community, with adequate policies and capabilities to agilely connect to commercial partners.

A key concept in this vision is that data can be processed directly inside the data cloud or externally through a content delivery system, minimizing the possible impacts due to network latency or capacity. In the LHC case all reconstructed data beside the final analysis sets would be kept in the cloud. Having all of the data virtually co-located in this way may open the way to radically new analysis models, and strongly supports today's models such as analysis trains. It would also permit the increased use of economical high latency media, such as tapes, as an active store for organized analysis, again helping with cost. This type of model also allows cost optimization through the use of hybrid centres: HEP owning compute resources at a level that is guaranteed to be fully used is very cost effective, and supplementing this with elastically provisioned resources as needed, presumably with some form of cheap spot-market pricing. This would allow an agile control of the cost, and can evolve as the commercial markets evolve.

Building and operating a HEP DataCloud may require new funding models and management methods and several aspects will need to be investigated and clarified: the capability to easily procure commercial resources at large enough scale to get economy together with the political implication of purchasing from the largest cloud vendors. At the moment, it is clear that real cost efficiency and elasticity require the use of spot-market style pricing and the implications for the procurement process need to be understood. Today we have many Tier-2 centres that also provide significant resources, together with other opportunistic resources. We should distinguish in terms of scale rather than role. Large scale centres could participate as “compute plug-ins” in the above model, while others primarily provide simulation resources. It should be considered that some 50% of LHC compute load is simulation, and the same will be true for forthcoming large experiments. Depending on the type of resource, some centres may be best suited to specific types of workload (e.g. HPC for event generation). This model will also easily accommodate the anticipated case of a funding agency moving academic computing internally to cloud credits rather than in-house facilities. The data cloud model also may be very interesting for other sciences, e.g. SKA regional centres, as it provides resiliency and long-term preservation capabilities. Scale-out is also inherent to the concept, although for practical reasons we might imagine loosely coupled US, European, and Asia-Pacific instances.

Software

Software, for the future of HEP, is as important as the infrastructure itself. Software and infrastructure must be considered together as the separation is a source of inefficiency and thus cost. WLCG has started a Cost Model working group to understand the interplays between software, workflows and infrastructure and to optimize the overall spending, including hardware, staffing and operational costs.

The HSF Community White Paper[5] identified key areas in the software domain which should be top priority in the future strategy of HEP computing with particular impact on WLCG on the timescale of HL-LHC. The software frameworks and algorithms of the WLCG experiments were designed many years ago and today cannot leverage efficiently all features and architectures of modern hardware (e.g. vectorization and use of accelerators). Modernising the software in this direction requires skill-sets not broadly available in the HEP community. Building such know-how requires dedicated effort in terms of training. It also needs the right form of recognition in terms of career opportunities of the software developers for which HEP is hardly competitive with industry and therefore has a problem in retaining expertise. It also requires a set of tools and procedures facilitating the process, such as elements of the build systems, tools for documentation, and advice on licensing among several others.

While part of this work is experiment and community specific, a large part of it can be achieved through a common effort at different levels: from common tools and procedures, to sharing methodologies, to actual software libraries shared by multiple experiments. We proposed in the section above the idea of a common infrastructure for scientific computing, intended as common set of tools, services and support for experiments to use, with no imposition of particular choices. Here we propose the same model to be applied to software.

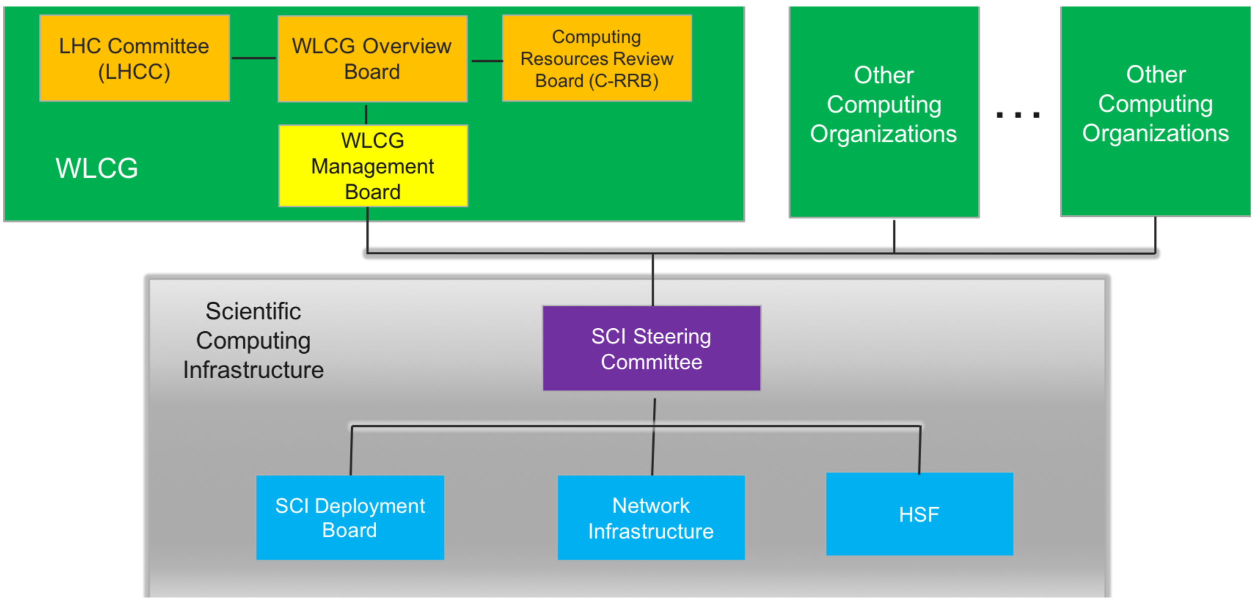
It is essential that this is recognised and supported. We have made significant progress in this area by setting up the HSF. It must be understood that there is no one-size-fits-all, but rather we need a community wide coordination of available tools covering the full stack from general workflow and data management tools to the application level. Common tools and libraries that can be used to build up the needs of an experiment are required. The HSF has made steps in this direction, covering many of the aspects of software for the HEP community, with the aim of collecting a set of tools contributed, developed, maintained and

[5] https://arxiv.org/abs/1712.06982

evolved by the community. It is also a mechanism for pursuing common R&D efforts in software, and for coordinating things like technology tracking, and developing software tooling for development and performance analysis. While the HSF is a good framework, this does not remove the need to engage appropriate levels of investment in this area. HEP must recognise that software efficiency and performance will be key to maintaining an affordable infrastructure. We must get ourselves into the position of being able to evolve our codes to make efficient and best use of the evolving computer architectures. This is not a one-off effort but will require sufficient and on-going investment in people and skill development and retention.

Scientific Computing Infrastructure governance: organization and steering

We propose to evolve the existing computing infrastructure in a way that could benefit a wider HEP community: The **Scientific Computing Infrastructure** (SCI). For such a process to happen we clearly need the buy-in of other potential stakeholders, and the community itself. What we have found that works best is a lightweight steering mechanism rather than strong governance. The experience in setting up the HSF along those lines is very clear as an example of effectiveness. The WLCG collaboration, while structured with a more formal governance and decision-making process, agreed in its Memorandum Of Understanding, de facto also reaches decisions by consensus. A possible organization of the Scientific Computing Infrastructure and its interplay with the major computing projects which it is part of are visualized in the figure.



From the current WLCG Organization, the **WLCG collaboration** would factor out the aspects specific to the LHC experiments. It would also manage the aspects of the infrastructure specific to LHC such as the Optical Private Network. The WLCG experiments would continue to negotiate with their funding agencies for both pledged resources and access to opportunistic ones. The WLCG collaboration would conserve its Memorandum Of Understanding (possibly amended with the modifications proposed here) and continue with the current resource management process: the **WLCG Overview Board** would oversee the functioning of the WLCG project and its role in the SCI; the project would report and respond to the **Computing Resources Review Board** about legal and resource matters, and to the **LHC Committee** concerning the technical matters and scientific aspects. The **WLCG Project Leader,** appointed by the CERN Director General in consultation with the Overview Board, would continue representing the project in front of the above-mentioned committees, the SCI and the outside. The **WLCG Management Board,** chaired by the WLCG project leader would retain its role of managing the day-by-day aspects of the project and represent the WLCG interests to the SCI.

HEP experiments or projects others than WLCG (see "*Other Computing Organizations*" in the figure) would have their own Computing Management and Organization which may or may not resemble that of WLCG. They also would have full autonomy on resource negotiations, usage policies and decisions about which services and tools to use or not to use. Clearly, they would need to fund the computing resources they will need. We do not expect nor propose to use the same resources for all experiments, but rather to try and use the same infrastructure, tools, software, and support as far as possible so that new projects are easier to support on existing facilities. Of course, this helps opportunistic use and sharing, but does not impose it. Experiments joining the SCI would likely produce their own MoU and will need to identify their own reporting lines to the SCI bodies.

The **Scientific Computing Infrastructure** would be driven by the major HEP sciences with a stake in the common infrastructure. Such sciences are represented by the WLCG project and the computing projects of other major HEP experiments who choose to participate, as just described above. A **Scientific Computing Collaboration Steering Group** should be set up. Its role would be to ensure that the physical and software infrastructures evolve in the direction that is suitable for the community and its projects. It should also be a mechanism to obtain or encourage funding and contributions of effort, through direct feedback to the Funding Agencies and laboratories. Finally, such a steering group would be an ideal forum within which to broker community-wide needs, such as licensing, joint procurements, agreements, and policies. For example, community-wide agreements with cloud vendors to get scale economies. It should also address political concerns, for example how to evolve the funding models.

The composition of the group should be discussed more widely, but most likely would include the heads of Information and Communication Technology of the major High Energy and Nuclear Physics laboratories worldwide, the computing project leaders from the major projects, facilities and experiments. This group could receive a mandate from a body such as the International Committee for Future Accelerators (ICFA), which is key to obtaining recognition in some countries. The governance should be very lightweight and decisions should be taken through consensus. The steering group would report on its activities to the involved sciences, through their representatives. The day by day activities would be organized leveraging existing mechanisms and initiatives, such as the HSF, HEPiX, existing working groups and task forces.

The **HSF** would be the vehicle by which the foreseen improvements in the area of software are addressed, as explained in the sections above. It would replace the role of the Architects Forum in WLCG and broaden its scope to provide a common set of tools, libraries and techniques for the different projects in the SCI. It would respond to the SCI Steering Group and inherently to the SCI projects.

The **SCI Deployment Board** would cover, in the new organization, the functions of the Grid Deployment Board in WLCG and therefore replace it. The representation would be broadened to experiments and computing centers of the full SCI, to discuss technical and related matters, prepare the decisions and plan the deployment and operations of the SCI services.

While we explicitly discuss this for HEP here, there is potential interest from other related scientific collaborations, for example the astronomy/astro-particle and 3rd generation gravitational wave communities. Negotiating such a broadened scope would also be welcome, and help address concerns of funding agencies and large scientific data centres about being able to more uniformly support a set of sciences with significant requirements.

Relationship to EOSC and other infrastructures

The model of infrastructure proposed here is fully aligned with that which is to be prototyped in the EC-funded ESCAPE project[6] which brings together HEP, astronomy, and astro-particle science ESFRI projects in Europe, to influence the European Open Science Cloud (EOSC) developments, and addressing the needs of Exabyte-scale data experiments compatible with the requirements of FAIR data management. The infrastructure would also leverage services and policies of the EOSC where appropriate in this global context, as well as offer services and tools to the EOSC catalogue.

The same is true of compatibility with other international infrastructures with which WLCG today interacts – for example Open Science Grid in the USA. What is proposed here does not introduce any incompatibility, but rather opens the door for broader collaboration with more science communities.

Conclusions

The WLCG organization has successfully provided a global computing service for the LHC experiments for more than a decade. In this paper we have considered the future HEP ecosystem and the likely evolution of HEP computing needs in the coming years. We propose a model where the current WLCG infrastructure evolves into a Scientific Computing Infrastructure, that could accommodate the needs of other experiments. Such an infrastructure would be driven by all of the major HEP and related sciences, reducing cost by leveraging economies of scale, common tools, services and operations. Each science would retain a high degree of autonomy in negotiating resources, setting policies and adopting appropriate tools and services.

[6] https://www.escape2020.eu, https://indico.in2p3.fr/event/18279/