PROCEEDINGS OF SCIENCE

Strategies of a WLCG Tier-2 site to meet the challenge of ever growing demands on delivery of computing resources

J. Chudoba ,[a,*] D. Adamová,[b] A. Mikula,[a,b] L. Míča,[a] M. Svatoš,[a] P. Šesták ,[a] J. Uhlířová[a] and P. Vokáč[a]

[a] *FZU - Institute of Physics of the Czech Academy of Sciences*
Na Slovance 1999/2, 182 00 Prague 8, Czech Republic
[b] *NPI - Nuclear Physics Institute of the Czech Academy of Sciences*
Hlavni 130, 250 68 Rez near Prague, Czech Republic
E-mail: jiri.chudoba@cern.ch, dagmar.adamova@cern.ch, mikula@fzu.cz, mica@fzu.cz, svatosm@fzu.cz, sestakp@fzu.cz, uhlirova@fzu.cz, vokac@fzu.cz

The current era of Exascale computing brings ever growing demands on the amount of available computing performance, storage capacity and network throughput. This affects also the massive computing infrastructure for management of data produced by the experiments at the LHC, the Worldwide LHC Computing Grid (WLCG). The standard financing used for many years enabling the resource growth of 10-20 % is no more sufficient and to close the resource gap different methods are pursued. The sites involved in the WLCG are encouraged to find non-grid external resources to be used for WLCG tasks. Probably the most popular among them are High Performance Computing (HPC) Centers.

In this contribution, we present an overview of one of the WLCG sites, the distributed Tier-2 center in Prague, Czech Republic. It is a standard example of a WLCG medium size Tier-2 center concerning the hardware resources, site management and the network connections within the WLCG, so a general picture of a WLCG Tier-2 site is provided. In addition, our site complies with the current trends supported by the WLCG. First it is the use of resources of the external national HPC center in Ostrava and second providing resources not only for the LHC experiments but also other particle and astro-particle experiments. This way we follow the recently adopted strategy towards a sustainable and shared infrastructure adapted to the needs of large Exascale science projects. In addition, we make use of BOINC which provides some minor contribution to our resources.

41st International Conference on High Energy physics - ICHEP2022
6-13 July, 2022
Bologna, Italy

*Speaker

PoS(ICHEP2022)1146

1. Introduction

The Large Hadron Collider (LHC) at CERN is a unique particle accelerator in the world not only due to the Physics results delivered during its operation but also due to the system built for processing and management of the data it has been producing. It was clear already from the first estimates in 1990s that LHC would require a special computing infrastructure delivering more resources than what was available at CERN for the mass of the future LHC data. In 2001, the concept of the Worldwide LHC computing Grid (WLCG) [1] was created and the new infrastructure building set off. Ever since its first campaign in 2003, the WLCG has been in operation without a break, while constantly extending and upgrading the infrastructure. Each of the countries participating in LHC experiments shares responsibility of dealing with LHC data to deliver Physics results. Without the WLCG this would be impossible.

2. The Computing Grid for LHC

The WLCG is a global computing infrastructure and consists of a network of connected computing centers which together can accomplish massive tasks. Currently WLCG provides about 1.4 million computer cores and 1.5 ExaBytes of storage from over 170 sites in 42 countries, worldwide, allowing for almost real-time access to LHC data. It runs over 2 million jobs per day and the global transfer rates exceed 260 GB/s.

The WLCG ecosystem consists of centers ranked as Tier-0, Tier-1 and Tier-2. Tier-0 is CERN, 15 Tier-1s are large computing centers and 160 Tier-2s are smaller size centers. All the sites provide services according to their hardware resources and connectivity level. Although Tier-2s are seemingly inferior to Tier-1s, their role in the WLCG ecosystem is crucial: in total they deliver the same amount of resources as Tier-1s.

3. CZ-Prague-T2: a distributed WLCG Tier-2 site in the Czech Republic

The Czech Republic has been participating in the LHC experiments ALICE and ATLAS since the beginning. Our contribution to WLCG resources is provided by the Tier-2 center CZ-Prague-T2 [2]. Our center took part in the very first WLCG campaign in 2003 and has been in operation ever since.

Currently CZ-Prague-T2 is a modest size Tier-2 site recognized as one of the 10 best performing Tier-2s in WLCG. It has distributed resources [2] with the main part installed at the Institute of Physics of the Czech Academy of Sciences (FZU) in Prague. The complete stack of resources includes 13,000 job slots, 11.07 PBytes of disk space and network connectivity ranging between 10 Gb/s and 100 Gb/s. The site services are used mainly by the CERN experiments ALICE and ATLAS, but also by other HEP and astroparticle experiments. It is anticipated that LHC will collect a factor of 10 to 100 more data in the next 3 to 7 years (High Luminosity LHC era (HL-LHC [3])) than the current data taking reaching the multi-Exabyte scale.

There will be potential shortfalls in the LHC Grid resources, the requirements exceeding up to 10 times the flat budget outcome. There is a number of projects proposing different strategies to balance the gap and create a high performance ecosystem for HL-LHC. From the point of view of our Tier-2 site, some of the strategies are crucial.

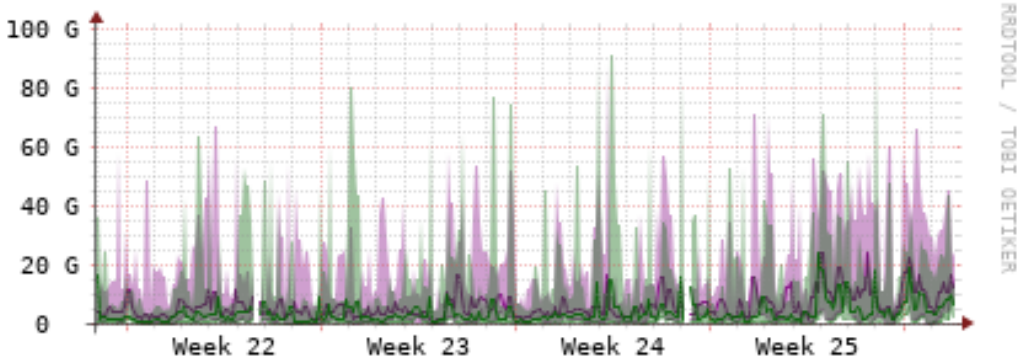

Figure 1: Network traffic between the main site of CZ-Prague-T2 and LHCONE in June 2022. Maximum throughput reaches to 91 Gb/s, the nominal capacity is 100 Gb/s.

There is a need to upgrade the network throughput to allow for a smooth streaming of data for processing at different sites. At CZ-Prague-T2, we are continually upgrading the connectivity, e.g. our connection to LHCONE [4] will soon be 2x100 Gb/s, see Figure 1. Considering the LHC data streams of PB/hour, the Grid system needs to be enhanced with other computing paradigms that could coexist with Grid, such as volunteer computing, cloud computing, and HPC [5].

In our case, clusters of the HPC center in the city of Ostrava, IT4I [6], have been used as an external resource for ATLAS simulation campaigns. The center is supported by the European High Performance Computing Joint Undertaking (EuroHPC JU) [5]. The profile of running CZ-Prague-T2 ATLAS jobs at the HPC IT4I clusters in the period from September to December 2021 is shown in Figure 2. CZ-Prague-T2 also uses the BOINC volunteer computing [7].

Considering the anticipated deficit of storage during the next LHC runs, a new strategy is in development in Data Organization, Management and Access (DOMA, [8]): WLCG Data Lakes [9]. CZ-Prague-T2 plans to gradually extend its storage capacity and get included into one of the Data Lakes.

Another path to create a robust computing infrastructure for ExaByte scale needs is to share fundings with other data intensive science projects like other HEP experiments or astroparticle projects etc. CZ-Prague-T2 follows this principle providing services to Auger, CTA, DUNE, NOVA and other projects.

4. Summary

There would be no Physics discoveries from the data produced during the LHC runs without the services of the WLCG. Approaching the ExaByte scale, a new concept of the computing grid for the LHC is crucial. The Czech Republic Tier-2 center is actively participating in the strategies leading to the WLCG infrastructure meeting the needs of future LHC runs.

Acknowledgments

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic project LTT17018.

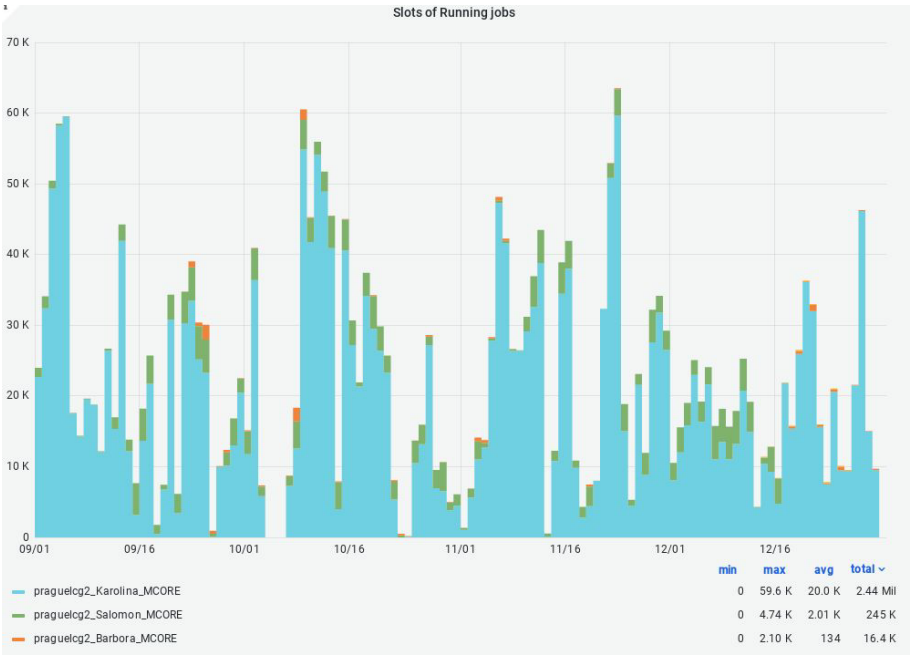

Figure 2: Running CZ-Prague-T2 ATLAS jobs at the HPC IT4I clusters Karolina, Salomon and Barbora from September to December 2021.

References

[1] Ian Bird et al, *Update of the Computing Models of the WLCG and the LHC Experiments*; https://cds.cern.ch/record/1695401/files/LCG-TDR-002.pdf

[2] Martin Adam et al, *Distributed resources of Czech WLCG Tier-2 center*, EPJ Web of Conferences **245**, 03034 (2020)

[3] High Luminosity LHC CERN; https://home.cern/science/accelerators/high-luminosity-lhc

[4] LHCONE; https://lhcone.web.cern.ch/

[5] The European High Performance Computing Joint Undertaking; https://eurohpc-ju.europa.eu/index_en

[6] IT4Innovations; https://www.it4i.cz/en

[7] BOINC Computing; https://lhcathome.cern.ch/lhcathome/

[8] Data Organization, Management and Access (DOMA); https://iris-hep.org/doma.html

[9] Ian Bird et al, *Architecture and prototype of a WLCG data lake for HL-LHC*, EPJ Web of Conferences **214**, (2019) 04024