# Computing for the LHC: operations during Run 2 and getting ready for Run 3

**Dagmar Adamova**[*]

*NPI AS CR Prague/Rez*
*E-mail:* adamova@ujf.cas.cz

**Maarten Litmaath**

*CERN*
*E-mail:* Maarten.Litmaath@cern.ch

The computing infrastructure for the LHC data handling (Worldwide LHC Computing Grid-WLCG) was well prepared for Run 2 and delivered a steady data processing since the first collisions. Over the past years including the Run 1 period of the LHC the WLCG developed a unique expertise in building and operating a very large scale infrastructure for unprecedented amounts of LHC-produced data. Together with other research projects the WLCG entered the era of Data Intensive Science. The cost of Computing infrastructures and services reached the level needed for building the LHC detectors. In this contribution, we will present ongoing and planned developments to evolve the computing models towards the needs for the High Luminosity (HL) LHC, including the disappearance of the distinction between "online" and "offline", analyses of the physics performance and corresponding computing costs, and inter-operation between commercial cloud services and publicly funded e-infrastructures.

*54th International Winter Meeting on Nuclear Physics*
*25-29 January 2016*
*Bormio, Italy*

---

[*]Speaker.

## 1. Introduction

The original concept of a huge computing grid needed to process and store the data produced by the Large Hadron Collider (LHC) [1] at CERN was conceived in 1999. Since then, the concrete implementations of the parts of the computing infrastructure were evolving along with technology and went through remarkable development.

The Worldwide LHC Computing Grid (WLCG) [2] has been built, expanded, gradually upgraded and systematically tested with Data Challenges [3] ever since 2003. As a result, WLCG was able to process efficiently the data pouring from the LHC experiments right from the beginning of the LHC operations in 2009, although the data rates today are much higher than anything originally planned for standard years of data taking.

At the time of writing this contribution, the LHC is in the middle of Run 2, the second continuous period of its operations. In contrast to Run 1, the current data taking is in the atmosphere of worries that a lack of computing resources might compromise the physics achievements. Fortunately the WLCG is still able to handle all the produced data although the data rates approximately doubled in comparison with Run1. In this article, we will present the WLCG performance during the first part of Run 2 and then ongoing and planned developments to evolve the model of data flow and other future challenges towards the needs of the High Luminosity (HL) LHC era.

## 2. Worldwide LHC Computing Grid: Current status

WLCG is a global distributed infrastructure which provides computing resources to store, distribute and process the data generated by the LHC experiments. It provides almost real-time access to LHC data and enabled achievement of exciting physics results including the discovery of the Higgs Boson in 2012. WLCG includes almost 170 computing centers in 42 countries worldwide (see Fig. 1) with a Tier-hierarchy structure, where CERN is the largest, central Tier-0 site [2, 3].
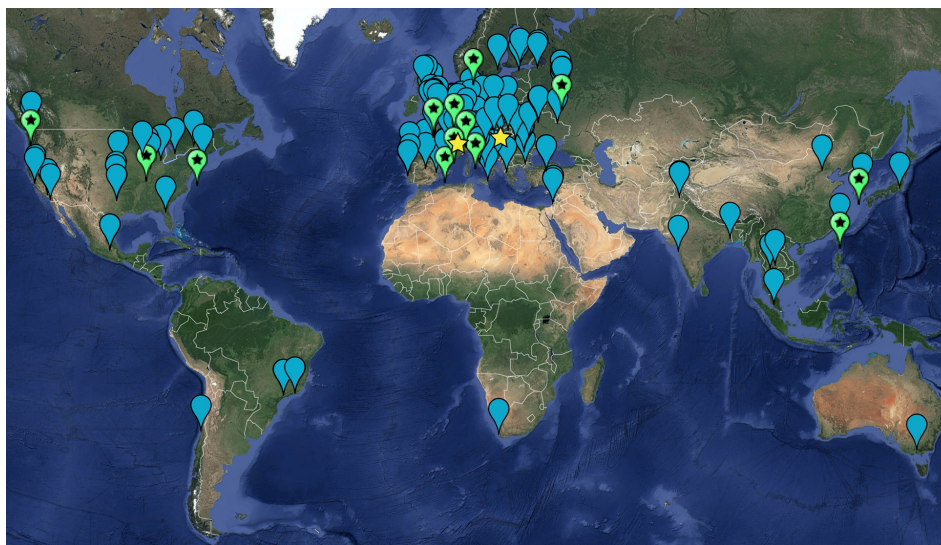


**Figure 1:** Map of the Worldwide LHC Computing Grid, showing individual computing sites, see [4].
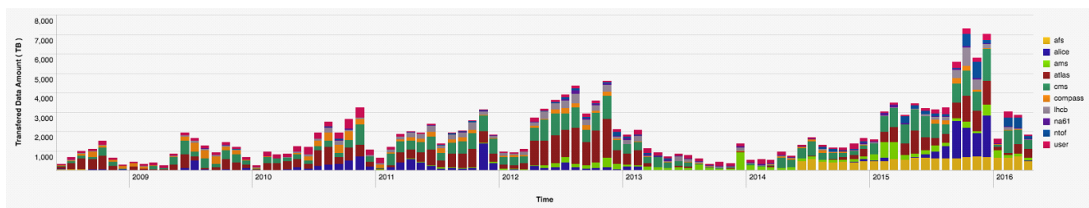
**Figure 2:** Data transfers from experiments to CERN Tier-0 during Run1 and Run2 , in TB/month. Collected during Run1: 15 PB/2010, 23 PB/2011, 27 PB/2012; during the 1st year of Run2: 32 PB [8].
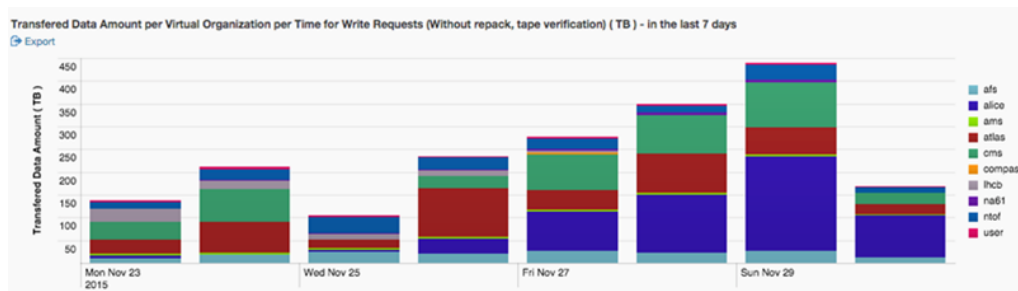


**Figure 3:** Data transfers from experiments to CERN during one week of the HI run in 2015, in TB/day.

The provided computing resources comprise up to 600 thousand of CPU cores which represents a performance of almost 7 million HEP-SPEC-06 [5] (special benchmarking units used by the High Energy Physics (HEP) community). The total provided online disk storage capacity is $\sim$ 320 PetaBytes (PB) and the total tape storage capacity is $\sim$ 390 PB. WLCG is coordinated by CERN and is the world's largest computing grid.

## 3. WLCG performance during Run 2

The year 2015 and the first months of 2016 were a commissioning period for the LHC. There was a one-month Heavy Ion (HI) run at the end of 2015 and p-p running during most of the operations. The original goal for 2015 was to reach 8-10 fb$^{-1}$ of integrated luminosity at 13 TeV c.o.m. p-p energy, 25 ns spacing, 2800 bunches and a maximum pile-up of 40 [6]. Due to various causes the final integrated luminosity achieved in 2015 was 4.3 fb$^{-1}$. The integrated luminosity during the HI run in 2015 (3 weeks of data taking for physics) was almost 700 pb$^{-1}$ [6]. Even so the volume of data collected by the LHC experiments in 2015 reached 32 PB, exceeding the data volumes of Run 1 (see Fig. 2).

In 2015 data rates from the experiments to the CERN Computing Center reached $\sim$ 8 GB/s sustained for 8 hours, with peaks $\sim$ 10.5 GB/s [7]. Per day, the data rates reached 600 TB, see Fig. 3 (previous maximum was $\sim$ 220 TB/day) [8]. In Fig. 4, the delivery of computing power for Run 1 and Run 2 is plotted. One can see a ramp-up almost by 1/3 in 2015 [8]. In Fig. 5 the data transfers from CERN to external sites are shown for the last 3 months of 2015.
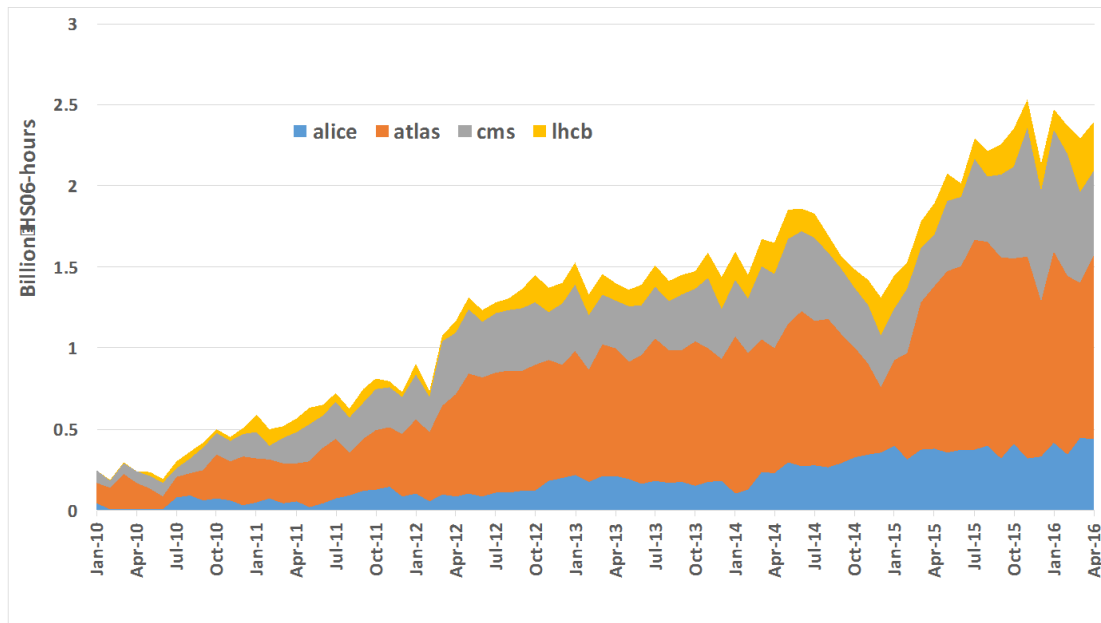
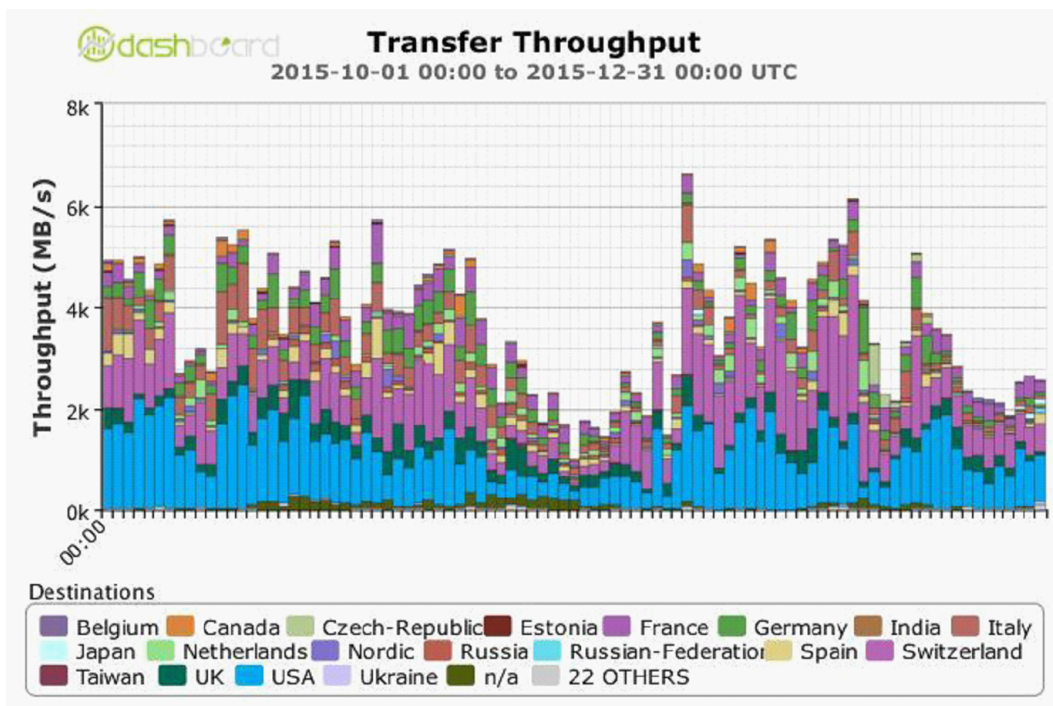**Figure 4:** Delivery of CPU resources during Run1 and Run2, in billion HepSpec06.



**Figure 5:** Data transfers form CERN to external sites during the last 3 months of 2015, in MB/s. Peaks up to 6 GB/s.

3

## 4. What is different in Run 2 in comparison with Run 1

The LHC computing grid is in continuous operation processing the data from the LHC, but in addition there are ongoing activities aiming to improve the system in multiple respects. These activities were envisaged already at the end of Run 1 when it became clear that the ever growing need for hardware, software and human resources cannot simply be covered by the flat budgets of the participating institutions. There is an intensive work in progress especially in the following areas:

1. *Moving to more flexible computing models:*
   Assigning less specific roles of different Tiers, more even distribution of tasks (reconstruction, simulation, analysis). The computing grid for LHC was originally built along the MONARC [9] model, where each participating site is ranked by its size and is assigned concrete tasks accordingly. So the large sites of the rank Tier-1 (14 of these in WLCG) were supposed to perform mainly reconstruction passes, while the smaller Tier-2 sites (150 in WLCG) were used for simulation and end-user analysis. This strict assignment of tasks was re-considered to better utilize the computing resources, e.g. to run simulation or organized-analysis campaigns also at Tier-1s.

2. *Use of HLT farms of the experiments for data processing during LHC technical stops and machine development periods:*
   The experiments further increased their large High Level Trigger (HLT) computing clusters for Run 2 providing tens of thousands of CPUs. During breaks in data taking these large resources can be used for production of simulated events, for an example see [10].

3. *Use of clouds and other opportunistic and commercial resources:*
   Since a few years a number of WLCG sites have been investigating how to use commercial cloud services to extend resources available for the processing of ever-growing LHC data volumes [11]. There have been several CERN procurements aimed at evaluating the possibility, costs, conditions and overall performance of processing WLCG jobs in the clouds.
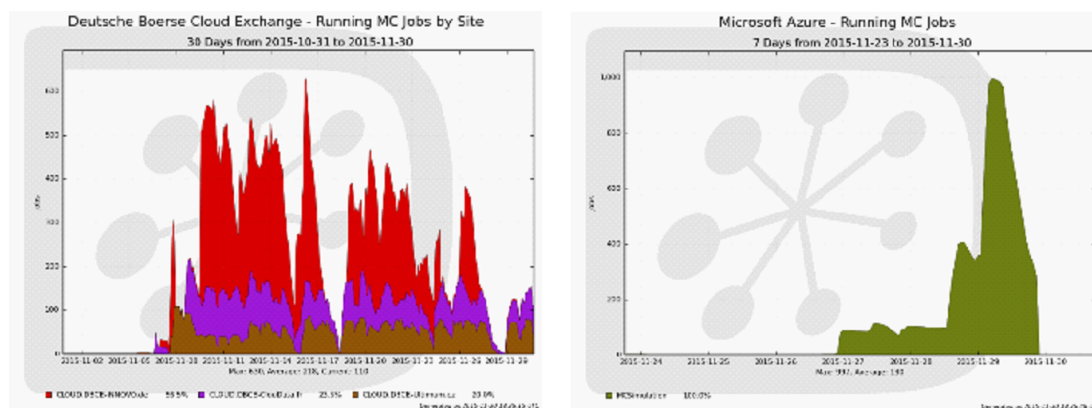


**Figure 6:** Use of the Deutsche Börse Cloud Exchange and Microsoft Azure cloud resources for LHCb simulation campaigns in 2015.

The evaluations began with Monte Carlo simulation campaigns using commercial cloud services [12]. Later the scope was widened to the full chain of data processing with a special focus on cloud storage and network throughput [13]. In Fig. 6, the use of the Deutsche Börse Cloud Exchange and Microsoft Azure cloud resources for LHCb simulation campaigns in 2015 is shown [14]. Figure 7 shows the number of events of ATLAS Geant4 simulation processed in the ATOS cloud during a several weeks campaign in 2015 [13]. The latest procurement of commercial cloud services has started in the spring of 2016. In addition to commercial clouds, other opportunistic resources can be obtained by Volunteer Computing [15] or usage of high-performance computing (HPC) center resources [16].

4. *Data popularity analysis and clean-up of unused data:*
   Because of increased pressure on available disk capacity, the experiments started to monitor more stringently the access to and use of different data sets, the so-called data popularity. There are regular clean-up campaigns following a careful analysis of data usage to obtain additional disk space. For example, the ALICE experiment analysis of the access to data showed that a small percentage of the data had not been accessed at all during 2015.

5. *Network access to data:*
   The computing models of the experiments were using the pattern of sending jobs to sites which had the data to be processed or copying datasets to the site where the corresponding jobs should run. Datasets turned out to get copied excessively to different sites, stressing the disk capacities. In general, moving data around is expensive, it needs additional disk and network resources. Although the rule that jobs should be sent where data are is still
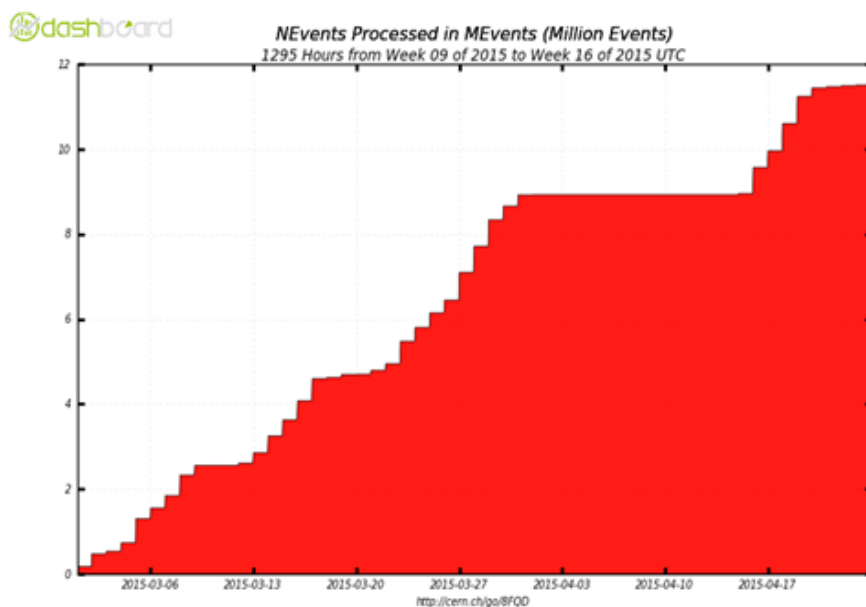


**Figure 7:** Number of events of ATLAS Geant4 simulation processed in the ATOS cloud during a several weeks campaign in 2015.

followed, the extensive data replication is no longer the norm. Instead, the WAN (Wide Area Network) access to data is increasingly used to stream data to jobs over the network, in case it is not available locally. The goal is to access any data from any site without the need to copy it first. However, as the overall network throughput is not yet sufficient, currently most of the processing still needs to be performed inside the centers holding the data or nearby via fast networks. Achieving efficient data access becomes quite difficult when many independent storage sites are involved, and therefore different types of data federations are under investigation. On a regional level, these represent groups of sites where datasets are held, connected via high capacity networks and allowing fast and reliable access to data without the need to copy it. For instance the CMS experiment is utilizing a framework based on the design principle "Any Data, Any Time, Anywhere" (AAA), which presents a global, unified view of storage systems – a "data federation" [17].

6. *Enhanced experiment software performance:*
The ever-growing data volumes to be processed together with advances in hardware technologies is driving the experiments to improve the efficiency of their software and to adapt it to new technologies, especially to multicore processors [18]. As a result, the experiments were able to speed up the simulation (up to factor 2) and improve the performance of reconstruction significantly (up to 4 times). Much software was re-engineered to utilize multicore processors with a much lower memory footprint per job [19].

7. *Other optimizations:*
One of the challenges for the LHC experiments is the calibration procedure or strategy. During Run 1, complicated calibration procedures required a number of re-processing cycles of the raw data, which occupied a significant part of the computing resources. Towards Run 2, the efficiency of calibration procedures improved substantially and as a result, fewer reprocessing cycles are needed. For instance the calibration and alignment procedures became faster in the ALICE experiment and online calibration was adopted by LHCb.
In the beginning of the data taking, also the analysis frameworks of the experiments were not tuned yet. The data formats were not always adequate for the analysis use cases and there were many individual physics groups and users performing chaotic, I/O-intensive analysis passes. This resulted in a low performance of analysis jobs and inefficient utilization of computing resources. Again, towards Run 2, the analysis frameworks were improved. This included new optimized data formats and use of organized (in contrast to chaotic) analysis, so called Analysis Trains [20] and the analysis efficiency improved significantly.
Large efforts were invested into improvements of the network system for the LHC data. For instance, the LHC Open Network Environment (LHCONE) [21], with the mission to provide a collection of entry points into a network that is private to the LHC Tier-0/1/2 sites, has been continuously extended involving a growing number of WLCG sites at all levels. Also, the transatlantic connectivity was upgraded significantly thanks to the US ESNet consortium [22] which provided 100 Gb/s links connecting the key sites in Europe and in the US.

## 5. Longer term prospects for the LHC computing

In the expectation of an exponential data growth towards the HL-LHC, the WLCG is in preparation for the future LHC runs. In Fig. 8 estimates of the data volumes produced by the LHC are shown.

The basic question is, what physics can be done with the computing which the LHC community can afford? It is almost impossible to make predictions 20 years ahead given the frantic developments in IT and the increases in the costs of HEP computing. During the LEP era, computing was affordable on commodity platforms. In the LHC era, the global yearly cost of WLCG hardware is approximately 100 million CHF/USD/Euro. Currently, the 5-year cost is almost the same as the construction cost of the ATLAS or CMS detectors. Most of the costs concern the continual yearly replacements, because the equipment has a lifetime of 3-5 years.

The costs include CPU, disks, tapes, networks, electricity, (cooling) facilities, system administrators. The costs of facilities and power motivate an idea that commercially provisioned computing may soon be more cost effective for HEP. For the growth of storage media costs some predictions are shown in Fig. 9. High energy physics can hardly go on without the distributed computing. But the support load and cost would be too high without the developments and changes already mentioned in the previous sections. Other important developments include:

1. A re-implementation of experiment software for reconstruction, simulation and analysis to work on CPU architectures that will be much less expensive (e.g., ARM or GP-GPU [23]).

2. First attempts to use Machine Learning [24] technology. All the LHC experiments are interested, e.g., in detection of anomalous behavior in production workflows and in data manage-
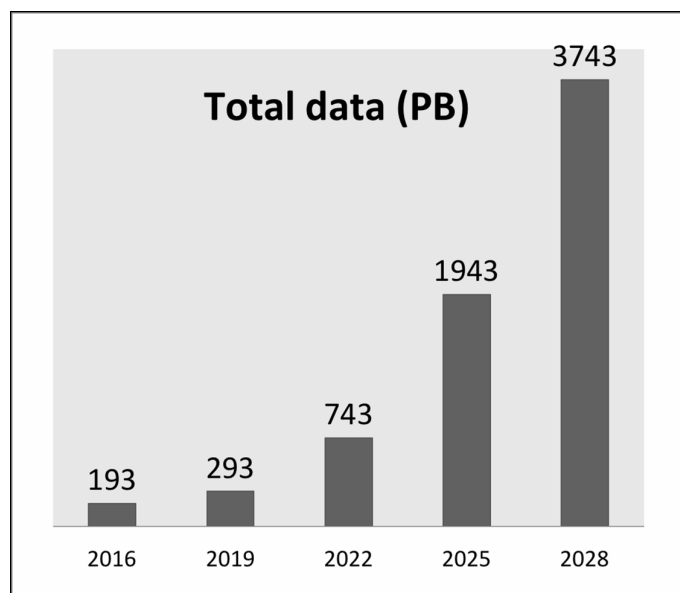


**Figure 8:** Predictions of the LHC data taking beyond 2016.

ment systems. It would be worthwhile to investigate common approaches in several experiments. Some work in this area was already done in ATLAS and LHCb.

WLCG was created to handle the data produced by the LHC. There exist, however, close collaborations with other international e-infrastructures which support other HEP experiments and other sciences. HEP experiments like Belle-II, Nova, AMS and others can profit from the WLCG infrastructure. It does not really mean the compute and storage resources, since most experiments have their own funds and resources, but the expertise developed by the WLCG that is unique and may also allow others to benefit from its structure: policies, networks, operations, support etc. In addition, many sites providing services for WLCG are used also by non-LHC projects, including not only HEP experiments but also, e.g., CTA, SKA and others. It is a common interest to have the infrastructures shared as much as possible, otherwise the support load and cost will be beyond our reach.

## 6. Summary and Outlook

The start-up of the LHC Run 2 was overall successful and the WLCG showed a very good performance, handling the LHC data without unexpected barriers while achieving much higher data volumes and data rates. These volumes and rates will grow further during the rest of Run 2, but there are ongoing activities and developments – as described in this article – to facilitate successful processing of all LHC-produced data.

WLCG is the world's largest computing grid and is one of the e-infrastructures built in Europe over the past decade. During this period, Europe developed world leading expertise in building and operating very large scale research infrastructures supporting international collaboration in
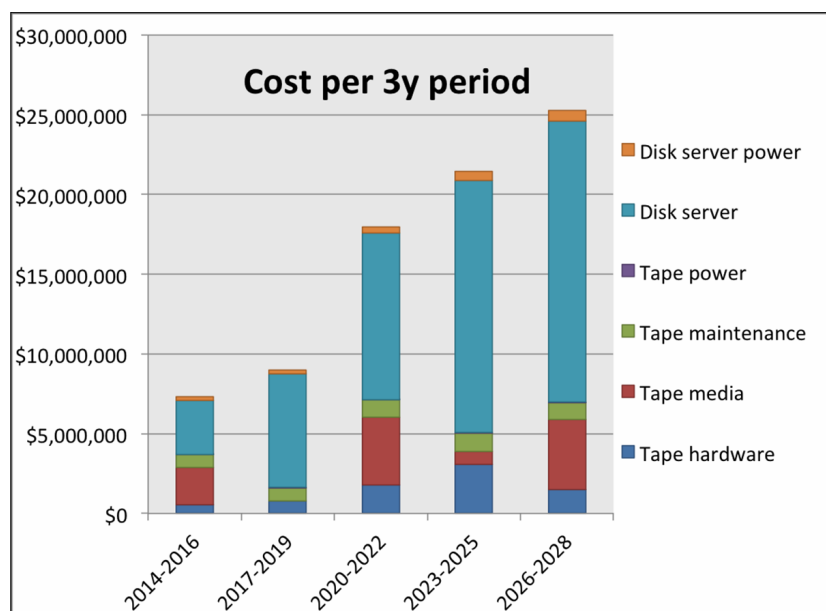


**Figure 9:** Prediction of the growth of storage media costs.

sciences [12]. With the expansion of these infrastructures, we have entered the era of explosive data growth. The endeavor to handle the "Big Data" motivated a transition from simple commodity computing ("Farms") to commodity distributed computing ("Grid") and more recently to commodity computing services ("Cloud").

Driven by the need to expand the capacity of IT services for the LHC experiments and other customers, CERN achieved a leading role in the effort to procure, develop, deploy and test the tools for interaction between commercial clouds and publicly funded e-infrastructures. These aspects are being investigated as a part of the Helix Nebula initiative [25] that was formed by CERN, EMBL [26] and ESA [27]. CERN also published a white paper on future IT challenges in scientific research [28].

The experience of recent years shows that the most likely roadmap to acquire the necessary IT resources involves inter-operation between WLCG and commercial clouds. Looking at the underlying trends in the cloud services market, the use of cloud computing for handling the LHC data is a matter of "when", not "if" [29]. The ongoing collaboration between CERN and European commercial cloud service providers shows that cloud computing in principle has the potential to reduce IT expenditures. However, the use of cloud computing services by public research organizations is complicated by many hurdles, e.g., technical standards and trustworthiness, which will take time to overcome [12].

In conclusion, the future challenges created by the growing data rates and data volumes from the LHC experiments can be successfully resolved combining various strategies described in this article. It seems beyond doubt that the inter-operation between WLCG and commercial clouds will have an important role. Nevertheless, although commercial clouds are a buzzword today, future technologies may bring yet better ways to get the most out of our data processing infrastructures!

## References

[1] The Large Hadron Collider at CERN; `http://lhc.web.cern.ch/lhc/`

[2] Worldwide LHC Computing Grid: `http://wlcg.web.cern.ch/`

[3] LHC Computing Grid: Technical Design Report, `http://wlcg.web.cern.ch/documents-reference`

[4] WLCG Public Page, `http://wlcg-public.web.cern.ch/`

[5] HEPSPEC2006-benchmarking of batch machines, `http://w3.hepix.org/benchmarks/doku.php`

[6] see e.g. F.Brody: *LHC Machine Status Report*, 125th LHCC Meeting-OPEN Session, March 2nd 2016, `http://indico.cern.ch/event/493464/`

[7] I.Bird: *WLCG Update*, LHCC Referees Session on Computing, December 1st 2015, `http://indico.cern.ch/event/461711/`

[8] I.Bird: *Status of the WLCG project*, LHC RRB Meeting, April 2016, `http://indico.cern.ch/event/503173/`

9

[9] The MONARC model,
http://monarc.web.cern.ch/MONARC/docs/monarc_docs.html

[10] see e.g. E. Lançon and S. Campana: *ATLAS computing report*, LHCC Referees Session on Computing, December 1st 2015, http://indico.cern.ch/event/461711/

[11] see e.g.. *Big science teams up with big business to kick-start European cloud computing*, CERN press office release March 1st 2012,
http://press.cern/press-releases/2012/03/big-science-teams-big-business-kick-start-european-cloud-computing

[12] *Towards the European Open Science Cloud*, Prepared by Bob Jones (CERN IT department) on behalf of CERN, March 2015.
https://zenodo.org/record/16001

[13] C.Cordeiro: *Integrating Commercial Clouds into the WLCG*, CERN IT Technical Forum, January 22nd 2016, https://indico.cern.ch/event/470426/

[14] S.Roiser: *LHCb Computing2015 Q3 Report*, LHCC Referees Session on Computing, December 1st 2015, http://indico.cern.ch/event/461711/

[15] LHC@home, volunteer computing for the LHC,
http://lhcathome.web.cern.ch/about/what-volunteer-computing
BOINC volunteer computing, https://github.com/BOINC/boinc

[16] High Performance Computing (HPC), see e.g., http://hpcinnovationcenter.llnl.gov/

[17] K.Bloom et al: *Any Data, Any Time, Anywhere: Global Data Access for Science*,
http://arxiv.org/abs/1508.01443

[18] WLCG Multicore Deployment Portal,
https://wlcg-ops.web.cern.ch/ops-coordination-task-forces/

[19] I.Bird: *Status of the WLCG Project*, LHC RRB Meeting October 2015,
https://indico.cern.ch/event/407749/other-view?view=lhcrrb&showSession=2

[20] Physics Analysis in the ALICE experiment, http://alice-analysis.web.cern.ch/

[21] LHCONE, http://lhcone.net/

[22] Energy Sciences Network (ESNet), https://www.es.net/about/

[23] About GPGPU, http://gpgpu.org/about

[24] Machine Learning: http://machinelearningmastery.com/

[25] Helix Nebula, The Science Cloud, http://www.helix-nebula.eu/

[26] European Molecular Biology Laboratory, http://www.embl.de/

[27] European Space Agency, http://www.esa.int/ESA

[28] CERN openlab Whitepaper on Future IT Challenges in Scientific Research, May 2014,
https://zenodo.org/record/8765#.V44o8BJGZD0

[29] Ian Bird, the head of the WLCG project, e.g. in *Symmetry*, December 21st 2011,
http://www.symmetrymagazine.org/breaking/2011/12/21/happy-10th-birthday-wlcg