Computing framework for the LHC: current status and challenges of the High Luminosity LHC future

Dagmar Adamova*

Nuclear Physics Institute, Czech Academy of Sciences
E-mail: adamova@ujf.cas.cz

During the first years of the LHC operations the computing infrastructure for the LHC data handling performed extremely well and enabled fast delivery of spectacular physics results. Towards the end of this period (Run 1), the LHC produced almost 30 PetaBytes (PB) of RAW data in one year and the computing and storage resources available within the LHC Computing Grid (WLCG) were used up to the limits. The ongoing LHC upgrade for higher luminosities and energies (HL-LHC) will bring data volumes dramatically larger than during Run 1. The current rate of Grid growth will not meet the new demands. For the upcoming HL-LHC, the experiments must adapt their computing models, speed up simulations, use parallelization at all levels and re-engineer experiment software. In addition, the LHC computing must reach out and use external resources and non traditional solutions like Big Data, High Performance Computing centers and Clouds. In this contribution, we will give an overview of the current status of the LHC computing infrastructures and the revisions under way in preparation for the HL-LHC.

---

*Speaker.

1. Introduction

The period of November 2009–February 2013 marked a unique event in the history of High Energy Physics (HEP): the Large Hadron Collider (LHC) [1] at CERN was in operation delivering data from proton-proton (p-p), proton-lead (p-Pb) and lead-lead (Pb-Pb) collisions (so called Run 1). The total volume of RAW data collected during this period reached $\sim$ 75 PetaBytes (PB). To store and process the LHC data a dedicated distributed infrastructure has been designed and built, the LHC Computing Grid (WLCG) [2]. It facilitated a fast delivery of amazing scientific results, of which the most significant was the discovery of Higgs boson [3] and the technologies developed during the WLCG building process were subsequently used in computing Clouds [4].

In this paper we present a short overview of the WLCG current status and its performance during Run 1 and the challenges which the WLCG will be facing in the future, in the era of the High Luminosity LHC (HL-LHC) [5].

2. Worldwide LHC Computing Grid: Current status

WLCG is a distributed infrastructure of about 150 computing centers in 40 countries worldwide, spanning over 6 continents (see Figure 1) with a Tier-hierarchy structure where CERN is the largest central Tier-0 site [2]. WLCG provides a computing power of over 300 thousand of computer cores (CPU) which represents a performance of $\sim$ 2 million HEP-SPEC-06 [6]. The biggest site provides $\sim$ 60 thousand of CPU cores and 12 sites (Tier-1) provide 2-30 thousand CPU cores. The total online disk storage capacity provided is $\sim$ 270 PB and the total tape storage capacity is $\sim$ 195 PB. The WLCG is the world's largest computing grid.
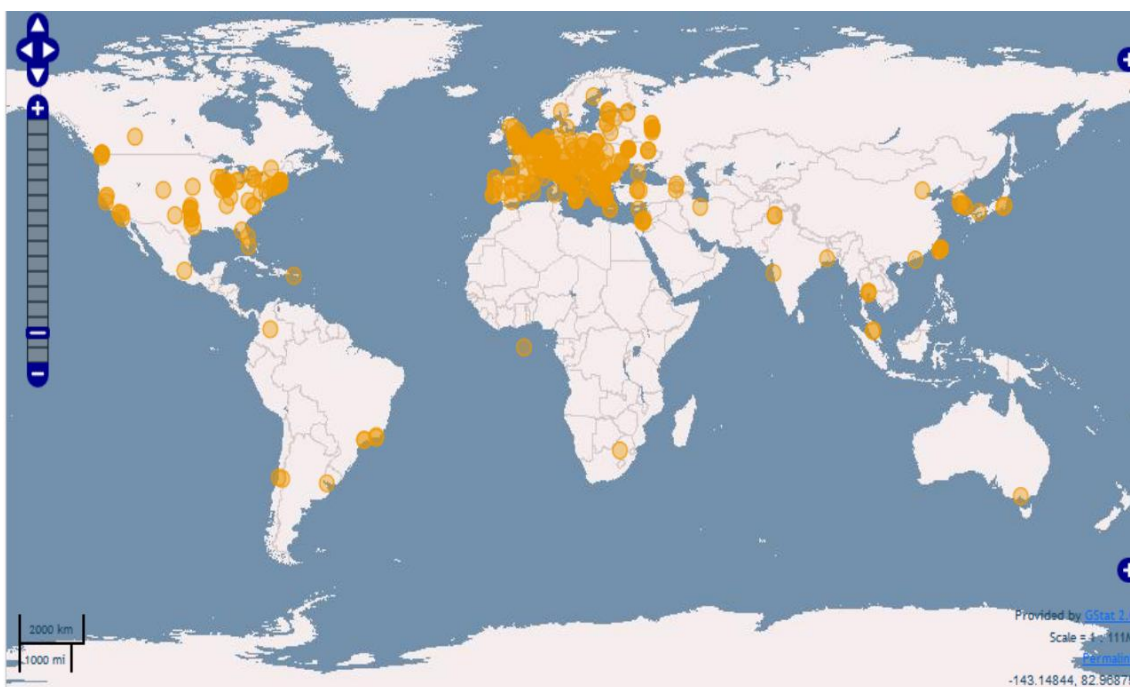
Figure 1: WLCG resources: spanning over 6 continents

3. Run 1: most remarkable results and performance of WLCG

During Run 1 (2009–2013) the accelerator delivered billions of events to the experiments from p-p, p-Pb and Pb-Pb collisions, with 50 ns bunch spacing and the pile-up up to 35. The luminosity delivered in years 2011+2012 in p-p collisions to ATLAS/CMS at energies $\sqrt{(s)} = 7/8$ TeV and used for physics was ~ 25 fb^{-1}. The plot of integrated p-p luminosity delivered in 2011+2012 is in Figure 2. On the 4th of July 2012, the Higgs boson 5σ discovery was announced [3], based on use of 10.7 fb^{-1} of data collected till June 18th 2012.

The enormous delivered luminosity resulted in high volumes of RAW data collected by the LHC experiments. The data was archived at the CERN tape storage system CASTOR [7] and the rates were 15 PB in 2010 (p-p and Pb-Pb), 22 PB in 2011 (p-p and Pb-Pb) and 27 PB in 2012+2013 (p-p and p-Pb). The total archive of the LHC and derived data on the CERN CASTOR reached ~ 75 PB. In addition, the WLCG infrastructure managed derived data: output of simulation campaigns, reconstruction and all kinds of analysis which represented additional tens of PB of data for each experiment.

The WLCG performance during Run 1 was remarkably stable and reliable. The data recorded by the experiments was processed within hours, which enabled the publication of scientific results within a couple of weeks after the data was recorded.
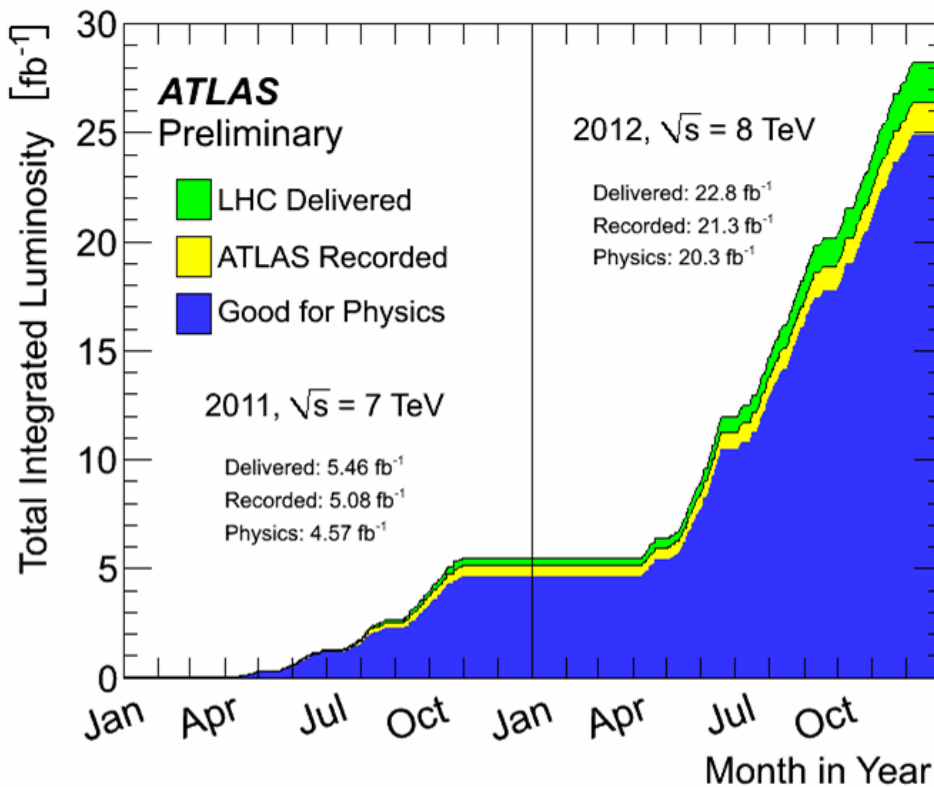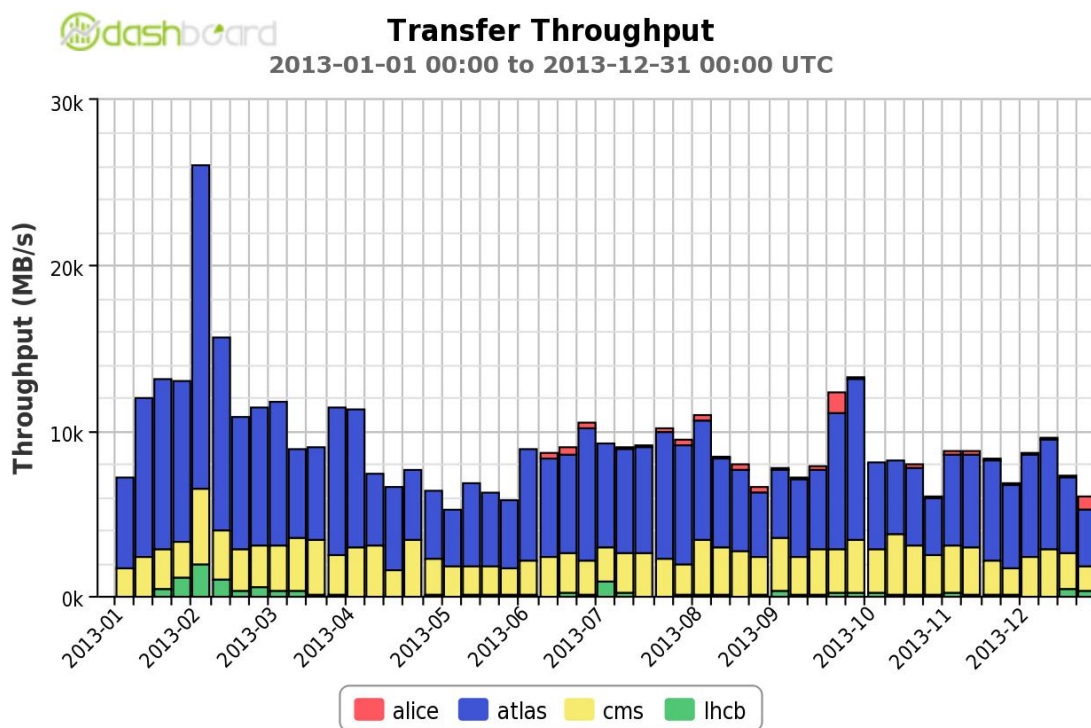

Figure 2: Integrated p-p luminosity 2011+2012

Figure 3: Average data transfer rates per week within the WLCG network system in 2013 (p-Pb + LS1), see [2]. Timescale in weeks on x-axis, data rates in MB/s on y-axis. There is a significant peak at the end of Run 1 in February 2013.

4. Activities during LS1

Now we are in the middle of the Long Shutdown 1 (LS1): the LHC is not in operation going through an upgrade for higher energies and luminosities. The computing infrastructure is in operation without any break with various different activities ongoing:

- Re-processing of data produced in Run 1

- Additional simulation campaigns

- Testing use of High Level Trigger (HLT) computing farms for offline processing as additional Tier sites

- ATLAS and CMS use Openstack (open source Cloud software) [8] to manage their HLT farms

- Optimizing usage of available resources in preparation for the next Run 2 (the system needs to continue operating while being improved)

Global data transfer rates within the WLCG network system during 2013 are shown in Figure 3: there is no break in the WLCG operations.

Run 2 (the upcoming period in the LHC operation expected to start in the beginning of 2015) will bring two times higher energy, pile-up and the number of recorded events. With the improvements and the regular upscale of resources done during LS1 the WLCG infrastructure should be capable of successful processing of the produced data, without fundamental changes.

5. Outlook for beyond Run 2

The future periods of the LHC operations, Run 3 expected to start in 2018/2019, and the next future runs will bring significantly enhanced luminosities (so called High Luminosity LHC). Run 3 should bring integrated luminosities of 250–300 fb^{-1} in one year and ~ 3000 fb^{-1} should be collected in about 10 years during several runs (the luminosity which should bring a chance to discover New Physics phenomena at p-p energy $\sqrt{(s)} = 14$ TeV [9]).

The enormous luminosities will show up in a significant increase of the volume of produced LHC data. The anticipated RAW data volume in Run 3 is ~ 130 PB/year, in Run 4 several 100 PB/year. In addition, there will be derived data, as mentioned in Part 3. The experiments current computing models do not scale accordingly and the standard Grid resources growth of 25% per year will not be sufficient to provide enough capacity for processing of such amounts of data. The anticipated data growth is shown in Figure 4.
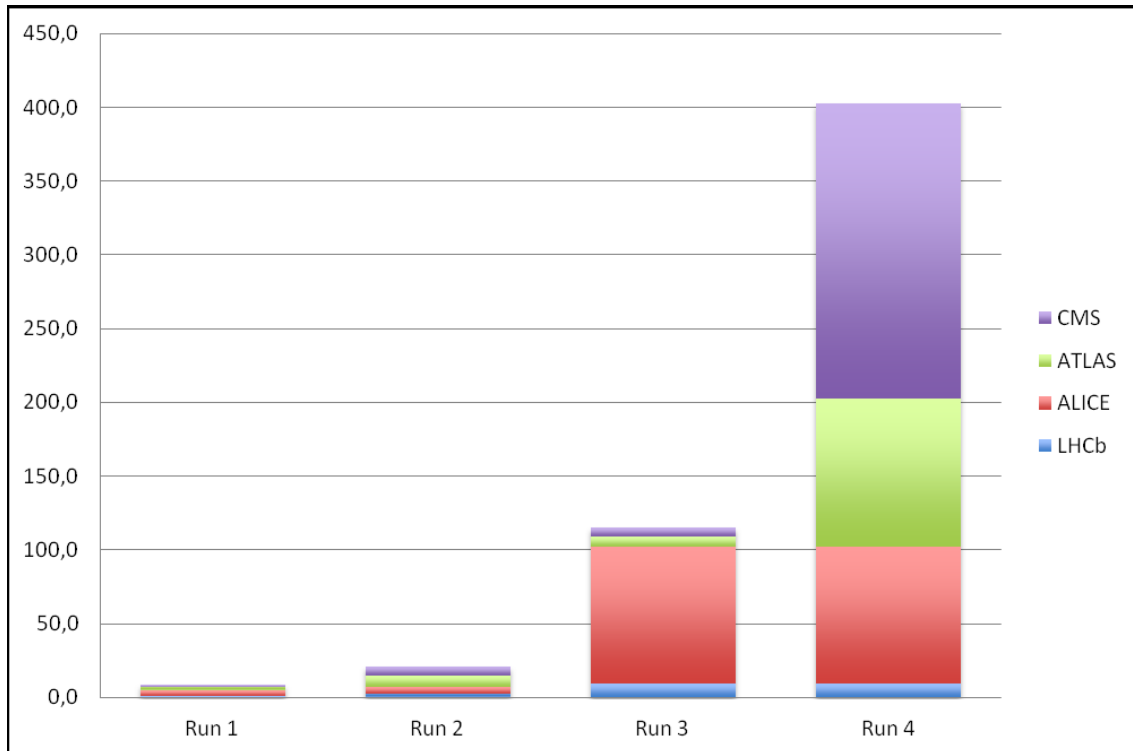
Figure 4: Data volumes expectation for HL-LHC (in PB)

6. Facing challenges of HL-LHC

To guarantee enough resources for Run 3 and beyond, there are activities in various different areas:

1. Update of the LHC experiments Computing models

   - Re-engineering experiment software to be able to make use of the new multi-core CPU and parallelism at all levels
   - Use of HLT farms for offline data processing

2. Re-structuring of Grid middleware layers

   - Open source Cloud technologies give a way to implement job submission and management on Cloud cluster, see e.g. [8]
   - Migration to more Cloud-like models

3. Improvements in Data management (the key issue)

   - Working towards transparent distributed data access enabled by today's efficient networks
   - Setting up data federations [10] (grouping storage clusters of a number of computing sites into one system) based on Xrootd and/or http – there are many commonalities between experiments
   - Optimizing data access from jobs: use of remote access, remote I/O
   - Introducing more intelligent data placement and caching
   - Developing data popularity services: data access monitoring shows that some data samples are much more frequently processed than other. Creating more replicas of the "popular" data samples makes the data processing more efficient

4. Use of opportunistic resources

   - An intensive effort to get access to High Performance Computing (HPC) [11] or Supercomputer centers worldwide for simulation campaigns
   - Tests of using commercial Clouds for simulations: ongoing, but the price is still too high for regular campaigns

7. Summary and Outlook

The WLCG project was approved in 2001 and the basic prototype and the Technical design report appeared in 2005. At that time, about 20 PB of data expected to be produced by the LHC in one year represented an enormous quantity. Tools for distributed data processing and management capable of handling such volumes of data were basically non-existent. The WLCG collaboration developed their own tools and services for handling the LHC data.
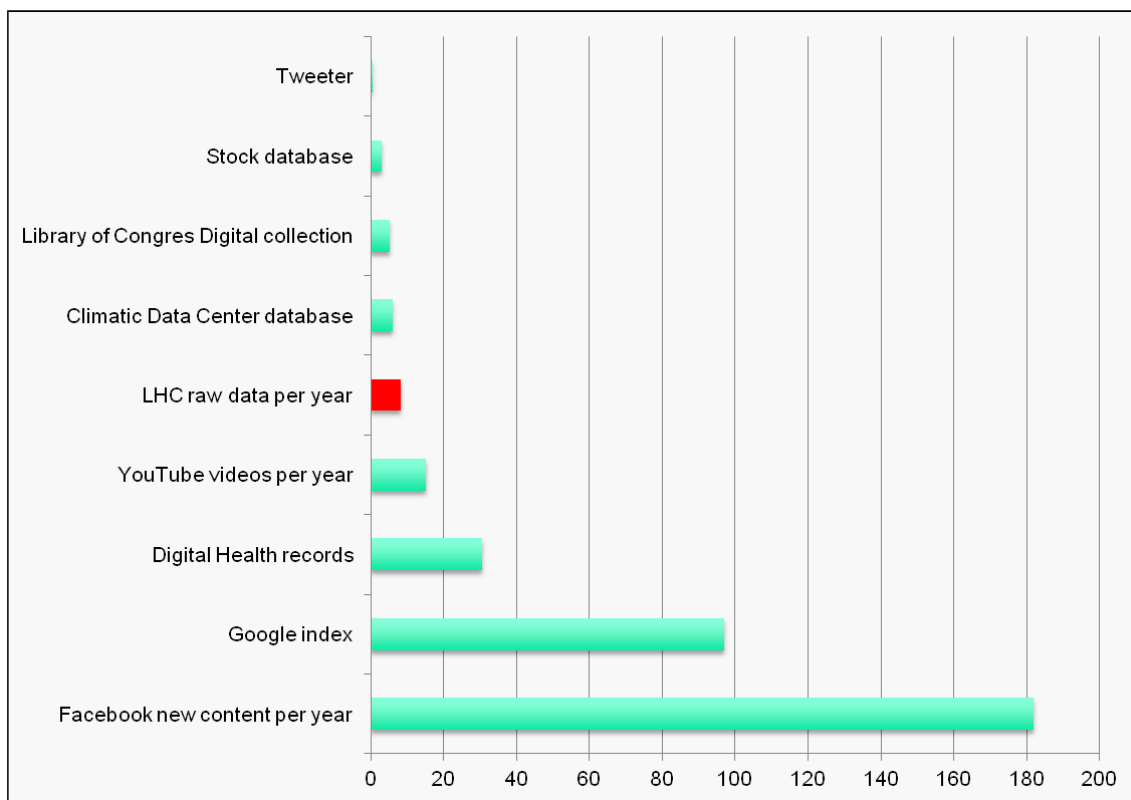
Figure 5: Large data producers

In the last 5–8 years, thanks to fast development of digital technologies, the price of creating, transferring and storing information/data dropped dramatically. Due to growing speed of network connections and availability of devices like smart phones, tablets or netbooks the volume of produced and stored data worldwide truly exploded in the commercial area, overtaking big science (see e.g. Figure 5). While in 2005 the data produced annually worldwide did not reach 1 ExaByte (1 thousand PB), in 2010 it reached 1 ZettaByte (1 million PB), see Figure 6. Along with this data explosion, technologies for storage and analysis of large and heterogenous data were developed for commercial use (like e.g. Hadoop [13]).

To what extent it will be possible for the WLCG collaboration to adopt the tools and solutions from the commercial world is not yet established. There are commonalities but the commercial solutions for handling large data are not necessarily suitable to solve the problems which the WLCG will be facing for Run 3 and beyond. But WLCG-specific and experiment-specific solutions will for sure be used together with some existing off-the-shelf products, in contrast to the situation before Run 1.

Acknowledgements
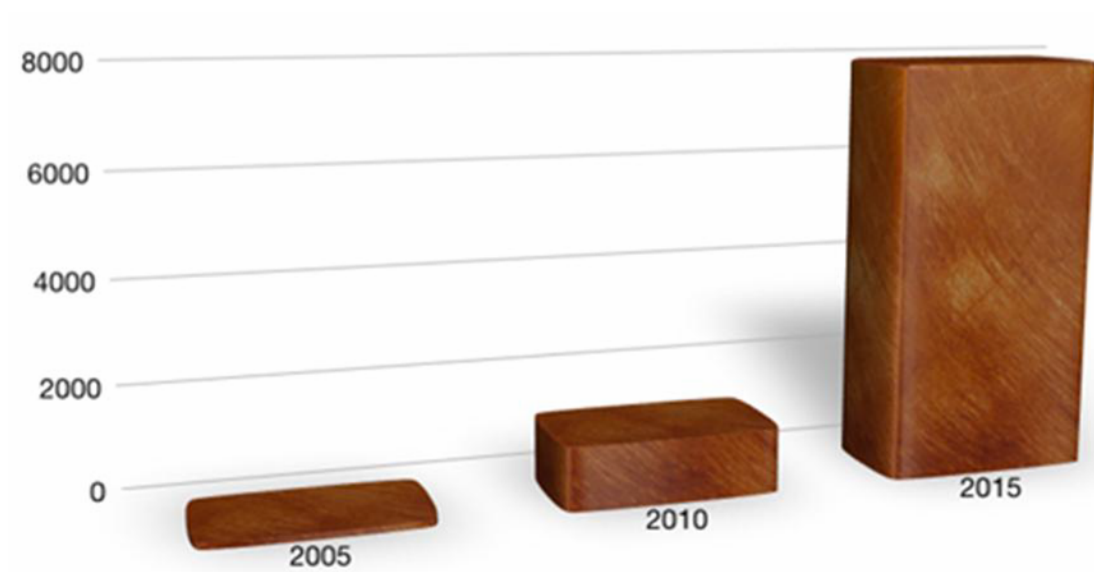
Figure 6: A Decade of Digital Universe Growth: Storage in Exabytes [12]

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] ATLAS Collaboration, *A Particle Consistent with the Higgs Boson Observed with the ATLAS Detector at the Large Hadron Collider, Science* **338** (2012) 1576-1582; CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, Physics Letters* **B716** (2012) 30–61.

[4] I. Foster et al, *Cloud Computing and Grid Computing 360-Degree Compared, Proc. of the Grid Computing Environments Workshop* 2008, GCE .08, Austin, Texas; `http://arxiv.org/ftp/arxiv/papers/0901/0901.0131.pdf`

[5] HL-LHC: High Luminosity Large Hadron Collider, `http://hilumilhc.web.cern.ch/HiLumiLHC/index.html`

[6] HEPSPEC2006 benchmarking of batch machines, `https://twiki.cern.ch/twiki/bin/view/FIOgroup/ProcRefHepSpec2006`

[7] The CERN Advanced Storage manager (CASTOR), `http://castor.web.cern.ch/`

[8] Open source software for building private and public clouds, `https://www.openstack.org/`

[9] Fabiola Gianotti, *The Physics Landscape*, `http://indico.cern.ch/event/260492/material/slides/1?contribId=4&sessionId=1`

[10] R. Gardner et al, *Data Federation Strategies for ATLAS using XRootD*, `http://cds.cern.ch/record/1622223`

[11] TOP500, `http://en.wikipedia.org/wiki/TOP500`

[12] J. Gantz and D. Reinsel, *Extracting Value from Chaos*, `http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf`

[13] Apache Hadoop, `http://hadoop.apache.org/`