ATLAS and DØ Computing at the GridKa Computing Centre

G. Duckeck, J. Elmsheuser, T. Nunnemann, and P. Schieferdecker

In January 2002 the "GridKa" computing centre [1] was founded at the Forschungszentrum Karlsruhe. Its main purpose is to provide the computing and storage capacity required to analyze the enormous data volume of the LHC experiments, in cooperation with similar centres in other countries. Until the start of LHC in 2007 GridKa is gradually build up and extended. It serves test productions of the four LHC experiments, as test site for Grid development and as computing centre for four running experiments, BaBar, CDF, Compass and DØ.

1. Atlas Data Challenge 1

In the years before LHC start-up Atlas performs a series of "data challenges" (DC) [2]. These are test productions, i.e. performing simulation and reconstruction of events, with increasing complexity and volume and aim to achieve several goals:

- Test and validation of the full Atlas offline software chain: Event generation, detector simulation, reconstruction, analysis.
- Generation of large samples of simulated events for physics and detector studies
- Test and deployment of the worldwide distributed computing using Grid tools as they become available.

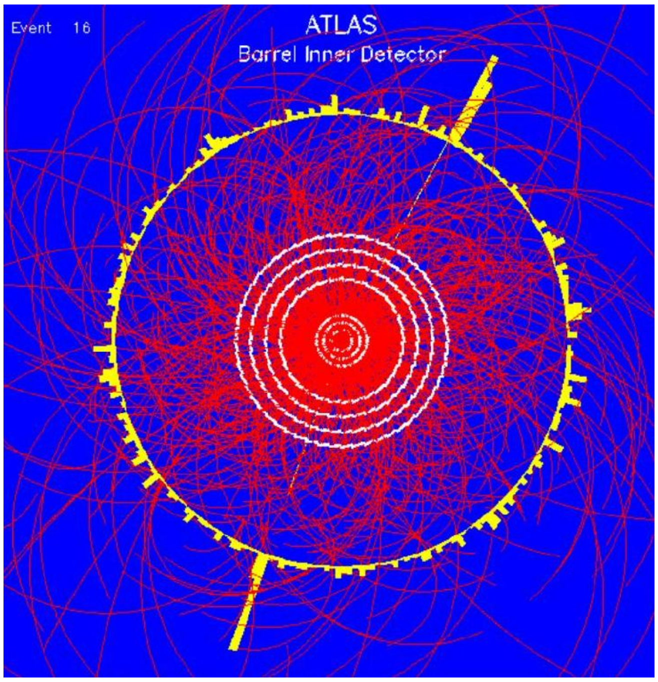

Fig. 1: Example for pile-up event: A Higgs decay into two photons hidden by many tracks and clusters from minimum-bias events.

The first Atlas Data Challenge (DC1) started in July 2002. Using 37 computing centres worldwide 10 million events have been generated adding up to a total volume of 30 TB. The GridKa centre operated smoothly and efficiently and contributed nearly 10 % to the produced volume, which was eventually delivered to the physics and detector-performance groups.

In a second step the "pile-up" was simulated: the overlaying of many so-called minimum-bias events (inelastic pp-scattering) simultaneously in the detector, as it will happen at the high luminosity foreseen at LHC. The proper simulation of pile-up is rather challenging: several data streams need to be read in parallel and many 100 events must be kept in memory in order to do the proper merging of events at hit level, before the digitization takes place. GridKa participated successfully also in this phase.

2. A Regional Analysis Center for DØ

The GridKa Computing Center with its CPU and storage resources is being employed as the first prototype remote analysis center for DØ, which provides CPU and storage resources for both Monte Carlo simulation and analysis of data collected with the DØ detector at Fermilab [3]. For the large-scale data transfer of the entire DØ dataset in a compressed format, the SAM facility and Globus authentication methods are used [4]. The data are imported and stored at a rate of $\sim$ 1 TB per month using a large disk cache and a tape archive, which are also employed as a data server for German institutes participating in the DØ experiment. SAM provides an efficient way of data set definition and distribution, e.g. if a data set is requested by an analysis job running locally in Munich, it is automatically imported – in the given order – either via a local cache, GridKa, Fermilab, or any other data cache. To minimize the transatlantic transfer, all data which are requested by a German institution are routed through and cached at GridKa.

For data reconstruction and analysis, DØ's software distributions have been installed and adapted. The DØ collaboration plans to set up a global distributed computing network (Grid), in which GridKa will contribute as a major computing center.

References

[1] *Requirements for a Regional Data and Computing Centre in Germany*, http://grid.fzk.de/grid/LHCComputing-1july01.pdf (2001).

[2] G. Poulard, *Atlas data challenges*, CHEP03, (2003).

[3] I. Bertram *et al.*, *Proposal for DØ Regional Analysis Centers*, DØ-Note 3984 (2002).

[4] A. Baranovski *et al.*, *SAM managed cache and processing for clusters in a worldwide grid-enabled system*, FERMILAB-TM-2175 (2002).