T 15: GRID Computing

Time: Monday 16:15–18:00 Location: T-H28

T 15.1 Mon 16:15 T-H28

Rucio Datamanagement: Expanding the capabilities for dealing with potentially problematic files — •Christoph Ames[1], Günter Duckeck[1], Rodney Walker[1], and Cédric Serfon[2] — [1]Ludwig-Maximillians-Universität, Munich — [2]Brookhaven National Laboratory

Rucio is a software framework that enables the management of large collections of data. Originally developed for the ATLAS experiment at the Large Hadron Collider (LHC), it has been expanded to encompass other scientific experiments, such as CMS and Belle II, whereby data is stored on and transferred between over 100 data centres across the world. Accessing files on these data centres doesn't always function without problems occurring, therefore dedicated services are used to track files that are being accessed or transferred. Files that cause problems are marked as suspicious by the tracking services and have to be handled by another service, the suspicious replica recoverer. The goal of this work is to expand the suspicious replica recoverer by adding more flexibility and nuance when deciding what should be done with a suspicious file. This is achieved by introducing a policy system based on file metadata, which allows each virtual organisation to individually set their own policies.

T 15.2 Mon 16:30 T-H28

Probing a tiered storage for a WLCG Tier-3 center — Michael Böhler, •Dirk Sammel, and Markus Schumacher — Albert-Ludwigs-Universität Freiburg

The storage of large amounts of data is an important topic in the HEP community, especially with the start of the High-Luminosity LHC in the near future. This data has to be available to local users at the institutes, which represent the Tier-3 level of the WLCG infrastucture. An interesting scenario would be a "tiered" Tier-3 system consisting of a combination of cold and warm storage. In such a setup, a parallel file system might serve as the warm storage and in addition, an object storage would be used as cold storage to avoid a too large occupancy of the parallel file system. In this talk, performance tests of a local S3 object store using a typical HEP use case are presented. The results are compared to read access to the local parallel storage and to other ATLAS grid sites within the WLCG.

T 15.3 Mon 16:45 T-H28

Modelling Large Scale Distributed Computing Systems for Identification of Efficient Architectures — •Maximilian Horzela, Manuel Giffels, Artur Gottmann, Günter Quast, and Achim Streit — Karlsruhe Institute of Technology

One approach to overcome the expected gap between available and required resources at future colliders like the High Luminosity Large Hadron Collider (HL-LHC) is to increase the efficiency of existing workflows, for example via caching of frequently used data. However, in complex environments with distributed computing systems and many users like the Worldwide LHC Computing Grid (WLCG), finding a solution is not trivial. Due to the complexity and size of such systems, it is not feasible to deploy several experimental test-beds at large scales.

Simulation of such systems has proved a scalable and versatile approach to identify efficient and practical computing architectures for High Energy Physics (HEP). A prominent example is the Monarc Simulation Framework, which lead to the present structure of the WLCG. However, the discontinuation of the old software and the demand for the ability to simulate recent scenarios including caching solutions requires a revision of suitable software.

In the context of this talk, a modern example for such applications based on the WRENCH and SimGrid simulation packages will be presented. Furthermore, first results obtained by the simulation will be shown.

T 15.4 Mon 17:00 T-H28

Dynamic and transparent provisioning of opportunistic compute resources for HEP — •Ralf Florian von Cube, René Caspart, Max Fischer, Manuel Giffels, Eileen Kühn, Günter Quast, and Matthias Schnepf — Karlsruhe Institute of Technology

The utilization of only temporarily available compute resources (opportunistic resources) not dedicated to HEP becomes more and more important for future HEP experiments. On the one hand, due to the unprecedented need and resulting scarcity of HEP compute resources. On the other hand, due to the desired integration of HPC, Cloud and locally available resources.

To meet the challenges posed by the resulting heterogeneous compute environment, the Karlsruhe Institute of Technology (KIT) developed the COBalD/TARDIS resource manager and an entire ecosystem around an HTCondor overlay batch system to allow for a dynamic, transparent and hassle-free integration of those resources into the World-wide LHC Computing Grid (WLCG) via a single point of entry.

In this contribution we will present the current status of COBalD/TARDIS developments as well as our experience with the integration of various opportunistic resources into the WLCG.

T 15.5 Mon 17:15 T-H28

AUDITOR: An accounting system for opportunistically used resources — Michael Böhler, •Stefan Kroboth, and Markus Schumacher — Albert-Ludwigs-Universität Freiburg

Computing clusters often experience varying workload over time, which may lead to suboptimal utilization of available hardware. Sharing resources between multiple clusters can mitigate this inefficiency. The software COBalD/TARDIS enables the integration of resources in an opportunistic, dynamic, and transparent manner and is successfully operated on various sites for a wide range of scenarios. In such a setup, the question of how to account for computations conducted on shared resources arises.

In this work the prototype of AUDITOR, the *AccoUnting Data handlIng Toolbox for Opportunistic Resources* is presented. At its core it consists of a database and a server which can be interacted with using a REST API. This allows *collectors* to store data relevant for accounting in the database and *plugins* to publish the data on an external platform or perform other actions. The extensible nature of AUDITOR enables handling of various use cases by combining the appropriate *collectors* and *plugins*. An example use case is described, where the fairshare of one batchsystem is transfered to priorities of another based on data collected from COBalD/TARDIS.

T 15.6 Mon 17:30 T-H28

Optimization of performance for HEP ML applications on GPU Clusters — •Tim Voigtländer, René Caspart, Manuel Giffels, Günter Quast, Matthias Schnepf, and Roger Wolf — Karlsruhe Institute of Technology, Karlsruhe, Germany

GPU clusters are gaining increased importance also in particle physics. To use GPUs most efficiently, concepts like multi-processing on a single GPU, multi-GPU usage for suitable applications or the balance between CPU and GPU resources must be considered. In particular, GPU support for applications in Machine Learning has become quite common, and they provide a wide variety of usage scenarios. The GPU performance in relation to CPUs depends on the complexity of the network topology, on the training strategy and other hyperparameters of the problem at hand. To illustrate the possible performance gains, a number of scenarios in neural network training on a shared GPU cluster attached to the TOpAS Tier3 at KIT are discussed.

T 15.7 Mon 17:45 T-H28

Belle II Grid Computing Developments at KIT — •Moritz Bauer, R. Florian von Cube, Torben Ferber, Manuel Giffels, Maximilian Horzela, Günter Quast, and Matthias Schnepf — Karlsruhe Institute of Technology

The Belle II experiment studies B-meson decays with high precision. Therefore, Belle II plans to record $50 ab^{-1}$. This will result in 50PB of raw data, which has to be reconstructed and analyzed. Furthermore, corresponding simulations have to be produced. To achieve this goal the Belle II collaboration uses several data centers around the world as a Grid, similar to the worldwide LHC Computing Grid.

While LHC experiments usually only analyze data from one data-taking period, Belle II analyses typically use the complete dataset over all periods. The high energy physics computing group at KIT works on several development projects to handle these challenges. In this presentation, we describe the challenges as well as the current development projects. These projects include the transparent integration of opportunistic resources to provide more computing resources and caching to provide additional copies of datasets to improve their accessibility automatically.