T 59: Grid-Computing und Software

Zeit: Mittwoch 16:00–18:15 Raum: S11

T 59.1 Mi 16:00 S11

Analysis software for the Belle II experiment. — •Sam Cunliffe — DESY

The Belle II experiment at SuperKEKB will benefit from an approximate factor 40 increase in the instantaneous luminosity with respect to Belle / KEKB. This, together with a longer data-taking plan will result in a factor 50 increase in the recorded data.

In order to cope with the challenges of such large data samples, a new software framework has been developed for Belle II including candidate-based high-level analysis tools. This talk will describe the key analysis tools, their integration into the framework, and the analysis model in general.

T 59.2 Mi 16:15 S11

Tuning ROOT parameters to optimize ATLAS Analysis Object Data — •Martin Errenst[1,2] and Attila Krasznahorkay[1] — [1]Bergische Universität Wuppertal — [2]Cern

In ATLAS, reconstructed physics objects are stored in the centrally managed xAOD format. xAOD is based on ROOTs file format and therefore offers a range of configuration settings. In a study where typical file access patterns in ATLAS analyses and other workflows are emulated, these settings are reviewed to optimize for small file sizes and good reading speeds.

The results of this study will be used to find optimal storage configurations for datasets used in different stages of ATLAS analyses. Finding good middle grounds between file size and file reading speed for different requirements of the final stage analyses and that of central data processing, is becoming even more important as ATLAS is preparing for LHC's Run3.

As a cross check to the analysis emulation, similar computing performance studies are done in an ongoing SUSY analysis, which is searching for R-parity violating decays in final states with one lepton and many (>8) jets of which 0 or >3 are b-tagged.

T 59.3 Mi 16:30 S11

Speeding Up Reconstruction of Low Energy Neutrino Events in IceCube with GPUs — •Maicon Hieronymus, Sebastian Böser, Bertil Schmidt, and Elisa Lohfink for the IceCube-Collaboration — Johannes Gutenberg-Universität, Mainz, Deutschland

The IceCube Neutrino Observatory can reconstruct neutrinos with just a few GeV of energy, even though only few hits per event are detected. The multi-dimensional parameter space where each event has an energy deposit at the vertex, potentially an energy and direction of the emerging muon track as well as the usual position and time parameters makes this a non-trivial problem. In addition, the light diffusion process can not be calculated ab-initio but has to be taken from simulation stored in so called photosplines as look-up tables. A maximum likelihood approach is used for the reconstruction. The vertex and the direction are optimized using MultiNest, a multi modal nested sampling algorithm that handles degenerated likelihoods. Track length and energy are determined separately. The likelihood is constructed from hit probabilities stored in spline tables. The overall process takes on average 10 mins per event on a CPU. The lookups in the spline tables are the bottleneck of the reconstruction. In this talk I present approaches using a CUDA-enabled GPU to significantly speed up the reconstruction. The focus lies on balanced workload, coalesced memory access and overlapping computation and communication, which might result in a speed-up of 15-20 over the CPU implementation for $O(10^4)$ likelihood evaluations.

T 59.4 Mi 16:45 S11

Grid Computing at Belle II — •Tobias Jenegger and Thomas Kuhr — Ludwig-Maximilians-Universität München

Over the last decades not just the setup of physics experiments and the measurement instruments have become considerably more sophisticated and complex but also the amount of retrieved data has exceeded what humans can interpret directly. The data can only be handled using large computing systems.

Belle II, a particle physics experiment in Japan designed to study the properties of B mesons, involves over 100 institutes in 25 countries and employs a large grid computing system that allows about 800 scientists to analyse the obtained data and to produce data sets of simulated events. Instead of working with one supercomputer a distributed grid computing system equipped with the most current software technology allows to exploit the computing resources of the individual research institutes in the most efficient way.

This talk will give an overview of the grid computing model at Belle II and present work in progress on developments of an improved user interface.

T 59.5 Mi 17:00 S11

Concept of federating German CMS Tier 3 Resources — •R. Florian von Cube, Manuel Giffels, Christoph Heidecker, Günter Quast, Martin Sauter, and Matthias J. Schnepf — Karlsruher Institut für Technologie

Computing resources located at the different institutes of the German CMS community (DCMS) provide a large infrastructure for end user analyses. Some groups also have access to HPC centers and cloud resources, not dedicated to HEP. Though, all of these resources are mostly only shared within the local group. The community, however, profits from making resources such as HPC centers and cloud resources available to all DCMS users for shared, efficient usage.

In this contribution we present a concept, how an overlay submission infrastructure based on HTCondor's routing and flocking mechanism can be used to connect different resource pools available to the DCMS community in order to allow a transparent and more efficient utilization of all resources. Beside integrating static resources, also dynamically allocated ones can be considered as well, provided that site specific access and usage policies are taken into account and implemented. Virtualization and container technologies such as virtual machines, Docker, and Singularity are key components to assure well-defined software environments on opportunistic resources, no matter where computing jobs are run.

We present the concept, and the current status of the proof-of-concept setup.

T 59.6 Mi 17:15 S11

Integrating Dynafed into the ATLAS workflow — •Benjamin Rottler[1], Frank Berghaus[2], Felix Bührer[1], and Markus Schumacher[1] — [1]Albert-Ludwigs-Universität Freiburg — [2]University of Victoria

In the current ATLAS grid computing model analysis jobs are sent to the grid site where the corresponding data is available. With the ever increasing amout of measured and simulated data this is not feasible anymore in the future. Furthermore, there are plans that smaller grid sides can only provide computing ressources, but not storage.

Dynafed allows to aggregate different storage endpoints where data access is possible via the WebDAV protocol. It supports traditional grid storage solutions like dCache and DPM as well as object stores such as Amazon S3 or Microsoft Azure. Dynafed allows to redirect the user request to the nearest storage endpoint where the data is available.

Our goal is to incorporate Dynafed into the ATLAS grid workflow. For this we need to evalute the performance of Dynafed and WebDAV based access. We present results of those benchmarks and compare Dynafed with traditional access methods. Additionally, we compare the performance of grid storage and object store endoints.

T 59.7 Mi 17:30 S11

HEP Analyses in a Cloud — •Matthias J. Schnepf, R. Florian von Cube, Manuel Giffels, Christoph Heidecker, Günter Quast, and Martin Sauter — Karlsruhe Institute für Technologie

Many user analyses in High Energy Physics (HEP) have a huge demand for computing resources. This demand is unpredictable due to peak loads, which makes it challenging to use the resources efficiently.

Additional computing resources that are not dedicated to HEP usage, so-called opportunistic resources, help to cover the resource demand. We allocate these resources from HPC centers or cloud providers on demand and integrate them transparently into our institute batch system. This results in a flexible cloud-like system with a huge number of heterogeneous resources and a single point of entry.

To improve the utilization of these resources, we developed COBalD (Opportunistic Balancing Daemon), a framework that makes decisions using feedback loops. This allows for reacting dynamically on the cur-

rent demand of resources while adjusting the amount of resources based on their utilization. For the integration of resources from different resource providers, we combined CoBalD with a multi-agent resource manager called TARDIS (Transparent Adaptive Resource Dynamic Integration System) in order to interact with various providers. TARDIS enables reacting on the current demand in a higher granularity which results in a higher resource utilization.

In this contribution, we give an overview of our developments as well as the experiences with our system.

T 59.8 Mi 17:45 S11

Testing XCache for ATLAS workflows — •Guenter Duckeck, Nikolai Hartmann, and Thomas Maier — Ludwig-Maximilians-Universität München

Caching servers are a promising and lightweight alternative or complement to permanent storage clusters. XCache is a service developed by XRootd team which acts as a proxy cache for remote storage. It is integrated into the ATLAS data management system Rucio and provides transparent access to the globally distributed ATLAS data. We have setup such an XCache server at the LRZ-LMU Tier-2, we tested several ATLAS workflows and will discuss possible use cases.

T 59.9 Mi 18:00 S11

Boosting data-intensive HEP analyses by coordinating distributed caches — •Christoph Heidecker, Martin Sauter, Matthias J. Schnepf, Max Fischer, Manuel Giffels, Eileen Kühn, R. Florian von Cube, and Günter Quast — Karlsruhe Institute of Technology

The ever-growing amounts of data processed by HEP user analyses results in challenges for the network and storage infrastructure, which can be tackled by introducing local caches for recurrently accessed data.

Efficient utilization of conventional caches placed within a distributed infrastructure requires both, coordination of data placement and sending work-flows to the most suitable host in terms of data locality. The coordinated and distributed caching approach thereby reduces redundantly stored data and improves the overall processing efficiency.

Thus, the KIT developed the NaviX coordination service, which connects an XRootD caching proxy infrastructure with an HTCondor batch system. The performance improvements of our concept are currently evaluated on opportunistic compute resources as well as on the Throughput-Optimized Analysis-System (TOPAS) cluster dedicated for data-intensive HEP user analyses, which is currently commissioned at KIT.

In this contribution, we give an overview of the coordinated and distributed caching concept, performance benchmark results and experiences gained.