T 35: GRID-Computing

Zeit: Montag 16:45–19:00 Raum: P15

T 35.1 Mo 16:45 P15

The Smart Grid Monitoring System - ANFIS — Gen Kawamura, •Erekle Magradze, Haykuhi Musheghyan, Jordi Nadal, and Arnulf Quadt — II. Physikalisches Institut, Georg-August Universität Göttingen, Friedrich-Hund-Platz 1 37077 Göttingen Deutschland

The demand on the various digital services is increasing and therefore the stable performance of the computing infrastructure is playing a critical role in the robustness of the provided services. Stability and high availability of the high performance computing infrastructure is achievable by an efficient administration, in time failure detection, the proper root cause analysis of the failures and finally by the proper action taking. All these tasks are mainly performed by systems administrators, which are widely using different monitoring and automated action taking systems. These systems were developed and continue evolving to help in the process of administration. The disadvantage of the monitoring or automated action taking tools is an absence of analytic properties, such as autonomous root cause analysis of the service degradation or short term predictions of a failure of the service or group of services. This fact served as a motivation to develop an Artificial Intelligence methods for development of the smart monitoring system. Latest results of the smart monitoring system are presented.

T 35.2 Mo 17:00 P15

HappyFace-progress and future development for the ATLAS experiment — Gen Kawamura, Erekle Magradze, •Haykuhi Musheghyan, Jordi Nadal, and Arnulf Quadt — II. Physikalisches Institut Universität Göttingen , Friedrich-Hund-Platz 1, D-37077 Göttingen

In grid computing importance of monitoring is growing due to the increasing complexity of the computing systems. One of the main tasks for administrators is the status check of hardware and services. The HappyFace project provides an efficient meta-monitoring framework, which allows grid computing infrastructure to aggregate, process and store monitoring data from remote grid resources to a common database and to display latest status information in a one place. Nowadays, we continue the development of modules and web-service based architecture for HappyFace. The purpose is to extend the functionality and the availability, in order to make it beneficial and open for the ATLAS computing environment. Further development is ongoing to achieve a flexible framework able to monitor computing infrastructures of different scales.

T 35.3 Mo 17:15 P15

Echtzeit Produktionsvalidierung mit Hilfe von JEM — •Frank Volkmer und Peter Mättig — Bergische Universität Wuppertal

Fehler in der Monte Carlo Massenproduktion werden oft übersehen oder zu spät erkannt da ein automatisches System zur Qualitätsauswertung fehlt. Dies führt zur Verschwendung von Resourcen, besonders wenn die Fehler erst spät erkannt werden und die generierten Daten bereits im produktiven Analyseeinsatz sind. Probleme entstehen wenn neue Softwareversionen eingesetzt werden deren Ergebnisse nicht ausreichend validiert sind oder wenn Massenproduktion mit falschen Konfigurationdateien gestartet wird.

Der Job Execution Monitor (JEM) ist eine, an der Bergischen Universität Wuppertal entwickelte, Software zur Überwachung von Grid-Jobs. Als nachladbares Modul des Pilot wird JEM genutzt um die Massenproduktion zu validieren. Die zu validierenden Tasks werden dem JEM Activation Service übergeben, welcher dann automatisch Jobs mit JEM instrumentiert bis die geforderte zu validierende Eventzahl erreicht ist. Den instrumentierten Jobs wird das Kommando übergeben, zusätzlich Qualitätshistogramme zum normalen Output zu erzeugen und diese an einen Server zu übertragen. Dort werdem sie mit vorhandenen Histogrammen anderer Jobs zusammengefasst und dann mit den entsprechenden Referenzhistogrammen ausgewertet. Die Ergebnisse werden dann automatisch auf einer Webseite zur Verfügung gestellt und via Mailingliste kommuniziert.

T 35.4 Mo 17:30 P15

Tier2 Report goes HammerCloud — •Michael Böhler, Adil Aktas, Anton Gamel, and Jan Erik Sundermann — Albert-Ludwigs-Universität, Freiburg, Deutschland

The sites of the Worldwide LHC Computing Grid for the ATLAS, CMS, and LHCb experiments are monitored to evaluate performance and reliability using a large number of different workload jobs that are experiment specific. These jobs are steered, controlled and monitored by the HammerCloud testing infrastructure.

A framework has been developed for the GridKa ATLAS computing cloud which extracts and visualizes the most relevant information in order to identify possible minor issues and problems at computing sites, which do not cause large scale failures. While such issues are usually not conspicuous enough to warrant the blacklisting of a specific site, they still can cause a few percent job inefficiencies.

The original design of the framework was to provide overview tables and plots as a status report for the weekly GridKa cloud meetings. The plots have been integrated into the HammerCloud monitoring webpage, which enables the continuous monitoring of site performance such that any problems can be dealt with immediately.

This talk describes the different development steps of this infrastructure, from its first conception to a wider usage in the entire GridKa computing cloud, and finally to global usage worldwide.

T 35.5 Mo 17:45 P15

Dynamic data allocation and long term storage access in the German ATLAS cloud — •Thomas Hartmann[1], Gen Kawamura[2], Kai Leffhalm[3], Friederike Nowak[3], Günter Duckeck[4], Torsten Harenberg[5], Marisa Sandhoff[5], Rodney Walker[4], Michael Böhler[6], Anton Gamel[6], Andrzej Olszewski[7], and Jan Erik Sundermann[6] — [1]KIT Karlsruhe — [2]Universität Göttingen — [3]DESY Hamburg — [4]LMU München — [5]Bergische Universität Wuppertal — [6]Universität Freiburg — [7]Instytut Fizyki Jadrowej Krakow

The global data processing and simulation of HEP data is a comparable effort as the actual data taking by the LHC experiments as ATLAS. For a successful analysis of the recorded events, several different Monte Carlo simulations have to be performed. Like the event data, the computing resources are distributed globally. Organizing and making these resources available is the task of the Worldwide LHC Computing Grid. Within the Grid, Gridka is the local organization of computing and storage resources at computing tiers in Austria, the Czech Republic, Germany, Poland, and Switzerland. Within Gridka, Tier-2 centres provide computing resources for user analyses and MC production and, as one of the larger Tier-1 centres, event data and computing resources are provided at KIT Karlsruhe. We present the cloud's implementation and experience of ATLAS' approach for dynamic data allocation ("federated access") in addition to the classic hierarchic approach. Furthermore, the talk will focus on the cloud's development of resources for long term storage access for end users.

T 35.6 Mo 18:00 P15

Datenmanagement für hochperformante Analysen — •Max Fischer[1,2], Christopher Jung[1] und Günter Quast[2] — [1]Karlsruher Institut für Technologie, Steinbuch Centre for Computing — [2]Karlsruher Institut für Technologie, Institut für Experimentelle Kernphysik

Computing in der Hochenergiephysik steht gleichermaßen wachsenden Anforderungen und Einschränkungen gegenüber. Zukünftige Herausforderungen sind nur durch effiziente Nutzung und Spezialisierung der Infrastrukturen zu bewältigen.

Im Rahmen der Arbeiten für High-Performance Data Analysis (HPDA) werden Strategien und Lösungen für die Entwicklung eines HEP-Analysezentrums entwickelt. Hardwarearchitektur, dynamische Jobslots sowie lokales und globales Datenscheduling bilden ein umfassend Konzept. Durch einen modularen Aufbau können auch bestehende Ressourcen hiervon profitieren.

Für die effektive Planung ist ein genaues Verständnis verfügbarer Komponenten in Analysesituation essentiell. Durch standardisierte, analyserelevante Tests wird die Performance von Systemen untersucht. Besonderer Schwerpunkt liegt dabei auf für Institute und Arbeitsgruppen relevante Hardware.

Der Vortrag vergleicht die Leistung aktueller Systeme für Nutzeranalysen. Erste Testergebnisse von Komponenten für Desktop- und Serverhardware werden präsentiert. Abschließend wird darauf aufbau-

end das HPDA Konzept und seine Komponenten vorgestellt.

T 35.7 Mo 18:15 P15

Tests of the Federated ATLAS XrootD system with HammerCloud — •Federica Legger, Guenter Duckeck, Johannes Elmsheuser, and Friedrich Hoenig — Ludwig-Maximilians-Universitaet Muenchen

The HammerCloud framework was developed to ensure the smooth operation of the ATLAS Grid infrastructure through the automatic validation of the site capability to execute both analysis and production jobs, and to test specific issues such as the validation of new sites, performance evaluation after hardware changes, and testing new software releases and configurations. In particular, HammerCloud has been extensively used to test the Federated ATLAS XrootD system (FAX), a storage federation aiming to treat Tier-1, Tier-2 and Tier-3 storage space as a single distributed storage system. We report on the results of such tests, aiming to evaluate the efficiency, use cases, and possible limitations of the FAX infrastructure.

T 35.8 Mo 18:30 P15

Statistical analysis of Grid activities in a WLCG Tier-2 site — •Gen Kawamura, Jordi Nadal, and Arnulf Quadt — II. Physikalisches Institut, Georg-August Universität Göttingen, Friedrich-Hund-Platz 1, 37077, Göttingen, Deutschland

The World-Wide LHC Computing Grid (WLCG) sites have handled massive data and job requests produced by the LHC and the ATLAS experiments. The dCache storage management system and the CREAMCE job management system have been widely used by the HEP community in the LHC experiments, and their massive grid activities have been recorded. However, such massive activities are being increasingly utilized to understand the actual use of the grid and timely monitoring. Thus we must use significant computational resources for analyzing such activities. We consider a data reduction technique of such data under the conditions that sites are operational and are in production for the LHC experiments. Thus, the log data containing statistical variables are not assumed to be Gaussian. Therefore in this case, it is necessary to consider bootstrap approximations based on statistics, since the distribution of data is unknown. A random data selection of the bootstrap method can be applied, thus reducing the cost of computations, and instantly obtaining up-to-date information. Using results from a large-scale distributed WLCG Tier-2 site, GoeGrid, we introduce a new procedure which presents a computationally efficient method of assessing quality among grid users, clients, grid components and individual networks.

T 35.9 Mo 18:45 P15

Centralized PanDA Blacklisting System — •Arwa Bannoura — Bergische Universität Wuppertal

PanDA is the distributed production and distributed analysis system for ATLAS. A concurrent update on the status of PanDA queues causes inconsistencies. This information can be updated by shifters and different agents like Hammercloud (a distributed analysis testing system). A centralized PanDA blacklisting system was implemented within AGIS (ATLAS GRID Information System) to manage the information about the status and update PanDA with the final status. The decision of the final status is based on the priority of the agent/shifter and how restricted is the proposed status. The system includes a RESTful API as well as a WebUI to manage and view the blacklisting status information of queues and clouds.