Third EELA Conference

Abstracts book

http://indico.eu-eela.org/conferenceDisplay.py?confId=96

Table of contents

Abstract ID : 0

Experience running an ATLAS distributed Tier2 and Tier3 at IFIC-Valencia

The ATLAS computing model describes a hierarchical distributed virtual computing facility within which are defined Tier-1 and Tier-2 computing centers having certainspecific MOU agreed roles and capacities to be used for the benefit and at the direction ofATLAS as a whole.In this model the primary functions of the Tier-1 are to host and provide long term storage for, access toand re-reconstruction of a subset of the ATLAS RAW data (20% in the case of the Tier-1), provide access to ESD, AOD and TAG data sets and support the analysis of thesedata sets. The primary functions of the Tier-2.s are simulation (they provide the bulk of simulation for ATLAS), calibration, chaotic analysis for a subset of analysis groups and hosting of AOD, TAG and some physics group samples.

Tier-3 sites are institution-level non-ATLAS funded or controlled centers/clusters which wish to participate in ATLAS computing, presumably mostfrequently in support of the particular interests of local physicists (physicists at the local Tier-3 decide how these resources are used). These are clusters of computers which can vary widely in size. It should be noted that substantial institutional funding to originate such clusters is potentially available, and that they could make a real contribution to the impact of ATLAS on the overall ATLAS physics output. As such, there is considerable value in providing some level of technical support to these sites.

In this talk the experience gained on running, maintaining, supporting and managing a Tier2 centre will be presented. Finally, a Tier-3 prototype at IFIC-Valencia is going to be discussed, in order to meet ATLAS data-taking requirements.

Primary authors : Dr. GONZALEZ DE LA HOZ, Santiago (IFIC - Instituto de Fisica Corpuscular Valencia)

Co-authors : Dr. SALT CAIROLS, Jose (IFIC - Instituto de Fisica Corpuscular de Valencia)

Presenter : Dr. GONZALEZ DE LA HOZ, Santiago (IFIC - Instituto de Fisica Corpuscular Valencia)

Track classification :

Contribution type : --not specified--

Submitted by : Dr. GONZALEZ DE LA HOZ, Santiago

Submitted on Monday 09 July 2007

Last modified on : Monday 09 July 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 4

Toward a Grid Workflow Formal Composition

To enhance the security and reduce the complexity of grid workflow applications such as e-science or e-business applications, large amount of computational resources are required and collaborative model development is needed since multiple parties could be involved in the development process and in the use of workflows. The use of a standard modelling language seems to be an appropriate solution for the description of different workflow parts and the communication between them.
In this paper, we survey existing approaches in workflow modelling and we focus on these using UML Activity Diagram (UML-AD) to model and develop workflows. In particular, this paper exposes an abstract syntax for the UML-AD as well as a formal foundation for grid workflow composition in form of a workflow algebra based on UML-AD. This composition fulfils the need of collaborative model development and the proposed algebra allows the definition of workflow views. These views are useful to limit the access to predefined users only in order to ensure the security of grid workflow applications. Finally, the need for the grid workflow normalization concept has arisen during the composition of workflow models.

Primary authors : Mrs. HLAOUI BENDALY, Yousra (Research Unit of Technologies of Information and Communication)

Co-authors : Dr. JEMNI BENAYED, Leila (Faculty of sciences of tunis)

Presenter : Mrs. HLAOUI BENDALY, Yousra (Research Unit of Technologies of Information and Communication)

Track classification :

Contribution type : --not specified--

Submitted by : Mrs. HLAOUI BENDALY, Yousra

Submitted on Friday 10 August 2007

Last modified on : Friday 10 August 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 5

A Worldwide Production Grid Service Built on EGEE and OSG Infrastructures - Lessons Learnt and Long-term Requirements

Using the Grid Infrastructures provided by EGEE, OSG and others, a worldwide production service has been built that provides the computing and storage needs for the 4 main physics collaborations at CERN's Large Hadron Collider (LHC). This service must cater for the needs of thousands of physicists in hundreds of institutes in tens of countries. A 24x7 service with availability of up to 99% is required with major service responsibilities at each of some ten "Tier1" and of the order of one hundred "Tier2" sites. Such a service - which has been operating for some 2 years and will be required for at least an additional decade - has required major manpower and resource investments from all concerned and is considered a major achievement in the field of Grid computing. We describe the main lessons learned in offering a production service across heterogeneous Grids as well as the requirements for long-term operation and sustainability.

Primary authors : Dr. SHIERS, Jamie (CERN)

Co-authors :

Presenter : Dr. SHIERS, Jamie (CERN)

Track classification :

Contribution type : --not specified--

Submitted by : Dr. SHIERS, Jamie

Submitted on Thursday 16 August 2007

Last modified on : Thursday 16 August 2007

Comments :

This talk will address several of the issues targetted for day 3 of the conference and is also related to High Energy Phys

Status : SUBMITTED

Track judgments :

Abstract ID : 6

Robust & Resilient Services - How to design, build and operate them

Grid infrastructures require a high degree of fault tolerance and reliability. This can only be achieved by careful planning and detailed implementation. We describe on-going work within the WLCG project to build and run highly reliable services. This talk will describe the "a priori" analysis based on the services and service levels listed in the Memorandum of Understanding that sites participating in WLCG have signed, together with an "a posteriori" analysis following over 2 years of production service. This work covers not only the services deployed at the Tier0 centre at CERN - which has the most stringent service requirements related to the acquisition of the raw data, the initial processing phase and the distribution of raw and processed data to Tier1 sites, but also a similar analysis for Tier1 and major Tier2 sites. The latter will be covered at a workshop that will take place shortly before the EELA conference and so will be very up-to-date.

Primary authors : Dr. SHIERS, Jamie (CERN) ; Dr. MCCANCE, Gavin (CERN)

Co-authors :

Presenter : Dr. MCCANCE, Gavin (CERN)

Track classification :

Contribution type : --not specified--

Submitted by : Dr. SHIERS, Jamie

Submitted on Thursday 16 August 2007

Last modified on : Monday 24 September 2007

Comments :

The techniques involved include hardware deployment, operational procedures as well as middleware requirements.

Status : SUBMITTED

Track judgments :

Abstract ID : 7

Towards A Scalable Scientific Data Grid Model and Discovery Service

Scientific Data Grid mostly deals with large computational problems. It provides geographically distributed resources for large-scale data-intensive applications that generate large scientific data sets. This required the scientist in modern scientific computing communities involve in managing massive amounts of a very large data collections that geographically distributed. Research in the area of grid has given various ideas and solutions to address these requirements. However, nowadays the number of participants (scientists and institutes) that involve in this kind of environment is increasing tremendously. This situation has leads to a problem of scalability. In this paper, we present a Peer-to-Peer (P2P) model for Scientific Data Grid that utilizes the P2P services to address the scalability problem. By using this model, we study and propose various decentralized discovery strategies that intend to address the problem of scalability. We then investigate the impact of data replication that addressing the data distribution and reliability problem for our Scientific Data Grid model on the propose discovery strategies. In this study, we used our own data grid simulation written using PARSEC. In this paper, we illustrate our P2P Scientific Data Grid model and our data grid simulation. We then analyze the performance of the discovery strategies with and without the existence of replication strategies relative to their success rates, bandwidth consumption and average number of hop.

Primary authors : Mr. ABDULLAH, Azizol (Universiti Putra Malaysia)

Co-authors : Dr. OTHMAN, Mohamed (Universiti Putra Malaysia) ; Dr. SULAIMAN, Md Nasir (Universiti Putra Malaysia) ; Dr. IBRAHIM, Hamidah (Universiti Putra Malaysia) ; Prof. OTHMAN, Abu Talib (Universiti Kuala Lumpur)

Presenter : Mr. ABDULLAH, Azizol (Universiti Putra Malaysia)

Track classification :

Contribution type : --not specified--

Submitted by : Mr. ABDULLAH, Azizol

Submitted on Wednesday 05 September 2007

Last modified on : Wednesday 05 September 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 8

LiveWN: CPU scavenging in the Grid Era

LiveWN was prototyped in 2006 as a grid-bootable CD supporting diskless, easy-to-deploy worker nodes and requiring virtually zero administration upon deployment, through automatic subscription. The system mixed key technologies including LiveCD (a self-booting Linux CD), gLite (the grid middleware stack), OpenVPN (virtual private network) and OpenAFS (open distributed filesystem), proving that the initial concept was correct and workable. In the meantime the same concept has been expanded in the larger DVD format gLiteDVD, in order to include a graphical environment, the ATLAS software, Povray, WINE, VLC and a list of other tools which we still keep expanding.

Hereby, we wish to present the currently existing technology, the advantages and disadvantages, as well as a series of issues that still have to be addressed in greater depth and offer opportunity for cutting-edge research for the near future:
* Internet Protocol Version 6,
* configuration management (based on a Content Delivery Network, Service Location Protocol, etc),
* incorporation of non-standard grid users (exploiting GridShib and OpenID over MyProxy),
* accounting and credit control (using DGAS and similar systems),
* benchmarking and optimal resource allocation (using LMbench),
* checkpointing and overall reliability improvement,
* automatic load-balancing, and fail-over in case of faults.
In fact, we have already made progress in about half of these areas and hope to expand.

Primary authors : Mr. GEORGATOS, Fotis (National Technical University of Cyprus) ; Mr. KOURETIS, Giannis (National Technical University of Cyprus)

Co-authors :

Presenter : Mr. GEORGATOS, Fotis (National Technical University of Cyprus) ; Mr. KOURETIS, Giannis (National Technical University of Cyprus)

Track classification :

Contribution type : --not specified--

Submitted by : Mr. GEORGATOS, Fotis

Submitted on Monday 10 September 2007

Last modified on : Monday 10 September 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 9

Metagenomic Analysis on the EELA Grid

In many cases, the sequencing of the DNA of many microorganisms is hindered by the impossibility of growing significant samples of isolated specimens. Many bacteria cannot survive alone, and require the interaction with other organisms. In such cases, the information of the DNA available belongs to different kinds of organisms. Metagenomic studies aim at processing samples of multiple specimens to extract the genes and proteins that belong to the different species. This can be achieved through a process of extraction of fragment, comparison and analysis of the function. By the comparison to existing chains, whose function is well known, fragments can be classified.

This process is computationally expensive and requires several iterations of alignment and phylogenic classification steps. Source samples reach several millions of sequences, which could reach up to thousands of nucleotids each. These sequences are compared to a selected part of the "Non-redundant" database which only implies the information from eukaryotic species. From this first analysis, a refining process is performed and alignment analysis is restarted from the results. This process implies several CPU years.

An environment has been developed to fragment, automate and check the above operations. This environment has been tuned-up from an experimental study which has tested the most efficient and reliable resources, the optimal job size, and the data transference and database reindexation overhead. The environment should re-submit faulty jobs, detect endless tasks and ensure that the results are correctly retrieved and workflow synchronised. The paper will give an outline on the structure of the system, the lessons learnt and the results obtained.

Primary authors : Dr. BLANQUER, Ignacio (Universidad Politécnica de Valencia) ; Prof. HERNáNDEZ, Vicente (Universidad Politécnica de Valencia) ; Mr. APARICIO, Gabriel (Universidad Politécnica de Valencia)

Co-authors :

Presenter : Dr. BLANQUER, Ignacio (Universidad Politécnica de Valencia)

Track classification :

Contribution type : --not specified--

Submitted by : Mr. BLANQUER, Ignacio

Submitted on Monday 10 September 2007

Last modified on : Monday 10 September 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 10

VOMS Server replication process in I2G and EELA Projects

Virtual Organization Membership Service server (VOMS) is the Core Service for managing authorization data in the E-Infrastructure Shared Between Europe and Latin America (EELA Grid Project) and in other gLite based grid infrastructures. The VOMS server provides a database of users, groups, roles and capabilities that are grouped in Virtual Organizations (VO's). When needed, users query the VOMS Server in order to get their VO grid credentials for access to the grid resources. Other VOMS-enabled services (Computer Element, Storage Element and others) use VOMS plug-ins to validation user access requests.

Due to the fact that the VOMS Server is a critical core service it is very important to be available all the time. In other words this service requires scalability, fault tolerance and high availability. This can be achieved with a simple service replication process, keeping database synchronization and providing a distributed and transparent access for users and grid services.

The VOMS authorization service performance can be affected when the number of requests received both from users and other services is high. However, it is possible to deploy VOMS server replicas in other sites to provide some load balance.

Furthermore, the grid infrastructures that are deployed in large geographic areas, such as EELA, can take advantage of this distributed access to the authorization infrastructure by using the nearest VOMS replica. Moreover, with VOMS service replication a fault tolerant system can be implemented in order to transparently distribute the requests to other replicas when the main server is unavailable.

This paper presents a description of the VOMS replication process in the EELA and int.eu.grid projects including a detailed description of all problems and the respective solutions implemented at LIP in Portugal. It also shows how to extend this approach to other core service such as the LCG File Catalog (LFC). Finally we show some VOMS replication problems that don't have an implemented solution yet.

Primary authors : Mr. SILVA, Bruno (LIP) ; Mr. GOMES, Jorge (LIP)

Co-authors :

Presenter : Mr. SILVA, Bruno (LIP)

Track classification :

Contribution type : --not specified--

Submitted by : Mr. SILVA, Bruno

Submitted on Tuesday 11 September 2007

Last modified on : Tuesday 11 September 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 13

Porting of bioinformatic tools for plant virology on a computational grid

The goal of TriGrid Project is the creation of the first Sicilian regional computational grid. In particular, it aims to build various software-hardware interfaces between the infrastructure and some scientific and industrial applications. In this context, we have integrated most innovative computing applications in virology research inside GRID infrastructure. Particularly, we have implemented, all in a complete workflow, tools for pairwise or multiple sequence alignment and phylogeny tree construction (ClustalW-MPI), phylogenetic networks (SplitsTree), detection of recombination by phylogenetic methods (TOPALi) and prediction of RNA secondary structures (KnetFold) . This work will show how created porting applications are able to decrease the execution time of the analysis programs; to improve the accessibility to the data storage and to manage a lot of metadata for data processing.

Primary authors : Dr. LANZALONE, Gaetano (INFN Catania)

Co-authors : Dr. LOMBARDO, Alessandro (PI2S2-Catania) ; Dr. MUOIO, Annamaria (INFN Catania) ; Dr. IACONO MANNO, Marcello (INFN Catania)

Presenter : Dr. LANZALONE, Gaetano (INFN Catania)

Track classification :

Contribution type : --not specified--

Submitted by : Dr. MUOIO, Annamaria

Submitted on Wednesday 19 September 2007

Last modified on : Wednesday 19 September 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 14

Optimizing the configuration of magnetic confinement devices with evolutionary algorithms and grid computing

We present a proposal for enhancing the configuration of a stellarator device - the TJ-II Flexible Heliac - in order to improve the performance of this magnetic confinement device. To achieve this goal, we propose to use grid computing with genetic and evolutionary algorithms. Grid computing allows performing many experiments in a parallel way. Genetic algorithms will allow us to obtain good solutions without exploring the whole solution space, because the number of parameters involved in the configuration of these devices and the number of combinations among these values make impossible to explore all the possibilities.

Primary authors : Mr. GóMEZ-IGLESIAS, Antonio (CETA-CIEMAT)

Co-authors : Dr. VEGA-RODRíGUEZ, Miguel A. (UNEX) ; Dr. CASTEJóN-MAGAñA, Francisco (CIEMAT) ; Mr. RUBIO DEL SOLAR, Manuel (CETA-CIEMAT) ; Mr. CáRDENAS MONTES, Miguel (CETA-CIEMAT)

Presenter : Mr. GóMEZ-IGLESIAS, Antonio (CETA-CIEMAT)

Track classification :

Contribution type : --not specified--

Submitted by : Mr. ANTONIO, Gómez-iglesias

Submitted on Saturday 22 September 2007

Last modified on : Thursday 27 September 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 15

Enabling efficient access to ATLAS data for Latin American collaborators.

The ATLAS Collaboration consists of 168 institutions, in 37 countries of 6 continents. Approximately 2000 physicists participate in this truly world-wide enterprise.
In order to enable fair and speedy access to the data to all Collaboration members, a complex distributed computing infrastructure has been designed and is now in the last phases of commissioning.
Main cornerstones of this infrastructure are the Distributed Data Management System (DDM), the Distributed Production System (ProdSys) and the Distributed Analysis System (Ganga/pAthena), supplemented by the Information System and the Monitoring and Accounting tools. ATLAS can make use of three Grid middleware stacks, provided by the projects EGEE (world-wide), NorduGrid (in the Scandinavian countries) and OSG (in the USA).
Several Latin American institutions joined ATLAS during the last couple of years, therefore we had to develop a work model that enables full and efficient access to all ATLAS data for small communities of geographically separated physicists.
This talk will describe the ATLAS Data and Computing Models, with particular attention to the issues of accessing data and running distributed analysis applications on the Grid.

Primary authors : Dr. BARBERIS, Dario (CERN & Genoa University/INFN) ; Dr. POULARD, Gilbert (CERN) ; Dr. DOVA, Maria Teresa (La Plata)

Co-authors :

Presenter : Dr. BARBERIS, Dario (CERN & Genoa University/INFN)

Track classification :

Contribution type : --not specified--

Submitted by : Dr. BARBERIS, Dario

Submitted on Tuesday 25 September 2007

Last modified on : Tuesday 25 September 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 16

Virtual Institute for Integrative Biology (VIIB): an eScience and High Performance Grid Computing Paradigm for Latin America

This presentation examines the case of the Virtual Institute for Integrative Biology (VIIB) as a Latin American paradigm for achieving global collaborative e-science using cyber-infrastructure and high performance grid computing. Biology has emerged as one of the major areas of focus of scientific research worldwide, providing new challenges for the sharing of huge databases, the use of computer intensive calculations and the development of novel bioinformatics algorithms. Whereas major efforts have been mounted in the USA, Europe and Asia to meet these challenges, less appears to have been accomplished in Latin America and the VIIB was developed to fill this need. VIIB is housed in the Center for Bioinformatics and Genome Biology (CBGB) and is part of the Fundación Ciencia para la Vida (FCV) (www.cienciavida.cl) in Santiago, Chile. Strong links have been initiated with other groups in Latin America, European and North America and it is anticipated that additional collaborators will be incorporated into the VIIB shortly. The FCV operates a 128 cpu computer cluster running Windows 2003 and is connected to other Latin American centers and to Europe (EELA) via REUNA and RedCLARA.

The scientific agenda being addressed by the VIIB includes: construction and management of databases for comparative genomics of organisms of particular relevance to Latin America, the development of a database of alternate reading frames for over 600 sequenced genomes, including organisms of medical and biotechnological importance, bioinformatics services and protein simulations. Human resource development through shared teaching, co-sponsored students and seminars is also an integral component of the collaborative effort. E-science challenges include: connectivity concerns, high performance computing (HPC) limitations, development of a customized Grid framework, language issues, maintenance of open access without compromising security and coherent approaches to the dissemination of scientific and technical information. Finally, it was recognized that computational frameworks, flexible workflows and visualization techniques were required to efficiently exploit shared resources without causing impediments to the user who has little interest in the underlying information technology (IT).

Overall, the VIIB is proving an effective way for small teams to transcend the critical mass problem, to overcome geographic limitations and to harness the power of large scale, collaborative science; as such, it may prove a useful model for promoting e-science and high performance grid computing in Latin America and other emerging regions.

Primary authors : Dr. HOLMES, David (Center for Bioinformatics and Genome Biology, Fundación Ciencia para la Vida)

Co-authors :

Presenter : Dr. HOLMES, David (Center for Bioinformatics and Genome Biology, Fundación Ciencia para la Vida)

Track classification :

Contribution type : --not specified--

Submitted by : Dr. HOLMES, David

Submitted on Wednesday 26 September 2007

Last modified on : Wednesday 26 September 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 18

Interoperability studies between the GTRS and EUMEDGRID e-Infrastructures

Grid computing enables the sharing, selection, and aggregation of a wide variety of geographically distributed computational resources such as supercomputers, clusters, storage systems and data sources. A grid presents them as one unified resource for solving large-scale and data intensive computing applications.
A middleware supports applications in distributed computing environments by providing services that enable the interconnectivity and interoperability of applications, systems and machines. Considering the evolution of this type of middleware, it is important to regroup several national and international grids, by creating a gateway between middlewares, to gather more power and resources.
In this setting, our work consists on making the Tunisian national grid GTRS interoperable with the grid infrastructure of EU funded project EUMEDGRID. A new concept of Super Worker Node is proposed in this work to reach the interoperability between the two grids of GTRS and EUMEDGRID.

Primary authors : Mr. ABBES, Heithem (UTIC)

Co-authors : Dr. BARBERA, Roberto (University of Catania and INFN) ; Dr. JEMNI, Mohamed (UTIC)

Presenter : Mr. ABBES, Heithem (UTIC)

Track classification :

Contribution type : --not specified--

Submitted by : Mr. ABBES, Heithem

Submitted on Thursday 27 September 2007

Last modified on : Thursday 27 September 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 19

e-Science perspectives in Venezuela

The work we present illustrates where Venezuela is aiming in terms of e-Science, seen in eyes of the Centro Nacional de Cálculo Científico Universidad de Los Andes (CECALCULA), Mérida, the Universidad de Los Andes (ULA), Mérida, and the Instituto Venezolano de Investigaciones Científicas (IVIC), Caracas. We will present the plans for Grid Venezuela and what has been done so far as well as Grid ULA, supported by Internet2. We will show different web based scientific applications we are working on or have developed for quantum chemistry, atomic physics, structure damage analysis, biomedicine and bioclimat, developed mainly within the framework of the E-Infrastructure shared between Europe and Latin America (EELA).

Primary authors : HOEGER, Herbert (EELA Member - Universidad de Los Andes - Mérida - Venezuela)

Co-authors : DíAZ, Gilberto (ULA) ; NúñEZ, Luis (ULA) ; HAMAR, Vanessa (Cecalcula) ; UZCáTEGUI, Maylett (Cecalcula) ; RUIZ, Nicolás (Cecalcula) ; TORRENS, Rodrigo (Cecalcula) ; FLóREZ-LóPEZ, Julio (ULA) ; MENDOZA, Claudio (IVIC)

Presenter : HOEGER, Herbert (EELA Member - Universidad de Los Andes - Mérida - Venezuela)

Track classification :

Contribution type : --not specified--

Submitted by : HOEGER, Herbert

Submitted on Thursday 27 September 2007

Last modified on : Thursday 27 September 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 20

Distributed Analysis Experience using Ganga on an ATLAS Tier-2 infrastructure

The ATLAS detector will explore the high-energy frontier of Particle Physics collecting the proton-proton collisions delivered by the LHC (Large Hadron Collider). Starting in spring 2008, the LHC will produce more than 10 Petabytes of data per year. The adapted tiered hierarchy for computing model at the LHC is: Tier-0 (CERN), Tiers-1 and Tiers-2 centres distributed around the word.

The ATLAS Distributed Analysis (DA) system has the goal of enabling physicists to perform Grid-based analysis on distributed data using distributed computing resources. IFIC Tier-2 facility is participating in several aspects of DA. In support of the ATLAS DA activities a prototype is being tested, deployed and integrated.

The data processing applications are based on the Athena framework. Ganga, developed by LHCb and ATLAS experiments, allows simple switching between testing on a local batch system and large-scale processing on the Grid, hiding Grid complexities. Ganga deals with providing physicists an integrated environment for job preparation, bookkeeping and archiving, job splitting and merging.

The experience with the deployment, configuration and operation of the DA prototype will be presented. Experiences gained of using DA system and Ganga in the Top physics analysis will be described.

Primary authors : Dr. FASSI, Farida (IFIC) ; Dr. CABRERA, Susana (IFIC) ; Ms. VIVES, Roger (IFIC)

Co-authors : Dr. SALT, Jose (IFIC) ; Dr. GONZALEZ DE LA HOZ, Santiago (IFIC) ; Mr. MARCH, Luis (IFIC) ; Mr. SANCHEZ, Javier (IFIC) ; Dr. MOHAMMED, Kaci (IFIC) ; Mr. FERNANDEZ, Alvaro (IFIC) ; Mr. LAMAS, Alejandro (IFIC)

Presenter : Dr. FASSI, Farida (IFIC)

Track classification :

Contribution type : --not specified--

Submitted by : Dr. FASSI, Farida

Submitted on Friday 28 September 2007

Last modified on : Friday 28 September 2007

Comments :

Grid Communities and Applications: High Energy Physics

Status : SUBMITTED

Track judgments :

Abstract ID : 21

GrEMBOSS: EMBOSS over the EELA GRID

With the growth of genome databases and the complexity for processing such information within bioinformatics research, there is a need for computing power and massive storage facilities which can be provided by Grid infrastructures.

EMBOSS is a free Open Source software analysis package (more than 170 applications) specially developed for the needs of the bioinformatics and molecular biology user community. This work describes the deployment of EMBOSS over a GLite middleware-based infrastructure, such as the EELA GRID.

This work is focused on rewriting the I/O EMBOSS libraries (AJAX) to use the GFAL library from the LCG/EGEE middleware. This library allows the use of files registered on the catalog service which are contained in the storage elements of a GRID.

Submitting a job into a Grid is not an easy task. This work also describes an adhoc mechanism to allow bioinformaticians to concentrate on the EMBOSS command, instead of acquiring advanced knowledge about Grid usage.

The results obtained so far demonstrate the functionality of GrEMBOSS, and represent an efficient and viable alternative for "gridifying" other bioinformatics applications.

Primary authors : Mr. BONAVIDES-MARTINEZ, Cesar (CCG, UNAM, Mexico)

Co-authors : Mr. MURRIETA, Eduardo (DGSCA, UNAM, Mexico) ; Mr. VERLEYEN, Jérôme (IBt, UNAM, Mexico) ; Mr. HERNANDEZ, Alfredo (CCG, UNAM, Mexico) ; Mr. ZAYAS, Romualdo (CCG, UNAM, Mexico) ; Mr. RODRíGUEZ, Roberto (IBt, UNAM, Mexico) ; Mr. BRANGER, Pierre-alain (CCG, UNAM, Mexico) ; Dr. VALVERDE, José R. (CIEMAT, Spain)

Presenter : Mr. BONAVIDES-MARTINEZ, Cesar (CCG, UNAM, Mexico)

Track classification :

Contribution type : --not specified--

Submitted by : Mr. BONAVIDES-MARTINEZ, Cesar

Submitted on Saturday 29 September 2007

Last modified on : Saturday 29 September 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 22

An Optimized Ring-Based Election Algorithm for Trusted Computational Grid Environment

As computational Grid grows and gains more popularity, increase of its security concerns must be considered and resolved. In this way, one important task is to keep trustworthiness between connecting nodes. We are going to design a trust-based Grid environment in which Grid is divided into some parts called region. In each region one node is specified as leader that has a key role in performance of trust management mechanism. In our approach leader does not work as a time coordinator, but it performs tasks such as communicating to other regions and exchanging security information. Regarding importance of leader role in each region, we require an algorithm to dynamically specify a new leader when current one is failed to perform its tasks. In this paper we concentrate on leader nodes and propose an algorithm for leader election in each region. Our approach is based on Ring algorithm and decreases number of required messages efficiently by applying some modifications to original algorithm. We name this algorithm as GT-Ring. Our simulation results show that our approach is scalable and optimum in a Grid environment.

Primary authors : Mr. DELDARI, Hossein (Department of Computer, Ferdowsi University of Mashhad, Iran)

Co-authors : Mr. TAGHIPOOR, Rasool (Department of Computer, Ferdowsi University of Mashhad, Iran)

Presenter : Mr. DELDARI, Hossein (Department of Computer, Ferdowsi University of Mashhad, Iran)

Track classification :

Contribution type : --not specified--

Submitted by : Mr. TAGHIPOOR, Rasool

Submitted on Saturday 29 September 2007

Last modified on : Saturday 29 September 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 23

Gridification of the JPEG2000 standard for the compression of Gigabytes satellite images

Image processing is one of the domains that require high calculation capacity and storage facilities especially when we are talking about huge images' manipulation like ultrasonic and satellite images. In fact, by manipulating these latters we cannot afford to lose any kind of information that might be important. For this purpose, those images are generally used and manipulated with no kind of compression and this is because of the fact that the actual compression algorithms do not offer the possibilities of giving a good lossless compression ratio that can override their slowlyness applied to these exceptionally huge images. But lately, a new compression Standard, namely JPEG2000, was introduced. This standard offers, in addition to other interesting characteristics, the possibility of lossless image compression, under the same core and with extremely competitive compression ratio. This use of the standard fits perfectly with the demanded lossless compression of gigabytes images, but presents, at the same time, some handicaps related to the nature of these images and to the algorithm itself some of these are the following: at first, none of the implemented versions of this standard can handle huge size images like the one we're considering, at second no machine can support the task of compressing these images at once due to the luck of memory and processing performance (eg. the JPEG2000 standard necessitates a virtual memory 10 times equal to the size of the considered image; thing impossible to happen if we're talking about images with 10 Gb and more). The use of parallel machines, in this case, could be a possible solution, but the availabilities of these machines is not evident and this could be one the disadvantages of this solution.
In this paper we will describe the gridification and implementation of the JEPG2000 standard on the EUMEDGRID infrastructure. This gridified version of the Jasper implementation of the JPEG2000 standard follows three steps, based on each worker node of the grid. Each Worker Node extracts (remotely from the Storage Element based on Gfal routines) a portion of the considered image, compresses it and finally saves the result back into the Storage Element. A merge task is performed after the gathering of all image portions. We will demonstrate that, this gridified algorithm, not only, compresses gigabyte images successfully but also efficiently and in a highly competitive processing time.
This paper will be divided into the following parts. The first one will give a brief presentation of the JPEG2000 algorithm and its different limits related to domain we are concidering. In the second part we will describe the additional CoDec that we developed to extend the capabilities of the program to the type of images we are manipulating. The third part will describe the different steps that we have developed for the compression of the considered image. We will finalize with some experimental resultats, that will support the use of Grids for this kind of treatments, and a conclusion.

Primary authors : Dr. KHANFIR, Sami (UTIC)

Co-authors : Prof. JEMNI, Mohamed (UTIC) ; Mrs. VALERIA, Ardizzoni (INFN)

Presenter : Dr. KHANFIR, Sami (UTIC)

Track classification :

Contribution type : --not specified--

Submitted by : Dr. KHANFIR, Sami

Submitted on Monday 01 October 2007

Last modified on : Monday 01 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 24

GENECODIS-Grid: An online gird-based tool to predict functional information in gene lists

Nogales, R., Mejía, E., Vicente, C., Tirado, F., Pascual-Montano, A.

Computer Architecture Department. Complutense University of Madrid, Spain
pascual@fis.ucm.es

GENECODIS [1] is a web-based tool that integrates different sources of information to search for annotations that frequently co-occur in a set of genes and rank them by statistical significance. The analysis of concurrent annotations provides significant information for the biologic interpretation of high-throughput experiments and may outperform the results of standard methods for the functional analysis of gene lists. This software, publicly available at http://genecodis.dacya.ucm.es/, uses a very complex and computing intensive datamining methodology to predict all possible combinations of functional information that are significant enriched in a list of genes. These algorithms make the system very inefficient when large sets of genes or proteins are provided.
Due to the large popularity of this tool, that has registered more than 11000 visits since its publication in January 2007, there is a strong need to facilitate users from different sites to access the system simultaneously. In addition, the complexity of some of the statistical tests used in this approach has made this technique a good candidate for its implementation in a Grid environment.

In this paper we introduce GENECODIS-Grid, a grid-based tool based on gLite 3.0. It takes advantage of two independent networks managed by a meta-scheduler and a web server to host the application. When the user submits a job through the web server, the meta-scheduler determines the less workloaded network to execute the job. This approach represents a cost-effective alternative to improve the throughput of these kinds of methods, highly used in Bioinformatics.

References

[1] Pedro Carmona-Saez, Monica Chagoyen, Francisco Tirado, Jose M Carazo and Alberto Pascual-Montano. GENECODIS: A web-based tool for finding significant concurrent annotations in gene lists. Genome Biology. 2007 Jan 4;8(1):R3

Primary authors : Mr. NOGALES, Ruben (Computer Architecture Department. Complutense University of Madrid, Spain)

Co-authors : Mr. MEJIA, Edgardo (Computer Architecture Department. Complutense University of Madrid, Spain) ; Mr. VICENTE, Cesar (Computer Architecture Department. Complutense University of Madrid, Spain) ; Mr. TIRADO, Francisco (Computer Architecture Department. Complutense University of Madrid, Spain) ; Mr. PASCUAL-MONTANO, Alberto (Computer Architecture Department. Complutense University of Madrid, Spain)

Presenter : Mr. NOGALES, Ruben (Computer Architecture Department. Complutense University of Madrid, Spain)

Track classification :

Contribution type : --not specified--

Submitted by : Mr. PASCUAL-MONTANO, Alberto

Submitted on Thursday 04 October 2007

Last modified on : Thursday 04 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 25

Grid Portal For Image and Video Processing

Users are typically best served by "Grid Portals", which are web servers that allow the user to configure or run a class of applications. The server is then given the task of authenticating the user with the Grid and invoking the required grid services in order to launch the user's application. PHP is a widely-used general-purpose scripting language that is especially suited for Web development and can be embedded into HTML. PHP is powerful and modern server-side scripting language producing HTML or XML output which easily can be accessed by everyone via web interface (with the browser of your choice). PHP can execute shell scripts on the server side. The aim of our work is development of Grid portal for image and video processing. The shell scripts contains gLite and globus commands for obtaining proxy certificate, job submission, data management etc. Using this technique we can easily create web interface to the Grid infrastructure. The image and video processing algorithms are implemented in C++ language using various image processing libraries.

Primary authors : Mr. DIMITROVSKI, Ivica (ID)

Co-authors : Prof. BUCKOVSKA, Aneta (AB) ; Prof. LOSKOVSKA, Suzana (SL)

Presenter : Mr. DIMITROVSKI, Ivica (ID)

Track classification :

Contribution type : --not specified--

Submitted by : Mr. DIMITROVSKI, Ivica

Submitted on Tuesday 09 October 2007

Last modified on : Tuesday 09 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 26

Numerical climate simulations on the GRID. Experiences in the EU-project EELA

Recent trends in climate modeling find in GRID computing a powerful way to achieve results by sharing computing and data distributed resources. In particular, ensemble prediction is based on the generation of multiple simulations from perturbed model conditions to sample the existing uncertainties. In this work, we present a GRID application consisting of a sequence of two state-of-the-art climate models (one global model and one regional model), operable through a web portal (based on Genius). The main goal of the application is providing ensemble-based regional predictions. This requires managing a complex workflow involving long-term jobs and job dependencies in a user-transparent way. On one hand, a single climate simulation may last longer than 2 CPU-days, thus requiring the restart and reallocation of the jobs. On the other hand, the regional model uses as boundary conditions the output of the global model to achieve higher resolution (from ~200 down to ~10km) and, therefore, DAGs should be managed to wait for the global model. In doing so, we identified the weaknesses of current middleware tools and developed a robust workflow by merging the optimal existing applications with an underlying self-developed workflow application based on the communication with metadata catalogs (currently AMGA) storing application status and dynamic model output generation.

As an illustrative scientific challenge, the application is applied to study the El Niño phenomenon, by simulating an El Niño year with different forcing conditions and analyzing the precipitation response over south-american countries subject to flooding risk.

Primary authors : Ms. FERNANDEZ-QUIRUELAS, Valvanuz (University of Cantabria) ; Dr. FERNANDEZ, Jesus (University of Cantabria) ; Mr. BAEZA, Claudio (University of Concepcion) ; Dr. COFIñO, Antonio S. (University of Cantabria) ; Dr. MIGUEL, Richard (SENAMHI) ; Dr. ABARCA, Rodrigo (University of Concepcion) ; Dr. GUTIERREZ, Jose Manuel (University of Cantabria)

Co-authors :

Presenter : Ms. FERNANDEZ-QUIRUELAS, Valvanuz (University of Cantabria)

Track classification :

Contribution type : --not specified--

Submitted by : Dr. FERNANDEZ, Jesus

Submitted on Wednesday 10 October 2007

Last modified on : Wednesday 10 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 27

Grid-oriented PSO for the Energy Generation Expansion Problem

This work presents a grid-oriented mapping of a Particle Swarm Optimization based solver for the SATyrus platform. SATyrus is a Pseudo-Boolean Lyapunov Function synthesizer – it assembles an unified objective function from a target problem specification, including viability and optimality constraints written on a propositional style specification language. Particle Swarm Optimization (PSO) is a stochastic optimization method with some traits of evolutionary computing. Originally formulated for continuous-space problems, the approach has been modified to handle binary-dimensional problems. The grid mapping is illustrated by experiments on the energy Generation Expansion Planning (GEP) problem, and its solution optimizes the costs of power distribution through construction/placement and utilization of different kinds of power plants.

Primary authors : Mr. DIACOVO, Ramon (COPPE - UFRJ) ; Prof. FRANçA, Felipe (COPPE - UFRJ) ; Prof. LIMA, Priscila (CDERJ)

Co-authors :

Presenter : Mr. DIACOVO, Ramon (COPPE - UFRJ)

Track classification :

Contribution type : --not specified--

Submitted by : Mr. DIACOVO, Ramon

Submitted on Wednesday 10 October 2007

Last modified on : Wednesday 10 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 28

Tuning of transfers over fat pipe network links between Europe and Latin America

The EELA Grid infrastructure is deployed over the Geant2 and RedCLARA continental backbone networks. Those
networks are interconnected trough the ALICE transatlantic link, which exhibits characteristics of high bandwidth
over long round trip delay transmission circuits, also called fat pipes. Normally TCP transfers are known to present
subpar performance on fat pipes when default configuration parameters are used. In order to tune TCP transfers
between both sides of the Atlantic, a Service Challenge was devised to study the best configuration parameters. This
work focuses on the topology of the network as relevant to the Service Challenge, and a review of the TCP
congestion control and window size mechanisms, and how to tune those properly under linux. After proper tuning
the infrastructure is able to achieve sustained throughputs of over 500Mbps on a link of 622Mbps.

The EELA project is co-funded by the European Commission under the contract number IST-2006-026409.

Primary authors : CARVALHO, Diego (CEFET-RJ) ; RAUSCH BELLO, Pedro Henrique (UFRJ) ; NOBREGA DUARTE, Alexandre (CERN/UFCG) ; PEREZ-GRIFFO, Javier (CIEMAT)

Co-authors : SALINAS-ESCRIBANO, Roque (CIEMAT) ; MORALES-RAMOS, Enrique (CIEMAT) ; OLIVEIRA, Frederico (UFRJ) ; PORTO, Eriko (CLARA)

Presenter : CARVALHO, Diego (CEFET-RJ) ; RAUSCH BELLO, Pedro Henrique (UFRJ) ; NOBREGA DUARTE, Alexandre (CERN/UFCG) ; PEREZ-GRIFFO, Javier (CIEMAT)

Track classification :

Contribution type : --not specified--

Submitted by : CARVALHO, Diego

Submitted on Wednesday 10 October 2007

Last modified on : Friday 12 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 29

Operating a Transatlantic Grid Infrastructure

Grids have the potential to revolutionize computing by providing ubiquitous, on demand access to computational services and resources. However, grid systems are extremely large, complex and prone to failures, what requires a constant monitoring of its services and prompt reaction to detected problems. Our experience in the operation of the EELA infrastructure, composed by sites in several countries in Latin America and in Europe, showed that there is a non-negligible effort involved in keeping the grid services running with the level of quality expected by its users. Moreover, the operation of a service involving institutions spread around several different countries with different organizational procedures and working on different time-zones can make this an ever harder problem.

In this paper we will present the operation procedures adopted in the EELA project and the tools used to keep the infrastructure running within the expected levels of availability. We start by describing the operation workflows created to cope with the reality of operating a wide spread service like a transatlantic grid. We present two different workflows, one describing the process to integrate a new site in the infrastructure and another one describing the process used to keep the sites already present in the infrastructure up and running.

Then we present the monitoring, diagnose and support tools used in the daily operation of the grid. These tools include fabric and application-layer monitoring and a ticketing system used to track problems and solutions and to provide a knowledge database about grid problems.
Finally, we discuss some of the difficulties found during the operation of the EELA infrastructure, how these difficulties where surpassed and the lessons learned during the operation of a widely distributed grid infrastructure.

The EELA project is co-funded by the European Commission under the contract number IST-2006-026409.

Primary authors : CARVALHO, Diego (CEFET-RJ) ; RAUSCH BELLO, Pedro Henrique (UFRJ) ; NOBREGA DUARTE, Alexandre (CERN/UFCG)

Co-authors :

Presenter : CARVALHO, Diego (CEFET-RJ) ; RAUSCH BELLO, Pedro Henrique (UFRJ)

Track classification :

Contribution type : --not specified--

Submitted by : CARVALHO, Diego

Submitted on Wednesday 10 October 2007

Last modified on : Friday 12 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 30

Building a Grid in Latin America: The EELA Project e-Infrastructure update

Several international projects and collaborations have emerged in the last years due to the increasing demand for
grid distributed resources around the world. One important aspect of these initiatives deals with the ``gridification''
of computing intensive scientific applications otherwise difficult to run efficiently. The EELA Project (E-Infrastructure
shared between Europe and Latin America) is a collaboration between Latin American and European institutions
that has developed a potent e-Infrastructure for e-Science applications. This paper describes the state of the art at
the end of the project of the consolidated EELA e-Infrastructure and the progress achieved so far.

The EELA project is funded by the European Commission under the contract number IST-2006-026409.

Primary authors : CARVALHO, Diego (CEFET-RJ) ; RAUSCH BELLO, Pedro Henrique (UFRJ) ; NOBREGA DUARTE, Alexandre (CERN/UFCG) ; GOMES, Jorge (LIP) ; STANTON, Michael (RNP/CLARA)

Co-authors :

Presenter : CARVALHO, Diego (CEFET-RJ)

Track classification :

Contribution type : --not specified--

Submitted by : CARVALHO, Diego

Submitted on Wednesday 10 October 2007

Last modified on : Friday 12 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 31

Progress of Grid technology in Argentina: Lessons learned from EELA

The EELA project aimed to create a collaboration network between Europe and Latin American for training in Grid technologies and the deployment of a pilot Grid infrastructure for e-science applications. Grid computing has emerged as an important new field, and its development in Argentina is particularly important for a number of reasons, such as that Argentina has recently joined the ATLAS collaboration at CERN and the increasing interest in future biomedical applications. The potential of the GRID technology is well known, however, its adoption is not a trivial task as it requires significant investment in several areas. In this paper, the achievements and progress in Argentina through close collaboration with EELA are presented. Among these are the deployment of a Grid Certification Authority infrastructure that is a crucial component in the activities of the e-Science community in the country; the deployment, integration and validation of a small local EELA node; installation and running of an analysis ATLAS application on the EELA infrastructure. The experience gained in participating in EELA dissemination events allowed us also to actively promoting GRID and training different target audiences in Argentina and in LA.

Primary authors : Prof. DOVA, Maria Teresa (Physics Department, Universidad Nacional de La Plata)

Co-authors : GRUNFELD, Christian (Physics Department, Universidad Nacional de La Plata) ; MONTICELLI, Fernando (Physics Department, Universidad Nacional de La Plata) ; VEIGA, Alejandro (Physics Department, Universidad Nacional de La Plata) ; AMBROSI, Viviana (CESPI, LINTI and Fac. Informatica, Universidad Nacional de La Plata) ; BARBIERI, Andres (CESPI, LINTI and Fac. Informatica, Universidad Nacional de La Plata) ; DIAZ, Javier (UCESPI, LINTI and Fac. Informatica, niversidad Nacional de La Plata) ; LUENGO, Miguel (CESPI, LINTI and Fac. Informatica, Universidad Nacional de La Plata) ; MACIA, Nicolas (CESPI, LINTI and Fac. Informatica, Universidad Nacional de La Plata) ; MOLINARI, Lia (CESPI, LINTI and Fac. Informatica, Universidad Nacional de La Plata) ; VENOSA, Paula (CESPI, LINTI and Fac. Informatica, Universidad Nacional de La Plata) ; ZABALJAUREGUI, Matias (CESPI, LINTI and Fac. Informatica, Universidad Nacional de La Plata)

Presenter : Prof. DOVA, Maria Teresa (Physics Department, Universidad Nacional de La Plata)

Track classification :

Contribution type : --not specified--

Submitted by : Prof. DOVA, Tere

Submitted on Thursday 11 October 2007

Last modified on : Thursday 11 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 32

Experience of developing a grid infrastructure in Cuba

In the frame of EELA Project and a national initiative for developing a Cuban Network for Science a grid infrastructure was deployed at Centro de Gestión de la Información y Desarrollo de la Energía (CUBAENERGIA). A stand-alone model was adopted to overcome connectivity limitations. The e-infrastructure is based on gLite-3.0/gilda middleware and is fully compatible with EELA-infrastructure.

The work was focused on grid applications. The biomedical application GATE was deployed from the early beginning for biomedical users. Further, two applications were deployed on the local grid infrastructure: MOODLE for e-learning and AERMOD for assessment of local dispersion of atmospheric pollutants. Additionally, our local grid infrastructure was made interoperable with a Java based Distributed System for Bioinformatics Calculations.

This experience could be considered as a suitable approach for national networks with weak Internet connections.

Primary authors : Mr. LóPEZ ALDAMA, Daniel (CUBAENERGIA) ; Mr. DOMíNGUEZ, Maikel (CUBAENERGIA)

Co-authors :

Presenter : Mr. LóPEZ ALDAMA, Daniel (CUBAENERGIA)

Track classification :

Contribution type : --not specified--

Submitted by : Mr. NEUMANN CIUFFO, Leandro

Submitted on Thursday 11 October 2007

Last modified on : Thursday 11 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 33

GACS: a Generic Framework for Grid Accounting and Charging

Grid Accounting and Charging System (GACS) is a generic accounting and charging framework for the Grid. GACS is based on a concept envisioned by the researchers of the Charging and Accounting Technology for the Internet (CATI) project. The main achievement of this concept is that it defines the recommended modules of such architecture: metering, accounting, pricing, charging and billing. The authors of this article have applied this concept in the design of the GACS framework.
This framework is set up to support various Grid systems by defining simple and unified interfaces. The resulting architecture is highly generic, modularized, scalable and transaction-safe. Its interfaces are defined in XSD schema. GACS is fully GSI enabled.
In this article, authors present the architecture, features and functionality of this framework. The real operation of GACS is demonstrated in a sample application, which migrates APEL accounting records collected on EGEE computing elements deployed in Hungary.

Primary authors : Mr. SOMOGYI, Csongor (BME Department of Control Engineering and Information Technology)

Co-authors : Dr. SZEBERéNYI, Imre (BME Department of Control Engineering and Information Technology)

Presenter : Mr. SOMOGYI, Csongor (BME Department of Control Engineering and Information Technology)

Track classification :

Contribution type : --not specified--

Submitted by : Mr. CSONGOR, Somogyi

Submitted on Thursday 11 October 2007

Last modified on : Thursday 11 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 34

Events simulation production for the BaBar experiment using the grid approach

The BaBar experiment is taking data since 1999, investigating the violation of charge and parity (CP) symmetry in the field of High Energy Physics.

Event simulation is an intensive computing task, due to the complexity of algorithm based on Monte-Carlo method implemented using the GEANT engine.

Data needed as input for the simulation, stored in the ROOT format, are classified into two categories: conditions data for describing the detector status when data are recorded, and background triggers data for including noise signal necessary to obtain a realistic simulation.

In order to satisfy these requirements, in the traditional BaBar computing model events are distributed over several sites involved in the collaboration where each site manager centrally manages a private farm dedicated to simulation production..

The new grid approach applied to the BaBar production framework in discussed along with the schema adopted for data deployment via Xrootd servers, including data management using grid middleware on distributed storage facilities spread over the INFNGrid network.

A comparison between the two models is provided, describing also the custom application developed for performing the whole production task on the grid and showing results achieved.

Primary authors : Dr. ANDREOTTI, Daniele (INFN) ; Dr. FELLA, Armando (INFN) ; Prof. LUPPI, Eleonora (Ferrara University)

Co-authors :

Presenter : Dr. FELLA, Armando (INFN)

Track classification :

Contribution type : --not specified--

Submitted by : Dr. ANDREOTTI, Daniele

Submitted on Thursday 11 October 2007

Last modified on : Thursday 11 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 35

Grid monitoring in EUChinaGrid infrastructure

Grid Monitoring in EUChinaGrid infrastructure

Lanxin Ma
European Organization for Nuclear Research, Geneva - Switzerland
Institute of high energy physics, Beijing-China
Lanxin.Ma@cern.ch

Abstract

EUChinaGrid is a Project which is founded by EU in 6th framework. The project aims to support the interconnection and interoperability of Grids between Europe and China. Until now, EUChinaGrid infrastructure contains 11 sites which cover 4 counties in China and in Europe. According to the GStat, there are more than 1000 CPUs, 66TB of storage space in EUChinaGrid infrastructure to support many applications, such as HEP(CMS, ATLAS), astro physics(ARGO-YBJ), biology etc. So, it is important to provide reliable grid services, improve the reliability of the grid infrastructure and provide stakeholders with views of the infrastructure allowing to understand the current and historical status of the service. For this purpose, SAM, GridIce and GStat etc. are used to monitoring EUChinaGrid infrastructure. In this paper, we present tools which are able to check if a given grid service works as expected for a given user or set of users on the different resources available on a grid.
This paper also aims to evaluate the current existing grid monitoring status in EUChinaGrid infrastructure and to give the thinking what could be improved in the future to increase the quality and quantity of monitoring information.

Primary authors : Mrs. MA, Lanxin (CERN/IHEP)

Co-authors :

Presenter : Mrs. MA, Lanxin (CERN/IHEP)

Track classification :

Contribution type : --not specified--

Submitted by : Mrs. MA, Lanxin

Submitted on Thursday 11 October 2007

Last modified on : Thursday 11 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 36

Grid for Mesoamerican Archaeology

Mesoamerican archeology requires working with a great amounts of information that is both disperse and
diverse. To generate knowledge from this information in a more efficient way, thus creating a better
understanding of history, it is important to include new working methods that optimize the acquisition, conservation, retrieval and analysis of the information; it is important to note that this information is
continuously expanding and it encompasses a large geographical area including Mexico, Guatemala, Belize,
Honduras and El Salvador; this large amount of data includes text, coordinates, raster graphs and vector graphs. All this information represents elements like tepalcates, constructions (archeological sites), mural paintings, high and low relieves, topography, maps and information of the fauna and the soil.

The Grid offers a solution in handling this information; it respects the necessity of independence for each generating group and at the same time it offers a platform to process and compare all the information obtained. Additionally, concerning the space-time analysis, remote visualization techniques become available through the Grid, drawing on different solutions that can include geographical information systems and virtual reality.

Primary authors : LUCET, Genevieve (UNAM)

Co-authors :

Presenter : LUCET, Genevieve (UNAM)

Track classification :

Contribution type : --not specified--

Submitted by : LUCET, Genevieve

Submitted on Thursday 11 October 2007

Last modified on : Thursday 11 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 37

How to access databases from EGEE-II grid environment: a comparison of tools and middlewares.

Grid infrastructures currently in use for production purposes are strongly computing-oriented, and they have proven to be capable of providing storage and computing resources suitable for scientific communities whose applications require intensive computation and data storage.

New e-Science projects have a wider perception of the Grid, as their applications require not only traditional computations but also the use of complex data operations that require on-line access to databases.
In this paper we examine the comparative evaluation of four tools to access data resources exposed onto Grids: G-DSE, GRelC, OGSA-DAI and AMGA. We compare the ability of the tools in accessing different types of data resources, their peculiarities and their limits.
The evaluation tests address the needs of two scientific communities that needs the adoption of a grid database access layer at the base of their research activities: the Bioinformatics one (BioinfoGRID and the LIBI projects) and of the
Astrophysical community (Istituto Nazionale di Astrofisica).

The test plan spans from very simple queries, use cases provided by the Bioinformatics and Astrophysical communities, up to stress tests with multiple queries submitted simultaneously from different sites.
Some tests are based also on very complex query coming from Italian Public Administration
requirements in order to test the flexibility of the tools.

The need to integrate databases and databases technology in the Grid environment has been recognized as a core research activity by the Grid community. This work represents an important baseline for the different scientific communities that want to
adopt a grid middleware to access databases and data sources and also for the middleware developers that want to integrate a database access system in addition to their Grid facilities.

Primary authors : Dr. TAFFONI, Giuliano (INAF - Oats)

Co-authors : Dr. FIORE, Sandro (University of Salento, Italy) ; DONVITO, Giacinto (INFN Bari, Italy) ; JAIN, Atul (INFN Bari, Italy) ; Prof. MAGGI, Giorgio (INFN Bari, Italy) ; Prof. ALOISIO, Giovanni (University of Salento, Italy) ; Prof. VUERLI, Claudio (INAF - Oats)

Presenter : Dr. TAFFONI, Giuliano (INAF - Oats)

Track classification :

Contribution type : --not specified--

Submitted by : Dr. GIULIANO, Taffoni

Submitted on Friday 12 October 2007

Last modified on : Friday 12 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 38

Solving Inductive Logic Programming problems in the EELA infrastructure

Inductive Logic Programming (ILP) is a technique intended to discover hidden data from relational databases. Given a background knowledge, a set of positive examples, a set of negative examples, and a language bias, the objective is to generate first order rules that (almost) perfectly describe all positive examples and none of the negative examples. We have been working with several domains, when applying ILP: drug discovery, analysis of mammograms, link discovery, among others. These domains present very large databases and sets of examples. In this work we present results on extracted models obtained from these data when running on the EELA infrastructure with some figures on new discovered knowledge and utilization of resources that made it possible to achieve these results.

Primary authors : Dr. DUTRA, Inês (Universidade do Porto)

Co-authors :

Presenter : Dr. DUTRA, Inês (Universidade do Porto)

Track classification :

Contribution type : --not specified--

Submitted by : Dr. DUTRA, Inês

Submitted on Friday 12 October 2007

Last modified on : Friday 12 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 39

Experience with Large Simulations on the EGEE Grid for the AUGER collaboration

The Pierre Auger Cosmic Ray Observatory is studying ultra-high energy cosmic rays, the most energetic and rarest of particles in the Universe. These highly energetic particles initialize extensive air showers while crossing the Earth atmosphere. CPU intensive Monte Carlo simulations
are needed to compare predictions of different models with observed data. Resources from EGEE grid were also used to run these simulations.

We tested grid sites supporting the AUGER VO in a CPU challenge in April and May 2007. The first production using CORSIKA with EPOS model was finished in September 2007 followed by a second production with different input parameters. Experience from these productions will be presented.

Primary authors : Ms. SCHOVANCOVA, Jaroslava (CESNET, Prague) ; Dr. CHUDOBA, Jiri (FzU and CESNET, Prague) ; Dr. TRAVNICEK, Petr (FzU, Prague)

Co-authors :

Presenter : Ms. SCHOVANCOVA, Jaroslava (CESNET, Prague)

Track classification :

Contribution type : --not specified--

Submitted by : Ms. SCHOVANCOVA, Jaroslava

Submitted on Friday 12 October 2007

Last modified on : Friday 12 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 40

Grid Technology and Quality Assurance

Grid is one of the potential architectures of the coming years to support both the research and the commercial environment. Quality assurance techniques needs both to adapt to these new architectures and exploit them to improve its effectiveness.

Software quality is a key issue in the Digital Era and Industries as well as Public administrations devote time to check and verify the quality of ICT products and services supplied. The definition and measurement of quality metrics is a key point for implementing any QA method.

As an example, according to "Metrics Best Practices" [1] only 25% of companies are at least at CMM level 2 and not more than 70 organization reach level 5 in the world (50 of them are in India).

These results show that certification for software quality is a very important issue of ICT companies worldwide. At the same time these figures show the difficulty for business organizations to be compliant to high-quality standards.

Currently several models and standards (e.g. ISO9126, CMMi, ITIL) have been proposed and are recognised to be effective in guiding organisations (from managers to developers) to produce quality software and to better managing the related activities. Such models certify that the company uses effective development processes, which may lead to good quality software.

But why adopting these guidelines in your organisation seems to be difficult?

Typical QA models are based on standard processes and parallel activities that need to be adequate to the volume of the business. Small and Medium Enterprises rarely start their business having a certification (say ISO or CMM) unless the market strongly requires it (i.e. to work in defense markets, such as with NATO, require the AQAP160). In general the overall effect of several factors (e.g. competences on QA, time to be spent on activities perceived as no-core, and a general scepticism in the actual possibility of getting results) cause many companies to avoid any type of certification.

In ETICS [2] we have first started from developing a powerful system designed to support research projects in automating build and test processes. On the basis of the availability of the system and its ability to fully automate the build and test procedures required for software production, we have developed the idea to asses the quality of a software application through the collection of quality metrics and trends as part of the execution of the build and test procedures procedures. In this way gathering quality measures on code, software structure, build run and test executions is done automatically and gives evidence of the status of the software and the trends during the development.

In this paper we propose a quality certification model, named Grid Quality Certification Model (GQCM), that uses automatically calculable metrics to asses the quality of software applications; this model can be integrated in any automatic build and test tool to have a fully independent suite that can be used to test software programs and to certify their quality level.
As an example, an implementation of GQCM using the ETICS system will be shown.
[1] Kulik P. and Weber C., "Metrics Best practices", 2001, KLCI research group http://www.klci.com

[2] http://www.eu-etics.org/

Primary authors : Dr. RIPPA, Adriano (Engineering Ingegneria Informatica S.p.A.) ; Dr. BEGIN, Marc-elian (CERN) ; Dr. DI MEGLIO, Alberto (CERN) ; Dr. MANIERI, Andrea (Engineering Ingegneria Informatica S.p.A.)

Co-authors :

Presenter : Dr. RIPPA, Adriano (Engineering Ingegneria Informatica S.p.A.)

Track classification :

Contribution type : --not specified--

Submitted by : Dr. RIPPA, Adriano

Submitted on Friday 12 October 2007

Last modified on : Friday 12 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 41

Data transfer for the EuchinGrid project

Into the EuChinaGrid Project a Data Transfer tool has been developed to move data from the data aquisition site and the main grid sites in China and Italy.This application was developed using FTS service.
In this paper we want to summarize the performances of this service and also some tests performed using other services with different customizations.
This is one of the most interesting items involved in the real world to perform the best delivery of data to different Storage Elements distributed into the grid.

Primary authors : Dr. CELIO, Paola (INFN Roma TRE - Dipartimeto di Fisica Roma TRE) ; Dr. CELLINI, Simone (INFN Roma TRE) ; Dr. BITELLI, Federico (INFN Roma TRE - Dipartimeto di Fisica Roma TRE) ; Dr. GARGANA, Riccardo (INFN Roma TRE) ; Mr. SITENI, Maurizio (INFN Roma TRE)

Co-authors :

Presenter : Dr. CELIO, Paola (INFN Roma TRE - Dipartimeto di Fisica Roma TRE)

Track classification :

Contribution type : --not specified--

Submitted by : Dr. CELIO, Paola

Submitted on Friday 12 October 2007

Last modified on : Friday 12 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 42

EELA infrastructure a governance case

Grid computing has emerged as an important area, distinguished from conventional distributed computing by
its focus on large-scale resource sharing, innovative applications, and high-performance orientation.
This technology unveils the relationship between institutions which provide computing and human resources in order to build the Grid Infrastructure. The Computer Network can be extended as an Institutional
Network, a set of independent organisations whose iterations are not exclusively regulated by formal
legal contracts. The main goal of this set is co-ordinated by agreements that regulates sharing of responsibilities and risks. This arrangement defines responsibilities and roles which configure the set configuration,
partner roles, resource allocation, decision power, etc. This work presents the EELA Project Organisation,
through the present governance status, against standard models. It is discussed governance solutions for
academic grids and possible extensions for other sectors.

The EELA project is co-funded by the European Commission under the contract number IST-2006-026409.

Primary authors : CARVALHO, Diego (CEFET-RJ) ; BARBASTEFANO, Rafael (CEFET-RJ) ; FLEXA, Raquel (CEFET-RJ) ; FRANçA, Felipe (COPPE/UFRJ)

Co-authors :

Presenter : CARVALHO, Diego (CEFET-RJ)

Track classification :

Contribution type : --not specified--

Submitted by : CARVALHO, Diego

Submitted on Friday 12 October 2007

Last modified on : Friday 12 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 43

gLibrary/DRI: A grid-based platform to host multiple repositories for digital content.

In this work we present the gLibrary/DRI (Digital Repositories Infrastructure) platform. We start from gLibrary, a system with a easy-to-use web front-end designed to save and organize multimedia assets on Grid-based storage resources. We add to the gLibrary features the DRI functionality: a platform to host any kind of repository, providing a common infrastructure and a set of mechanisms (APIs) that the repository providers use to define the data model, the access to content (by navigation trees and filters) and the storage model. DRI offers a generic
way to provide all this functionality; nevertheless the providers can add specific behaviors to the default functionality for their repositories.
The architecture is totally Grid based (VO system, data federation and distribution, computing power, etc). Some features of gLibrary/DRI are Grid authentication/authorization, multi-repository support, hierarchical browsing, item view, etc.
A working example based on Mammographies repository is also presented.

Primary authors : Mr. RAúL, Ramos Pollán (CETA-CIEMAT) ; Mr. MANUEL, Rubio Del Solar (CETA-CIEMAT) ; Mr. ANTONIO, Calanducci (INFN Catania)

Co-authors : Mr. DORIN, Tcaci (Maat-Knowledge) ; Mr. JUAN MANUEL, González Martín (Maat-Knowledge)

Presenter : Mr. RAúL, Ramos Pollán (CETA-CIEMAT)

Track classification :

Contribution type : --not specified--

Submitted by : Mr. MANUEL, Rubio

Submitted on Friday 12 October 2007

Last modified on : Friday 12 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 44

GENIUS/EnginFrame Grid Portal: VOMS Proxy creation, new features and enhancements.

Scientific domain knowledge and tools must be presented to the (non-expert) users in terms of applications without needing to know the underlying details of the Grid Middlewares. GENIUS Grid portal, powered by EnginFrame, is an increasingly popular mechanism for creating customizable, Web-based interfaces to Grid services and resources. This work describes new GENIUS portal's capabilities such as portal login, access control, display management, new approach to building reusable portal components as plugins, better performance results as consequence of EnginFrame core functions improvement. Finally will be described two new tools developed for VOMS Proxy creation and simple JDL composer.

Primary authors : FALZONE, Alberto (NICE srl) ; ARDIZZONE, Valeria (INFN Catania) ; BARBERA, Roberto (INFN Catania)

Co-authors : SCARDACI, Diego (INFN Catania) ; GIORGIO, Emidio (INFN Catania) ; NERI, Lorenzo (INFN Catania) ; VENUTI, Nicola (NICE srl)

Presenter : FALZONE, Alberto (NICE srl) ; ARDIZZONE, Valeria (INFN Catania) ; BARBERA, Roberto (INFN Catania)

Track classification :

Contribution type : --not specified--

Submitted by : FALZONE, Alberto

Submitted on Friday 12 October 2007

Last modified on : Friday 12 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 45

EELA Grid infrastructure MPI support, a success case

MPI support in grid environments is far from standardised. Each resource provider who supports MPI uses a diferent implementation from the other, making the task of using MPI on grids very difficult. This work focuses on our impressions on what MPI support should be like, an example taken from the EELA infrastructure, a survey on recent developments done by several grid projects, and the solution adopted by the EELA infrastructure team. At the end, two use cases are described and their results analysed.

The EELA project is co-funded by the European Commission under the contract number IST-2006-026409.

Primary authors : CARVALHO, Diego (CEFET-RJ) ; RAUSH BELLO, Pedro Henrique (UFRJ) ; NOBREGA DUARTE, Alexandre (CERN/UFCG) ; FRANçA, Felipe (COPPE/UFRJ)

Co-authors :

Presenter : CARVALHO, Diego (CEFET-RJ)

Track classification :

Contribution type : --not specified--

Submitted by : CARVALHO, Diego

Submitted on Friday 12 October 2007

Last modified on : Friday 12 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 47

Network support for e-Science in Latin America

Computer networks in Latin America have connected scientists in the region to their peers in other parts of the world since 1986. Starting with the creation of Internet2 in 1996, a new global research network has been extended throughout the world, providing communications infrastructure for large-scale international scientific collaboration. With the creation of the RedCLARA network and its links to Europe and the US between 2004 and 2005, this global network reached the majority of Latin America countries, setting the stage for much closer collaboration between scientists in Latin America and their counterparts in other countries.

In this article we describe the development of the research networking infrastructure currently available within the region together with its inter-regional connections, and how this infrastructure is being used for support of e-science. Particular attention is given to the role of the national research and education networks (NRENs) in the region, and of their association, CLARA, in providing networking support for e-science projects. CLARA and Latin American NRENs are active partners in the EU-supported EELA and RINGrid projects, and also are making significant supporting contributions to the success of other international projects with Latin American partners, in fields such as High-Energy Physics, Astronomy and Astrophysics and Space Geodesy, to single out the early adopters of advanced networking technologies. These contributions are described in the article.

The article concludes describing future trends in networking infrastructure in the region, in order to meet foreseeable demands for e-science support. These include the widespread adoption of optical networking and support for grid-based applications, as well as the provisioning of significantly higher international bandwidth to meet the declared needs for international collaboration in a number of fields including those mentioned above.

Primary authors : Dr. STANTON, Michael (Rede Nacional de Ensino e Pesquisa - RNP) ; Ms. MACHADO, Iara (Rede Nacional de Ensino e Pesquisa - RNP) ; Dr. FAERMAN, Marcio (Rede Nacional de Ensino e Pesquisa - RNP) ; Dr. MOURA, Ana Lúcia (Rede Nacional de Ensino e Pesquisa - RNP)

Co-authors :

Presenter : Dr. STANTON, Michael (Rede Nacional de Ensino e Pesquisa - RNP)

Track classification :

Contribution type : --not specified--

Submitted by : Dr. STANTON, Michael

Submitted on Friday 12 October 2007

Last modified on : Friday 12 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 48

LEMDist: e-learning and e-science work space

LEMDist is a implementation for remote access to laboratory equipment in a grid environment. The actual functionality for these applications include the remote data acquisition from real laboratory equipment in the grid environment. The access has been implemented for instruments with standard serial or USB interface. Experiments for Basic Chemistry and Food Engineering will be presented. The instruments are reached via authentication and authorization grid services and a interface grid device commands. Other services had been implemented for Food Engineering, they include a modelling process for freezing times of meat calculation and texture analysis from frozen meat images. Taking advantage of Grid infrastructure and experimental laboratory equipment the design model based on a categorical approach had been driven to build a technological platform to support different pedagogical approach in natural science teaching and e-science applications, implementing other services.

Primary authors : CRUZ GUZMAN, Jesus (Universidad Nacional Autonoma de Mexico, UNAM)

Co-authors : HERNADEZ DUARTE, Moises (Universidad Nacional Autonoma de Mexico, UNAM) ; GARZA RIVERA, Jose Luis (Universidad Nacional Autonoma de Mexico) ; ARJONA ROMAN, Jose Luis (Universidad Nacional Autonoma de Mexico, UNAM)

Presenter : CRUZ GUZMAN, Jesus (Universidad Nacional Autonoma de Mexico, UNAM)

Track classification :

Contribution type : --not specified--

Submitted by : CRUZ GUZMAN, Jesus

Submitted on Saturday 13 October 2007

Last modified on : Saturday 13 October 2007

Comments :

Status : SUBMITTED

Track judgments :

Abstract ID : 49

AQUAGRID: The subsurface hydrology Grid service of the Sardinian regional Grid infrastructure

AQUAGRID is the subsurface hydrology service of the Sardinian regional Grid infrastructure, designed to deliver complex environmental applications via a user-friendly Web portal. The service is oriented towards water professionals and decision makers providing them a flexible and powerful tool to solve water resources management problems and aid decision between different remediation options for contaminated soil and groundwater.
In this paper, the AQUAGRID application concept and the enabling technologies are illustrated. The heart of the service is the CODESA-3D hydrogeological model to simulate complex and large groundwater flow and contaminant transport problems. The relevant experience gained from the porting of the CODESA-3D application on the EGEE infrastructure, via the GILDA testbed (https://gilda.ct.infn.it/testbed.html), has contributed to the service prototype. Today AQUAGRID is built on top of compute-grid technologies by means of the EnginFrame Grid portal. The portal enables the interaction with the underlying Grid infrastructure and manages the computational requirements of the whole application system. Data management, distribution and visualization mechanisms are provided by the DatacroSSing Decision Support System (http://datacrossing.crs4.it). The DSS, built on top of the SRB data-Grid middleware, is based on Web-GIS and relational database technologies. The resulting reporting production environment allows the end-user to visualize and interact with the results of the performed analyses using graphs, annotated maps and 3D objects. Such a set of graphical widgets increases enormously the number of AQUAGRID potential users because it does not require any specific expertise of the physical model and technological background to be understood.

Primary authors : Mrs. LECCA, Giuditta (CRS4)

Co-authors : Mr. MURGIA, Fabrizio (CRS4)

Presenter : Mr. MURGIA, Fabrizio (CRS4)

Track classification :

Contribution type : --not specified--

Submitted by : Mr. NEUMANN CIUFFO, Leandro

Submitted on Monday 15 October 2007

Last modified on : Monday 15 October 2007

Comments :

Status : SUBMITTED

Track judgments :