



Project gridification: the UNOSAT experience

Wednesday, March 1, 2006 3:00 PM (15 minutes)

The EGEE infrastructure is a key part of the computing environment for the simulation, processing and analysis of the data of the Large Hadron Collider (LHC) experiments (ALICE, ATLAS, CMS and LHCb). The example of the LHC experiments illustrates well the motivation behind Grid technology. The LHC accelerator will start operation in 2007, and the total data volume per experiment is estimated to be

a few PB/year at the beginning of the machine's operations, leading to a total yearly production of several hundred PB for all four experiments around 2012. The processing of this data will require large computational, storage and associated human resources for operation and support. It was not considered feasible to fund all of the resources at one site, and so it was agreed that the LCG computing service would be implemented as a geographically distributed Computational Data Grid. This means, the service will use computational and storage resources, installed at a large number of computing sites in many different countries, interconnected by fast networks. At the moment, the EGEE infrastructure counts 160 sites, distributed over more than 30 countries. These sites hold 15000 CPUs and about 9PB of storage capability.

The Grid middleware will hide much of the complexity of this environment from the user, organizing all the resources in a coherent virtual computer centre.

The computational and storage capability of the Grid is attracting other research communities and we would like to discuss the general patterns observed in supporting

new applications, porting their application onto the EGEE infrastructure.

In this talk we present our experiences in the porting of three different applications inside the Grid like Geant4, UNOSAT and others.

Geant4 is a toolkit for the Monte Carlo simulation of the interaction of particles with matter. It is applied to a wide field of research including high energy physics

and nuclear experiments, medical, accelerator and space physics studies. ATLAS, CMS,

LHCb, Babar, and HARP are actively using Geant4 in production.

UNOSAT is a United Nations initiative to provide the humanitarian community with access to satellite imagery and Geographic System services. UNOSAT is implemented by the UN Institute for Training and Research (UNITAR) and manager by the UN Office for Project Services (UNOPS). In addition, partners from public and private organizations constitute the UNOSAT consortium. Among these partners, CERN participates actively providing the computational and storage resources needed for their images analysis.

During the gridification of the UNOSAT project, the collaboration with the developers of the ARDA group to adapt the AMGA software to the UNOSAT expectations was extremely important. The satellite images provided by UNOSAT have been stored

in

Storage Systems at CERN and registered inside the LCG Catalog (LFC). The files so registered have been identified with an easy to remember Logical File Name (LFN). The LFC Catalog is then able to map these LFN to the physical location of the files.

Due to the UNOSAT infrastructure, their users will provide as input information the coordinates of each image. AMGA is able to map these coordinates (considered metadata information) to the corresponding LFN of the files registered inside the Grid. Then the LFC will find the physical location of the images.

A successful model to guarantee a smooth and efficient entrance in the Grid environment is to identify an expert support to work with the new community. This person will assist them during the implementation and execution of their applications inside the Grid. He will also be the Virtual Organization (VO) contact person with the EGEE sites. This person will work together with the EGEE deployment team and with the responsible of the sites to set the services needed by the experiment or community, observing also the relevant security and access policies. Once these new communities attain a good level of maturity and confidence, a VO Manager would be identified in the users community.

This talk will report a number of concrete examples and it will try to summarize the

main lessons. We believe that this should be extremely interesting for new communities in order to early identify possible problems and prepare the appropriate

solutions. In addition, this support scheme would also be very interesting as a model, for example, for local application support in EGEE II.

Primary author: Dr MENDEZ LORENZO, Patricia (CERN IT/PSS)

Presenter: Dr MENDEZ LORENZO, Patricia (CERN IT/PSS)

Session Classification: 1c: Earth Observation - Archaeology - Digital Library

Track Classification: VO management - Portals