SA4.HEP Work Plan (Part B)

1.1 Work Package "SA4 HEP"

1.1.1 Overview and Effort

Start Date	M01
End Date	M36 (preferably)
Activity Type	SVC
Partner Acronym	**Effort in Person-Months**
CERN	540 (17 FTEs – 2 FTEs through INFN)
INFN	72 (guess)

1.1.2 Objectives

- Grid infrastructure support, operation and management; services deployed on top of generic communication and computing infrastructures to build and serve virtual communities in the various scientific domains;
- Deployment, quality assurance and support of services for existing International Heavy User Communities;
- Data and resource management (including secure shared access, global scheduling, user and application support services) to foster the effective use of distributed computing infrastructures; federated and interoperable services to facilitate the deployment and wide use of digital repositories of scientific information.
- Vertical integration of the different services in support of specific virtual research communities, including virtual laboratories for simulation and specific workspaces.

1.1.3 Description of Work

Task 1: VO specific service deployment for large-scale international user communities

Building on the powerful generic infrastructure of the underlying grids that they use, the LHC experiments have developed important complementary services particularly in the areas of data and workload management, as well as in support for analysis services. Such services, which extend the capabilities of the infrastructure by exploiting knowledge of the experiment's computing model, data placement policies and/or information in metadata repositories, allow these massive international communities to maximise the benefit of the grids that they use. For example, PhEDEx, the CMS data movement system, is able to source files belonging to a larger dataset (a concept that does not exist at the underlying FTS layer) from alternative sites, leading to additional robustness and performance.

As much as 50% of the data – possibly more – may be retrieved from such a source: functionality that cannot – by design – be provided at the FTS layer.

Whilst these experiment-specific solutions typically address today individual VOs, each such VO consists of thousands of users worldwide and corresponds to significant usage of grid infrastructure. Furthermore, experience has shown that some such solutions not only become adopted by other HEP VOs but later spread into additional communities and should be considered an important source of innovation (the driving force being the raw requirement but the actual realisation through the significant competence within these VOs must also be considered as a major source of "unfunded" effort that can benefit other communities worldwide). The "classic example" in this vein is AliEn, which is a lightweight Grid framework that is built around Open Source components using the Web Services model. It has been initially developed by the ALICE collaboration as a production environment for the simulation, reconstruction, and analysis of Physics data in a distributed way. The architecure of AliEn provided a blueprint[i] and a starting point for developing the gLite architecture. Other examples include the use of generic pilot jobs, originally adopted by both ALICE and LHCb but now becoming the preferred mode for all 4 LHC experiments.

This activity is of particular importance now as we enter the exploitation phase of the world's largest scientific machine – the Large Hadron Collider at CERN – and will allow us to capitalize on the investment made by the European Commission through its funding of three phases of the Enabling Grids for E-sciencE project. This has resulted in large scale production use of world-class Grid-based solutions by many key communities and has established Europe's leadership in this area. In the short to medium term it is expected that this will lead to significant advances in our basic understanding of the Universe around us, whereas in the longer term major spin-offs, both related to the advances in science as well as in Information Technology, can be expected.

Virtual Organization	FTE requirement	Tasks
ALICE	2	AliEn services, covering workload management, data management (built upon standard components), integration of these services into WLCG services, VO box services and support.
ATLAS	3.5	Distributed Data Management system, built upon underlying services such as FTS and LFC and monitored via the Dashboard framework.
CMS	3.5	PhEDEx Data Service, CMS Remote Analysis Builder (CRAB) and related workload management and data services.
LHCb	2	DIRAC workload management and data management services.

Task 2: Support of Generic Frameworks for the Heavy User Communities

In order to perform production and analysis tasks across a highly distributed system crossing multiple management domains powerful and flexible monitoring systems are clearly needed. To respond to the LHC experiments' requirements in this area, the experiment Dashboard monitoring system was originally developed in the context of the EGEE NA4/HEP activity. This framework, not only supports multiple grids / middleware stacks, including glite, OSG and ARC (NDGF), but is also sufficiently generic as to address the needs of multiple other communities including but not limited to HUCs. Furthermore, it covers the full range of the experiments' computing activities: job monitoring, data transfer (see FTS and VO services above) as well as site commissioning. It also addresses the needs of different categories of users, including:

- Computing teams of the LHC VOs;
- VO and WLCG management;
- Site administrators and VO support at the sites;
- Users running their computational tasks on the grid infrastructure.

Future Evolution:

The future evolution of the project is driven by the requirements of the LHC community which is preparing for LHC data taking at the end of 2009.

The main strategy is to concentrate effort on common applications which are shared by multiple LHC VOs but can also be used outside the LHC and HEP scope. Examples of such applications are: generic job monitoring application and user task monitoring, FTS monitoring, site status board, VO-specific site availability based on the results of tests submitted via Site Availability Monitor (SAM).

Impact:

Reliable monitoring is a necessary condition for establishing and maintaining production quality of the distributed infrastructure. Monitoring of the computing activities of the main communities using this infrastructure in addition provides the best estimation of its reliability and performance.

The importance of flexible monitoring tools focusing on the applications has been demonstrated to be essential not only for "power-users" but also for single users.

For the power users (such as managers of key activities like large simulation campaigns in HEP or drug searches in BioMed) a very important feature is to be able to monitor the resource behaviour to detect the origin of failures and optimise their system. They also benefit from the possibility to "measure" efficiency and evaluate the quality of service provided by the infrastructure. Single users are typically scientists using the Grid for analysis data, verifying hypotheses on data sets they could not have available on other computing platforms. In this case the monitoring / dashboard is a guide to understand the progress of their activity, identify and solve problems connected to their application.

This is essential to allow efficient user support by "empowering the users" in such a way that only non-trivial issues are escalated to support teams (for example, jobs on hold due to scheduled site maintenance can be identified as such and the user can decide to wait or to resubmit).

Ganga is an easy-to-use frontend for job definition and management, implemented in Python. It has been developed to meet the needs of the ATLAS and LHCb for a Grid user interface, and includes built-in support for configuring and running applications based on the Gaudi / Athena framework common to the two experiments. Ganga allows trivial switching between testing on a local batch system and large-scale processing on Grid resources.

A job in Ganga is constructed from a set of building blocks. All jobs must specify the software to be run (application) and the processing system (backend) to be used. Many jobs will specify an input dataset to be read and/or an output dataset to be produced. Optionally, a job may also define functions (splitters and mergers) for dividing a job into subjobs that can be processed in parallel, and for combining the resultant outputs. Ganga provides a framework for handling different types of application, backend, dataset, splitter and merger, implemented as plugin classes. Each of these has its own schema, which places in evidence the configurable properties.

As it is based on a plugin system, Ganga is readily extended and customised to meet the needs of different user communities. Activities outside of ATLAS and LHCb where Ganga is successfully used include Geant4 regression tests and image classification for web-based searches.

The number of Ganga users has steadily increased and today there are several hundred grid users using the tool in their daily work, some 25% of whom are not from HEP VOs. Whilst these other VOs and the successful “gridification” of numerous associated applications in a wide range of fields including Fusion, Material Sciences, Accelerator Studies and Biomedical applications, the effort requested here would focus on production service deployment to the WLCG VOs ATLAS and LHCb in the critical early years of the LHC’s operation.

Framework	Manpower Requirement
Dashboards	4 FTEs
Ganga	2 FTEs

---

[i] See http://glite.web.cern.ch/glite/alien/.