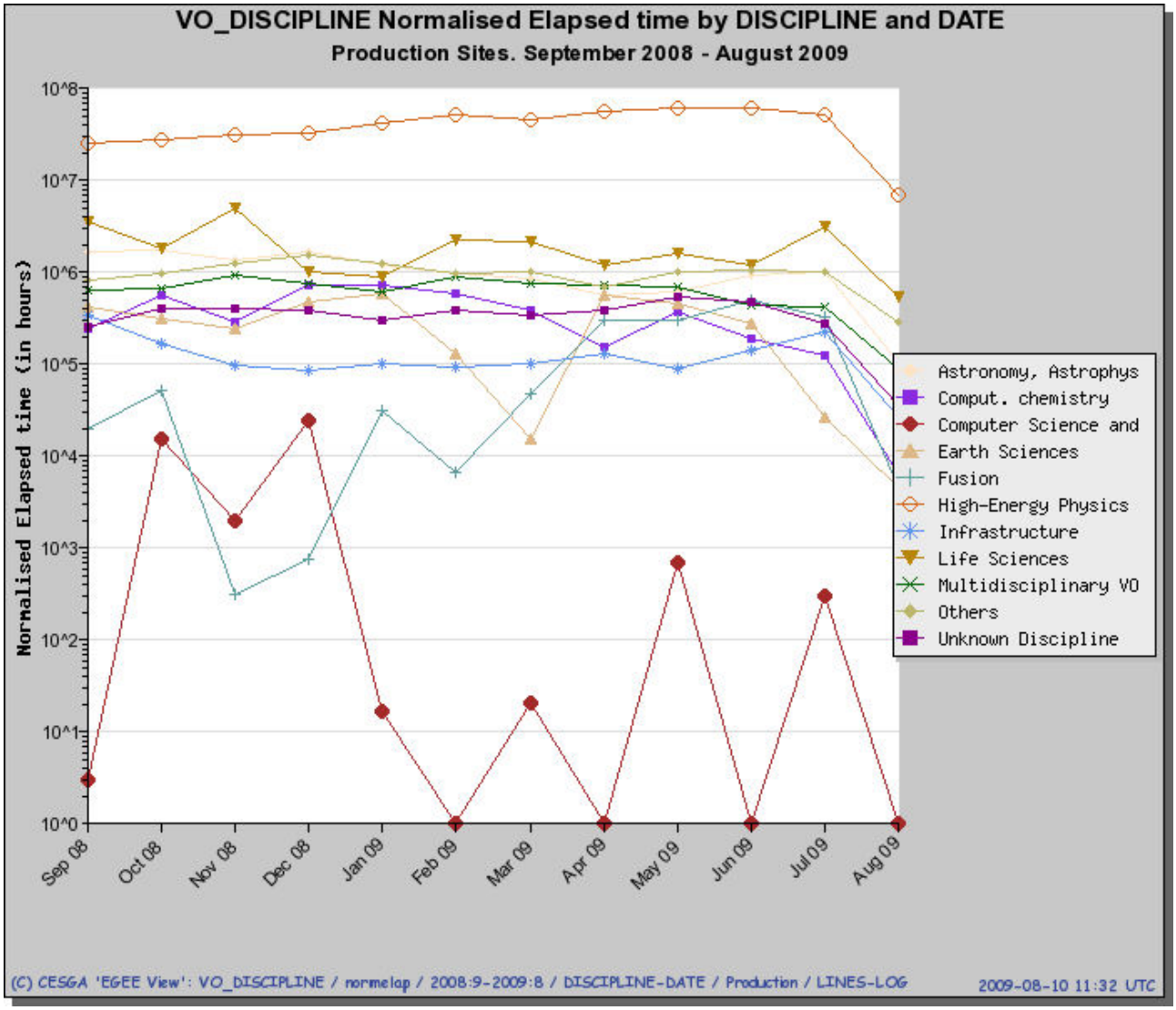
1.5.4 Activity SA4: Services for Communities of Heavy Grid Users

1.5.4.1 Activity Description

The support of the services needed by the present heavy users of the grid, is a key Service Activity in this Proposal, in response to the 1.2.1.2 sub-call. These services are build over the basic middleware components that in year 2008 (and beforehand) are provided by EGEE and other interoperable projects, and include both middleware services (e.g. FTS) and the frameworks that relay on the elementary components and services for performing functions tailored for specific communities (e.g. the Dashboard for the monitoring of the grid for specific VO's). The EGI support to these services includes their deployment, operation, and evolution for adapting to the needs both of the community that originally used the service, and of a larger set of users. The deployment and operation is part of this Activity, while the evolution is treated in JRA2

The European Grid Infrastructure is presently actively used by some scientific disciplines that heavily relay on this infrastructure for carrying on their research activity. The graph below, showing the monthly usage in the year from September 2008 to August 2009 by the different Scientific disciplines, gives clear indication of this massive usage.



The graph is taken from the EGEE accounting portal, the unit is Hours.1kSI2k elapsed time; in very first approximation a 10^6 in the vertical scale corresponds to about 2000 typical processors continuously working for the full month.

The Communities identified as Heavy Users Communities (HUCs) are

- High Energy Physics (HEP)
- Life Sciences (LS)
- Astronomy and Astrophysics (AA)
- Computational Chemistry and Material Sciences (CCMT)
- Earth Sciences (ES)

Besides the massive usage of the Grid, the HEP and LS have played a pilot role in EGEE, giving decisive contributions for bringing the grid at production quality level, via feedback on the deployed services efficiency and functionalities, stress tests of the infrastructure and selected components, etc. It is expected now that all the HUCs will be able to play a similar role for the services of their interest included in this Activity, and at least in some case for specific aspects of the EGI grid (e.g. further scalability in case of HEP)

A key element for the satisfaction of the needs of these users, and for ensuring the continuity and enlargement of the role of "pilot users", vital for the grid, is providing them with the higher level services they need. In general these services are already used by the community, with some support by EGEE and the other interoperable projects.

The first objective for SA4 is that Heavy Users Communities experiment no disruption of their activity with the transition of the e-infrastructure to the EGI support. In fact EGI aims at increasing the satisfaction of these main users, also in view of expanding the Grid usage within the disciplines they belong to, and toward new disciplines. SA4 will grant to the communities the continuity and good integration of services in the general grid, and will shape the provision of these services so that they are efficiently used by all the presently interested user communities and extend their usage to new communities.
SA4 will work so that most services will become more standard and easy to configure, deploy and operate. Some services will retain interest only for a small fraction of the user communities, and in perspective will be supported primarily by these communities with reduced contribution from EGI, while some other services will become more general in their use and will in the future become integral part of the EGI general infrastructure.

1.5.4.2 Assumptions on the general services available outside SA4

The services provided in SA4 are the services needed by the HUCs and not included in the general EGI infrastructure, nor supported by any other trusted provider.

The following Table.SA4.1 provides a list of the general services available outside SA4; when the service is not included in the general EGI infrastructure the provider is indicated

Compute Element from gLite and ARC	
LCG CE?	?
SE from GLite and ARC	
dCache	?
StoRM	
Castor	CERN
SRM	
GridFTP	
Info.services from gLite and ARC	
Accounting from gLite and ARC	
Authorization Services from gLite and ARC (e.g. VOMS, MyProxy, SCAS LCMAPS, LCAS, gLExec)	
Work Load Management services (e.g. WMS)	

Each HUC provides the indication of with detailed services in the list are of interest for it, specifying the use, the importance etc.

1.5.4.3 Task Description

1.5.4.3.1 TSA4.1 SA4 Management

The SA4 management comprises the full time Activity Manager, with the responsibility of supervising the services and coordinating their provision with the relevant communities.
He/she is assisted by representatives of the relevant communities, with the expertise necessary for providing feedback on the working of the services and input on the modification that may be needed, in configuration, operation and deployment of the services for the specific communities.
The effort needed from the representatives of the communities depend from the amount and complexity of the services requested by the specific community

The SA4 manager will be a member of the MCB

1.5.4.3.2 TSA4.2 Hosting of Community specific Services

Provision and operation by a small number of NGIs of Core Grid Services (O-N-8) explicitly needed to support this user community, but of potential benefit to other communities.
These centres will be experts and provide an SLA around the hosting of services such as FTS, LFC, Hydra, AMGA and VO specific services.
*This list is an **example** of services and the writers from the specific communities should introduce here the services of their interest,if any; for each service please provide*

- *The description and the motivations*
- *The evaluation of the effort*
- *The NGI(s) proposed for hosting and their share of the effort*
- *Indication of SLA/SLD*
- *Interested HUC(s)*

The four main experiments at CERN's Large Hadron Collider (LHC) – namely ALICE, ATLAS, CMS and LHCb – rely on a worldwide virtual computing facility that is implemented using grid technology. The architecture of this system is based upon a tier model that consists of CERN, ten major regional or national data-intensive "Tier1" centres as well as over one hundred "Tier2" sites that are located close to end users. In order to facilitate data movement between these sites – which has run at 1PB/day over prolonged periods – and to enable the required data to be located, file transfer and catalog services are required. These are implemented on top of *glite* components: namely the File Transfer Service (FTS) and the LHC (aka "local") File Catalog (LFC). These services run at the Tier0 and Tier1 sites and involve both application servers as well as backend database systems. Whilst the LFC has been adopted by numerous other HEP and non-HEP communities, the use of the FTS is currently predominately WLCG only. However, as more and more communities face limitations from power and cooling and move to distributed systems, the need for reliable wide-area file transfer services over high capacity network links can be expected to grow. It may also be needed in *cloud* computing environments to transfer data both into and out of "the cloud".
Both of these middleware components were carefully designed with service deployment and reliability in mind: they permit resilient and scalable deployment models (load balanced front-ends, database clusters as back-ends etc, "rolling upgrades" of both middleware and underlying O/S components and even migration of hardware!) WLCG Tier1 sites are nonetheless recommended to have 1 FTE to manage the database services behind these and other required services (detector conditions and / or storage services), whereas the Tier0 runs a more complex setup with – in the case of LHCb – data replication for the LFC to read-only replicas at the Tier0 and all Tier1 sites. Thus, to deliver these services a total of 1 FTE is required per Tier1 with a minimum of 2 FTEs at the Tier0. The sites involved are CERN and the 6 WLCG Tier1 sites in Europe (IN2P3 in France, FZK in Germany, NIKHEF/SARA in the Netherlands, NDGF in the Nordic countries, PIC in Spain and RAL in the UK – the remaining Tier1 sites in North America and Asia would not expect to be funded through this tasks, although a similar amount of effort, adjusted to the number of VOs supported by each site, is required).

1.5.4.3.3 TSA4.3 Hosting of VO specific Services

Provision and operation by a small number of NGIs of Core Grid Services (O-N-8) explicitly needed to support this VO, but of potential benefit to other users.
Justified if the VO users are a relevant fraction of the Grid users and/or they use a relevant fraction of the grid resources, or if the service is foreseen to become of more general interest during this Project.
The writers from the specific communities should introduce here the services of their interest, for each service please provide

- *The description, and the justification*
- *The evaluation of the effort*
- *The NGI(s) proposed for hosting and their share of the effort*
- *Indication of SLA/SLD*
- *Interested HUC(s)*

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 total estimated effort to support these services worldwide is 10FTEs.

1.5.4.3.4 TSA4.4 Support of Frameworks

The frameworks integrate different components and services for performing functions tailored for specific communities or VOs
An example are the VO Dashboards: VO Dashboards have been found to be very useful by large VOs to provide a VO view of the infrastructure for their community. Other examples may be GANGA, PHEDEX, DDM, WISDOM

This task in case of Dashboard includes the hosting of the service and the integration and development of VO specific tests, driven by the particular user community, necessary to verify the correct functioning of the infrastructure for their work. This will also draw on the generic service monitoring infrastructure and tests maintained by the NGIs.

The content is analogue in the case of the other Framework and should be described by the specific writers

The writers from the specific communities should introduce here the frameworks of their interest, if any; for each framework please provide

- *The description, and the justification*
- *The evaluation of the effort*
- *The NGI(s) proposed for contributing and their share of the effort*
- *Indication of SLA/SLD if applicable*
- *Interested HUC(s)*

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, which not only supports multiple grids / middleware stacks including glite, OSG and ARC (NDGF), but also is sufficiently generic as to address the needs of multiple other communities including but not limited to HUCs, 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;
- Physicists running their analysis tasks on the EGEE 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.

Impact:

Reliable monitoring is a necessary condition for 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 hypothesis on data sets they could not have available on other computing platform. In this case the monitor 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).

Effort: 4 FTEs

Ganga is a flexible and user-friendly tool to enable large user communities to move away from a niche usage of the Grid. It supports the access to major grid infrastructures (notably EGEE, OSG and NDGF) and to all major batch systems.EGEE has supported the Ganga project via the EGEE NA4 HEP activity.
Detailed information on the Ganga system has been described in a recent publication "Ganga: a tool for computational-task management and easy access to Grid resources", J. T. Moscicki et al. Users can seamlessly run their applications by combining the usage of both Grid and batch resources. It is a genuine tool to run applications on multiple Grids (interoperability) and to protect users from changes of the middleware layer (new APIs). Ganga's strengths are in its non-intrusive nature, allowing its users to run a applications without change on the Grid.
Impact:
Ganga is one of the most useful elements in our toolbox to support large user communities, enable new ones and collaborate with other infrastructure projects. The return of investment is ensured by the large user communities, by the existing and future activities in the area of support and by the outreach to new communities.
Effort: 2 FTEs