EUROPEAN MIDDLEWARE INITIATIVE

DJRA1.2.1 - DATA AREA WORK PLAN AND STATUS REPORT

EU DELIVERABLE: D5.2.1

Document identifier:	**EMI_DJRA1.2.1-1277615-Data_Area_Work_Plan_v1.0.docx**
Date:	**31/07/2010**
Activity:	**JRA1**
Lead Partner:	**DESY**
Document status:	**Final**
Document link:	**http://cdsweb.cern.ch/record/1277615?ln=en**

Abstract:
This deliverable contains the work plan of the Data Services technical area, which will be updated every year. It also covers the state-of-the-art in this technical area.

Delivery Slip

	Name	Partner / Activity	Date	Signature
From	Patrick Fuhrmann	DESY/JRA1	26/08/2010	
Reviewed by	Jens Jensen	STFC/External	26/11/2010	
Approved by	PEB		29/11/2010	

Document Log

Issue	Date	Comment	Author / Partner
1	1/7/2010	First Draft (TOC)	Patrick Fuhrmann/DESY
2	5/7/2010	State of the Art	Patrick Fuhrmann/DESY, Riccardo Zappi/INFN, Ralph Müller-Pfefferkorn/TUD, Oliver Keeble/CERN
3	12/7/2010	Work plan	Patrick Fuhrmann/DESY
4	3/8/2010	Harmonization and Evolution	Patrick Fuhrmann/DESY
5	11/8/2010	Detailed work plan	Same as (2)
6	13/8/2010	Merged review by EMI data and JRA1 lead	Patrick Fuhrmann
7	13/8/2010	Includes suggestions by Alex Sim, LBL	Patrick Fuhrmann
8	26/8/2010	First complete version sent for review	Patrick Fuhrmann
9	18/11/2010	First revision after J. Jensen review	Patrick Fuhrmann
10	26/11/2010	Second revision after J. Jensen second review	Patrick Fuhrmann

Document Change Record

Issue	Item	Reason for Change
1		
2		
3		

TABLE OF CONTENTS

1. INTRODUCTION

1.1. PURPOSE

This document represents the first deliverable in the data area of EMI. Its purpose is to identify components and their product teams within this area and describe the state of the art of those components at the point in time when the EMI project started. Based on this, the document elaborates on the work plans of the different product teams, responsible for the EMI data area components.

1.2. DOCUMENT ORGANISATION

According to the description of this deliverable, this document describes the state of the art of components within the data area of EMI, as well as the work plan that is being considered for development. The work plan is split into three sections. The section on *harmonization* describes work on which the production teams have to collaborate while the section on *evolution* concentrates on the plans for the different individual data components. Finally, a detailed schedule is presented for the first year of EMI and a brief plan for the remaining project years.

1.3. DATA AREA

The EMI data area covers

- data access by standard high performance wide and local area protocols,
- data catalogues, keeping track of meta data including data location,
- data storage, including mechanism to steer *access latency* and *retention policy* of stored data
- and scheduled data transfers between storage endpoints.

The goal of the harmonization in EMI, as described later in this document, is to make those building blocks of the data area work together seamlessly and to provide a professional service to EMI customers.

1.4. REFERENCES

R 1	GLUE2 Specification, OGF Grid Final Document No.147 - http://www.ogf.org/documents/GFD.147.pdf
R 2	DMC - http://www.nordugrid.org/documents/dmc.pdf
R 3	SRM, OGF Grid Final Document No.129 - http://www.ogf.org/documents/GFD.129.pdf
R 4	ARC UI - [http://www.nordugrid.org/documents/arc-ui.pdf
R 5	UNICORE Data Finder [http://www.dlr.de/sc/desktopdefault.aspx/tabid-1273/1756_read-3140/] [http://sourceforge.net/projects/datafinder/]
R 6	UNICORE [http://www.unicore.eu/]
R 7	dCache reference [http://www.dcache.org]
R 8	gLite [http://www.glite.org]
R 9	ARC [http://www.knowarc.eu/middleware.html]

1.5. DOCUMENT AMENDMENT PROCEDURE

This document can be amended by the authors further to any feedback from other teams or people. Minor changes, such as spelling corrections, content formatting or minor text re-organisation not affecting the content and meaning of the document can be applied by the authors without peer review. Other changes must be submitted to peer review and to the EMI PEB for approval.

When the document is modified for any reason, its version number shall be incremented accordingly. The document version number shall follow the standard EMI conventions for document versioning. The document shall be maintained in the CERN CDS repository and be made accessible through the OpenAIRE portal.

1.6. TERMINOLOGY

SE	GRID storage element	Abstraction of a data storage endpoint. Requires a defined minimum set of data access, control and information protocols.
SRM	Storage Resource Manager	Definition of a remote data management protocol defined by OGF
WLCG	World wide LHC Grid Computing	Grid optimized for High Energy Physics (HEP) tuned for experiments around the LHC
OGF	Open Grid Forum	Standardization body in the GRID domain
OSG	Open Science Grid	US GRID initiative

2. EXECUTIVE SUMMARY

The EMI data area is supposed to cover data access by standard high performance wide and local area protocols, data catalogues, keeping track of meta data including data location, data storage, including mechanism to steer *access latency* and *retention policy* of stored data and scheduled data transfers between storage endpoints. The goal of the harmonization in EMI, as described later in this document, is to make those building blocks of the data area work together seamlessly and to provide a professional service to EMI customers.

The vast majority of components in the EMI data suite have proven to work sufficiently well in their particular middleware environment. ARC, gLite and dCache managed the expected data flow and storage since the LHC has started producing data and UNICORE is in production since 2003, but having traditionally a more HPC-driven computing focus where cluster file systems take care of the majority of data-related issues. Thus the challenge within EMI is having components from different middle-wares seamlessly interoperating. One of the motivations for this is that HPC environments (managed by UNICORE) can actually use LHC data for computing that is in turn not conveniently for end-users possible today. Therefore, within *EMI data* interoperability will be achieved by implementing or introducing standard interfaces wherever possible to guarantee not only interoperability within EMI but to external, possibly industry components as well. In some of those areas, the different components will evolve to that common goal independently, but following specifications that are considered to be useful by EMI. In other areas, those goals can only be achieved by a close collaboration between the middleware or product teams. For the latter, the following areas of work within the data area of EMI (aka EMI data) have been identified.

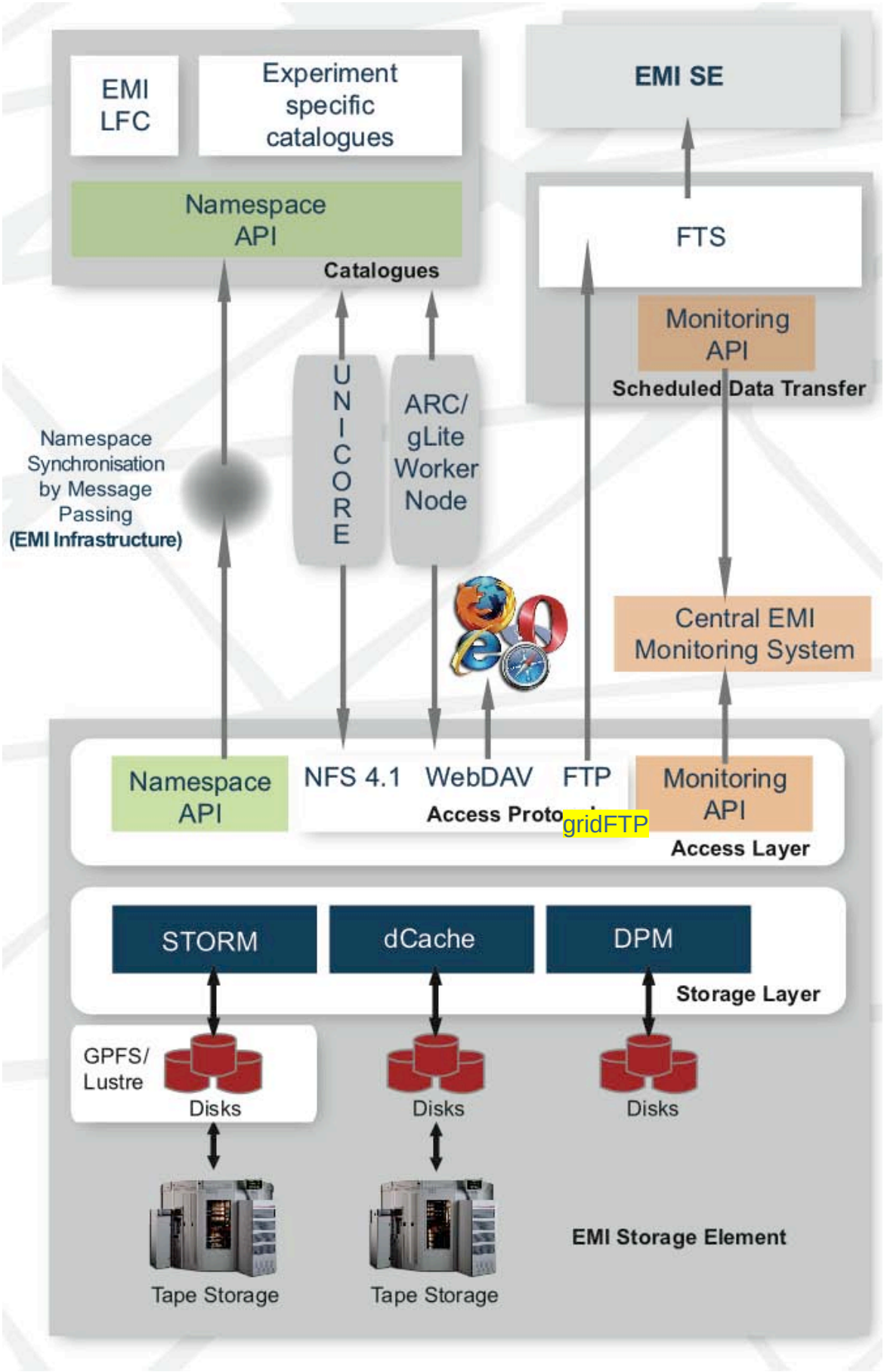
Most important is certainly providing access to data across EMI storage elements by offering POSIX semantics and mechanisms. *EMI data* agreed that NFS 4.1 is the most future proof approach in this area where native POSIX access it not available. For local and remote Web access, implementing either http(s) or WebDav on top of http is envisioned and part of the work plan.

Besides accessing data, remotely managing data is indispensable for large-scale global storage. The Open Grid Forum (OGF) defined a protocol named as the Storage Resource Manager (SRM) [3], which has been implemented in the WLCG framework to tackle this issue. However, experiences with this protocol have exposed issues with the specification, which need to be followed up on in order to ensure a more stable interoperability between storage elements. The main aspects will be to standardize the protocol security mechanism and to simplify the specification based on the experiences gained by known use-cases.

As a consequence of a decision made at a very early stage in the design of Grid Storage in the LHC community, file catalogue and storage element namespace get out of sync over time. The currently applied solution for this problem doesn't scale to the extent being required by EMI customers. *EMI data* will propose and implement an improved solution.

Within Grid infrastructures, the wide variety of components publishes information to enable clients to select the most appropriate service. GLUE [1] is an agreed open standards-based schema to publish this kind of metadata. Most recently, a migration from GLUE 1.3 to the more powerful GLUE 2.0 has been proposed. *EMI data* will allow data components to gradually and seamlessly migrate to GLUE 2.0.

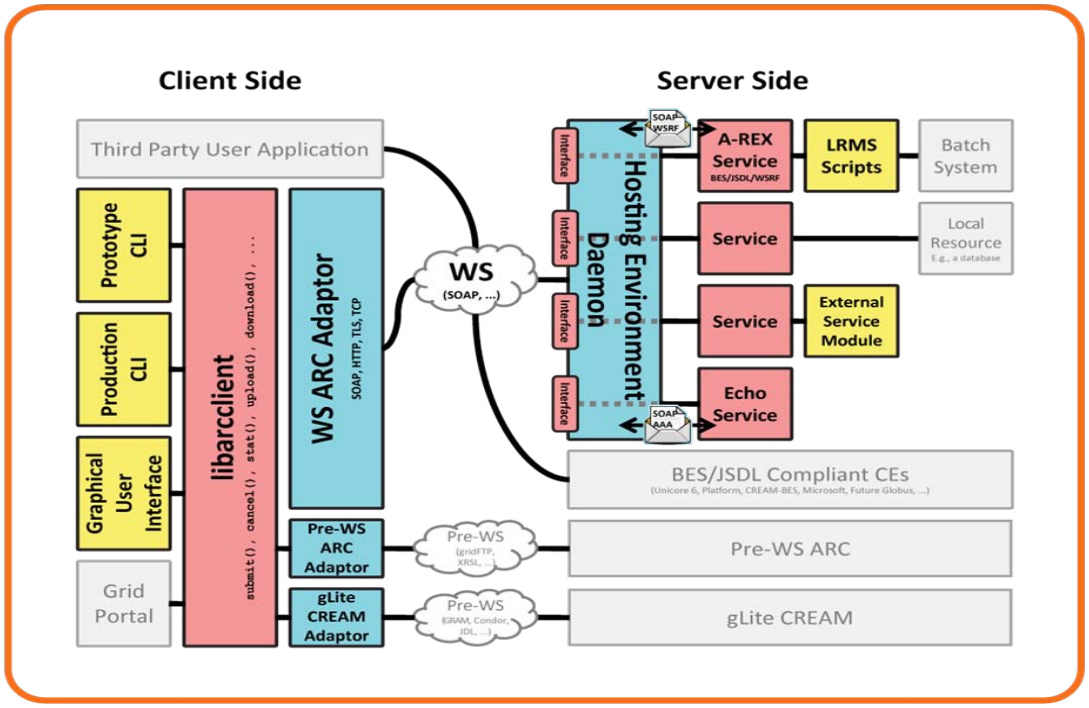
Another area of harmonization within *EMI data* is the consolidation of client data access libraries and the integration of features of the EMI ARGUS authorization system. Beside activities based on close interaction between components and consequently requiring establishing cross product-team working groups, each data component has room to evolve independently. This will happen in the context of monitoring, accounting, manageability and usability.


EMI LFC
Experiment specific catalogues
Namespace API
Catalogues
EMI SE
FTS
Monitoring API
Scheduled Data Transfer
Namespace Synchronisation by Message Passing (EMI Infrastructure)
UNICORE
ARC/ gLite Worker Node
Central EMI Monitoring System
Namespace API
NFS 4.1
WebDAV
FTP
Access Proto
gridFTP
Monitoring API
Access Layer
STORM
dCache
DPM
Storage Layer
GPFS/ Lustre
Disks
Disks
Disks
Tape Storage
Tape Storage
EMI Storage Element

3. STATE OF THE ART OF EMI DATA COMPONENTS

This section describes the current state of the art in the EMI data area.

3.1. ARC



ARC components

The libarcdata2 library and the corresponding arc* data clients provide a uniform way to move data from source to sink, supporting different data transfer protocols. This is a successor of the previous ng* data clients which are still widely used and will be maintained.

The libarcdata2 library has a modular structure to support different data transfer protocols. The core libarcdata2 library does not introduce any additional external dependencies. The plug-ins (DMCs) [2] for specific data access protocols can however have various external dependencies. This separation of external dependencies from the core library helps to reduce a minimum set of requirements for ARC, while allowing the support for additional access protocols requiring special dependencies to be installed by those who need it.

Most of these components and the clients are also available on different platforms (Linux, Windows, Mac, Solaris), and the libarcdata2 is also available via Python, and to some extent from Java.

Currently the following transfer protocols and metadata servers are supported: (for more information, see [4]

- **ftp**: File Transfer Protocol (FTP)
- **gsiftp**: GridFTP, the Globus-enhanced FTP protocol with security, encryption, etc.
- **http**: ordinary Hyper-Text Transfer Protocol (HTTP) with PUT and GET methods using multiple streams
- **https**: HTTP with SSL v3

- **httpg**: HTTP with Globus GSI
- **ldap**: ordinary Lightweight Data Access Protocol (LDAP)
- **rls**: Globus Replica Location Service (RLS)
- **lfc**: LFC catalog and indexing service of EGEE gLite
- **SRM**: Storage Resource Manager (SRM) service
- **file**: local to the host file name with a full path
- **arc**: for the Chelonia storage service

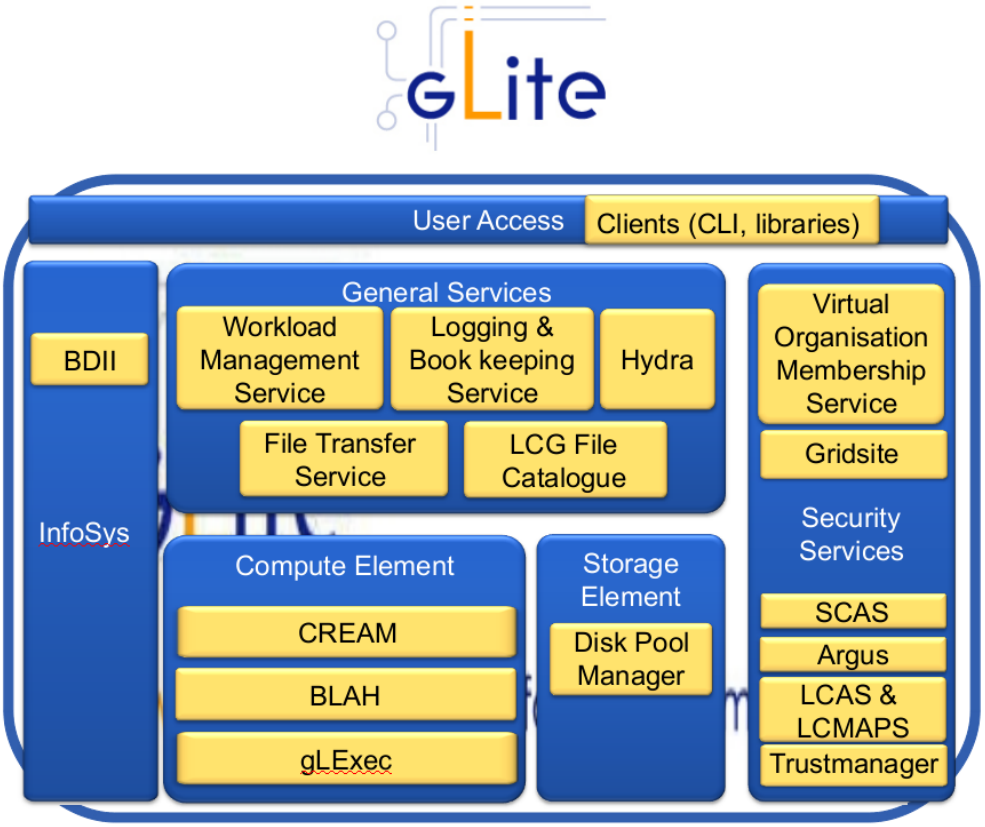
The ARC data clients provide posix-like access to EMI storage elements. The clients are as follows:

- **arcls** is a simple utility that allows to list contents and view some attributes of objects of a specified (by a URL) remote directory.
- **arccp** is a powerful tool to copy files from one data source to another data sink where both can use different protocols.
- **arcrm** is a command that allows users to erase files at any location specified by a valid URL.

In the ARC data management model all data transfer is performed on the Computing Element front-end on each site before and after the job runs on the worker node. An essential feature of ARC is the ability to dynamically cache input files on the front-end. When the front-end receives a job it first checks if the job is in cache and if not downloads the file to cache. Once a file is in cache it can be made instantly available on worker nodes by linking or copying from the cache to the worker node. To increase efficiency and satisfy the requirement that jobs go to data, an ARC Cache Index (ACIX) central service is used to store locations of cached files by pulling Bloom filters of the caches from the front-ends. Job brokers can then query ACIX and send jobs to sites where jobs are cached.

3.2. GLITE

gLite offers an integrated set of tools and services for Grid data management, with solutions for data storage, catalogues and reliable file transfer.



3.2.1 Disk Pool Manager, DPM

The Disk Pool Manager (DPM) offers a simple solution for a disk-only Storage Element (SE), supporting SRM for storage management, GridFTP for transfer, and a number of access protocols (HTTP, RFIO & xroot). DPM supports X509 and Kerberos authentication and VOMS authorization. DPM is installed at over 200 sites grid-wide.

3.2.2 LCG File catalogue (LFC)

The LCG File catalogue (LFC) offers a hierarchical view of files to users, with a UNIX-like client interface, and provides logical and physical mappings for file identifiers. LFC supports authorization on its namespace and is available for both MySQL and Oracle (which can be used with streams to implement a read-only distributed catalogue). It provides C and python APIs. LFC can be used locally, or as a global catalogue holding details of files stored on Grid storage elements.

3.2.3 Data access clients (gfal/lcg_util)

The clients, gfal/lcg_util, provide programmatic (posix-like) and command-line interfaces to storage and catalogues, enabling interaction with EMI storage elements via a single interface. SRM is supported, as is integrated storage and catalogue (LFC) manipulation.

3.2.4 File Transfer Service (FTS)

The File Transfer Service (FTS) is a data movement service for administering, monitoring and performing the transfer of files between storage elements. Bulk, asynchronous requests for transfers can be registered and will then be carried out according to VO policy, using either GridFTP or SRMCopy. Users interact with the service either via the Web Service API or via the provided clients. FTS currently manages the distribution of LHC data from CERN to partner sites.

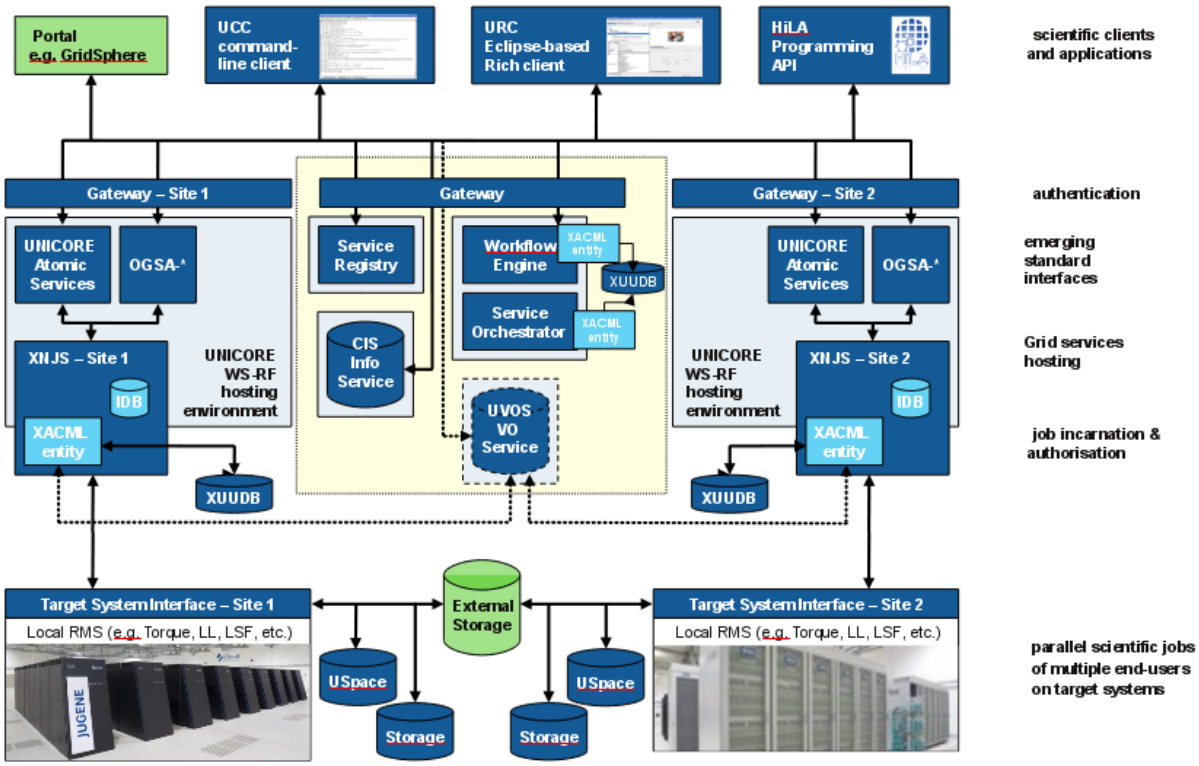
3.2.5 StoRM

StoRM is a service compliant with the standard SRM v2.2. It is able to manage any POSIX file systems supporting Access Control List (ACL) mechanisms, including high performance storage systems based on a cluster file system such as GPFS from IBM and Lustre from Oracle. StoRM decouples the file system from the SRM service, in the sense that it does not incorporate the storage system but it just relies on it; in other words StoRM specifically provides only management functionalities, but makes them available via a standard-compliant service interface. StoRM uses the underlying file system features in two ways:

1. The file system is used as final enforcer of authorization information (allowing the use of native access to the file system by authorized users and applications via ACLs as explained below)
2. The file system is used as a natural store and a manager of namespace metadata. Consequently StoRM does not make use of a database to store the namespace of managed resources (files).

StoRM uses the authorization mechanism offered by the underlying file system (ACL) to allow local and direct access using the native POSIX protocol (i.e. “file://”). Applications using standard POSIX operations without using external data access library (e.g. RFIO) considerably benefit in terms of I/O throughput, mostly in the presence of high performance cluster file systems. In addition, StoRM supports other non-standard Grid access protocols offered by data access libraries. (e.g. RFIO). Currently the supported data access and transfer protocols are: “file://”, “rfio://”, “gsiftp://”, “root://”.
The latest stable version of StoRM (v1.5.x) enables the management of hierarchical storage resources through a generic interface. In this configuration it is currently used at the Italian Tier-1 (CNAF) in Bologna in the context of LHC, where it manages a hierarchical system based on GPFS and Tivoli Storage Manager (TSM).

3.3. UNICORE

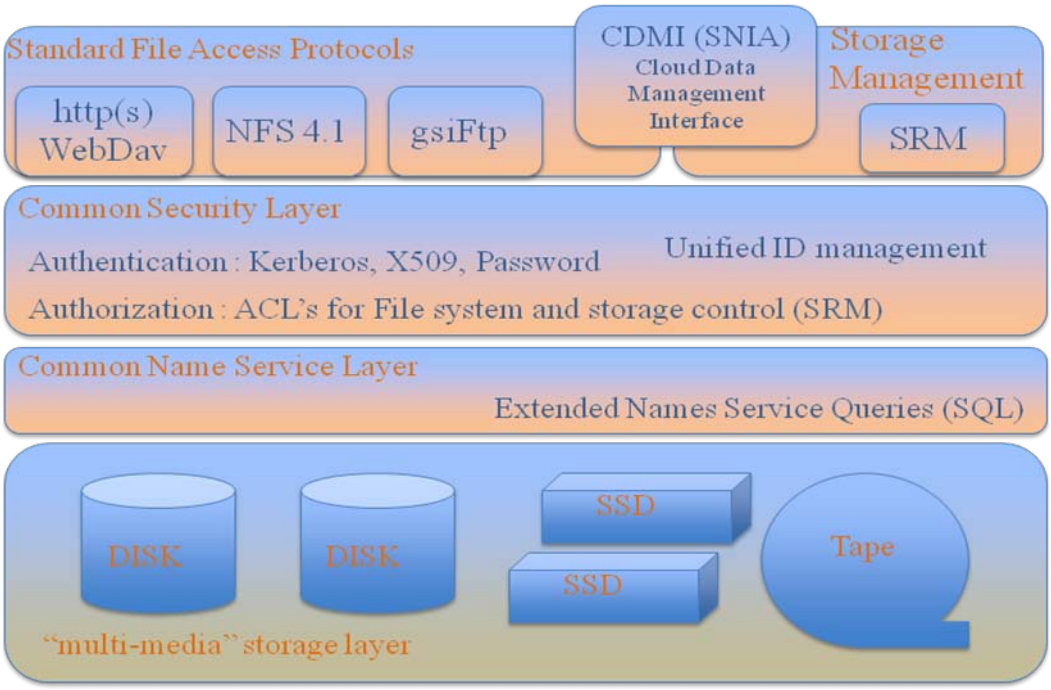


One of the available Web services of UNICORE is the Storage Management Service (SMS) that exposes storage resources to the system user. It is also able to initialize a file transfer between a user and a server or between servers. The transfer itself is done by the UNICORE File Transfer Service (UNICORE-FTS).

SMS defines an abstract file system like view that allows to plug-in any storage system. It provides operations such as listing a directory, copying, deleting and moving files and directories and changing permissions. Various back ends are already available (e.g. file system, Apache Hadoop, iRODS in prototype status, DataFinder). It is designed for extensibility (e.g. Amazon S3).

The UNICORE FTS allows the import/export of files from/to the client and send/receive files from/to other servers. Built-in are several file transfer protocols: BFT transfer (based on HTTPs, single open port needed, simple interface with bulk write and read supports byte ranges) and OGSA ByteIO (uses SOAP messages, single open port needed, rich interface with POSIX-like, block read/write, etc). Alternative mechanisms can be plugged in (e.g. UDT, GridFTP, parallel HTTP).

3.4. DCACHE



dCache design

dCache is a data management technology designed for storing, retrieving and managing huge amounts of data, distributed among a large number of heterogeneous server nodes, under a single virtual file system tree. A number of standard methods are offered to access and manage data. Depending on the persistency model, dCache provides mechanisms for exchanging data with backend (tertiary) storage systems and by doing so simulates unlimited direct access storage space. Data exchanges to and from the underlying Hierarchical Storage Systems performed automatically and transparently to the user. In the context of managing data, dCache provides space management, as described in the SRM 2.2 specification, data location management, dataset replication, hot spot determination and recovery from disk or node failures.

The dCache system is freely available directly from the dCache.org Web pages, as well as through the gLite and the OSG VDT distribution. It supports a large variety of data transfer, control and information service protocols (gsiFtp, http, NFS4.1, dCap, xrootd, SRM 2.2-WLCG) as well as the information schema GLUE1.3. In order to extend the attraction to non High Energy Physics communities, dCache is focusing on the implementation of industry standards. Most recently the NFS 4.1 (pNFS) and WebDAV (Web- based Distributed Authoring and Versioning) protocol were added to the dCache protocol-suite, enabling clients to take full advantage of the distributed manner dCache is storing data. In order to make dCache storage available via cloud protocols, the dCache community is working on the integration of the Cloud Data Management Interface (CDMI) into dCache. This effort is not funded by EMI and at present it is not planned to integrate CDMI into any other EMI component. dCache is in production at 8 out of 11 WLCG Tier I and about 40 large Tier II centers. More than 50% of the entire LHC data is stored and managed with dCache installations.

4. WORKPLAN

In this section the JRA1 activities within the data area foreseen for the first year are presented as well as a very brief estimation on the activities for the remaining time of the project.

The work-plan is split into three parts.

- *Harmonization*, describing areas of common interest, including standardization efforts, where planning, designing or development requires interactions between middle-wares or components.
- *Evolution*, describing improvements of components applied individually.
- *Detailed work plan*, describing tasks and timeline

4.1. HARMONIZATION

The idea of harmonization in the area of data management is build upon two pillars. The first is to have data management components from all EMI middle-wares interact seamlessly (in harmony), building one larger functional unit. The second is to have this new, compound component interfacing to non-EMI software components by either using well defined industry standards or, if not available, by defining new interfaces in conjunction with an appropriate standardization body. The virtue of this approach is that components within the EMI data bundle will become interchangeable to the convenience of the sites, running EMI data services, and that the EMI data components offer themselves as an "in place replacement" of possibly expensive industry solutions.

Within *EMI data*, only those protocols and interfaces will be further developed which are standardized or are in the process of being standardized by a well-known standardization body. (e.g. IETF, OGF, SNIA). This approach offers end users the flexibility to choose their technologies they need through this established interoperability by using open standards. The relevant protocols are the Storage Resource Manager protocol (SRM), POSIX file access and the well-known Internet standards: http(s) and WebDAV.

A gap analysis of the existing systems has exposed areas where additional work has to be done in order to achieve those high level goals, resulting in the following efforts:

4.1.1 Catalogue Synchronization

One of the critical issues within the current data management system is the synchronization of the data location catalogues (LFC) and the content of the actual data endpoints (SE). For various reasons they diverge over time, leading to inconsistencies which often cause system failures or unnecessary data delivery delays. EMI will propose and implement solutions to allow the EMI-LFC and SE's to interact directly and to determine and fix inconsistencies. The proposed solution should be sufficiently flexible to be adaptable by non-EMI catalogue systems, e.g. catalogues of the LHC experiments. As it is very likely synchronization will take advantage of message passing systems. Consequently, collaboration with the infrastructure area of EMI is foreseen.

4.1.2 File catalogue access for UNICORE

To allow further integration of the more High Performance Computing oriented world of UNICORE and the more High Throughput Computing oriented world of the other middle-wares, UNICORE needs to get access to data in gLite/ARC storage elements. Besides of the direct data access via the access protocols mentioned below, access to the location management of the file catalogue needs to be added to the SMS.

4.1.3 Consolidation of the Storage Resource Manager (SRM) protocol

Over the last years the WLCG community, under the supervision of the Open Grid Forum (OGF) has defined a protocol to remotely manage data (Storage Resource Manager, SRM). However, as parts of the specification turned out to be ambiguous, different implementations of the SRM show slightly different behaviors. In collaboration with OGF and SE providers, not being part of the EMI bundle, EMI data will clarify the SRM v2.2 specification based on the experience in operating SRM systems for years and promote existing SE implementations accordingly. At the time being, it is only planed to improve the specification by adding more user-friendly documentation. Two documents are envisioned in addition to the OGF specification: an analysis of the interpretation of the currently existing SRM implementations and a user guide to better understand the pure specification. The user guide might use different terms to describe the SRM functionality based on misunderstandings imposed by the current documentation.

4.1.4 Replacing the Globus httpg security protocol with the SSL/X509 (https) standard.

Some components in the data management area depend on the Globus proprietary GSI protocol, instead of being based on the widely used plain https (SSL/X509) security protocol. The advantage of this decision has been that GSI supports handling X509 proxies and that delegation is part of the protocol. In collaboration with the EMI security group, the data area will decide in which areas a move towards standard SSL/X509 will be beneficial and how the migration should be organized. As not all SRM storage elements and clients are part of the EMI data bundle and a successful deployment requires all those components to cooperate, the outcome of the latter goal is not clear, but EMI will investigate solutions to guarantee interoperability.

4.1.5 Providing standard access to data through a mounted file system.

In terms of data access, EMI-data will ensure that all supported Storage Elements will make their data available via standard POSIX file access. POSIX file access is achieved in different ways by different Storage Element implementations. While DPM and dCache will provide NFS4.1 access, StoRM uses the available mechanism of the underlying file system. However, the FUSE approach and NFS 4.1 are in discussion for StoRM as well. Accessing data through a mounted file system is essential for EMI components to compete with industry solutions.

4.1.6 Providing standard access to data via http(s) and WebDav

With the evolution of the World Wide Web, http(s) has become a standard way of accessing data within organizations as well as across large distances. So the goal of EMI data is to provide the same services to its customers. For more sophisticated use cases the well-established extension, WebDav, will be implemented by dCache and, depending on customer feedback, by StoRM and DPM.

4.1.7 Publishing GLUE 2.0 information

The GLUE specification is an information model for describing Grid entities by using the natural language and UML Class Diagrams. GLUE has evolved over time and, based on the experience of running production grids, a GLUE version 2.0 was made available and agreed on, which is rich enough to cope with requirements in the various areas of the currently deployed Grid infrastructures in Europe.

4.1.8 Integration of the ARGUS EMI authorization system

ARGUS is a system meant to render consistent authorization decisions for distributed services (e.g. compute elements, portals). For the data area, the blacklisting functionality of ARGUS is of particular interest, as it allows to centrally blacklist certificate owners for all VO associated services.

In order to harmonize authorization with EMI-data, all components agreed to access the ARGUS authorization at least for obtaining user blacklisting. Some SE's may choose to further integrate ARGUS wherever ARGUS provides additional services to the SE implementation. However, at the time being, it is not planned to replace all Policy Decision Points by ARGUS.

4.1.9 Consolidating data access client libraries

The ARC and gLite middleware are providing data access libraries individually, which have sufficient functionality in common to justify a merge of those libraries. EMI data will work on an agreement between ARC and gLite on such a merge and a subsequent implementation and migration

4.2. EVOLUTION

For each individual EMI data component, *harmonization* is certainly part of the evolutionary process as well. However, in some areas each component may evolve without the need to coordinate those activities with other components. Those areas have been identified to be monitoring, accounting, maintainability and usability.

4.2.1 Monitoring and accounting

In their current version, most EMI storage elements provide information, which is picked up by monitoring agents to be further processed in a site-specific way to allow service and performance monitoring. Some components e.g. DPM offer probes for particular, well established monitoring systems (Nagios). For other storage elements, probes are provided by EMI external distributors or user communities (e.g. OSG/VDT and the German Storage Support group for dCache). The goal within EMI is to consolidate those activities and offer mechanisms to better monitor user access and user accounting as well as to interface this information to pseudo standard monitoring tools like Nagios across all EMI storage elements. This will happen gradually. For user accounting purposes, a specification or mechanism has to be provided, which allows EMI and external storage elements to interface with the infrastructure group within EMI to make this information centrally available in compliance with country specific privacy laws.

While accounting is already partially established in the compute area, a proper definition of what accounting would mean for data is still missing. The goal in EMI, possibly in collaboration with standardization initiative (e.g. OGF), is to define an accounting record, commonly used by EMI data components and covering the different aspects (e.g. legal, financial).

4.2.2 Maintainability and Usability

In order to simplify user or automatic interaction with EMI data components, most of the services are planning to provide or improve Web interfaces or Command Line interfaces. We are planning to leave the evaluation of the necessary for the different components in this area to the individual project teams.

4.2.3 Evolution in wide area protocols

At the time being, gsiFtp is the only commonly accepted protocol for wide area transfers. With the experience of operating a world wide data Grid, flaws were found in the original gsiFtp specification, which led to an updated version (v2). Version 2 is currently being implemented in GLOBUS software and with that will become available for EMI Data components (DPM, StoRM) soon. dCache already provides this functionality for quite some time and a production version of FTS does as well.

The question of providing additional protocols for wide area transfers, e.g. http(s) or NFS4.1(pNFS) is interesting from the theoretical point of view; however it is not part of the work program of EMI.

4.3. DETAILED WORKPLAN

Based on the description of harmonization and evolution above, a detailed work plan has been prepared. For the first year the work has been broken down to months and for the remaining project to years. This paragraph lists the tasks and the timing in more detail.

EMI JRA 1 DATA Technical Workplan

			2010								2011					
			Mai	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr		
			1	2	3	4	5	6	7	8	9	10	11	12		
			Year 1												Year 2	Year 3
1.0	**SRM specification consolidation**															
1.1		Start dicussion within OGF														
1.2		Collect community feedback														
1.3		Initial document														
1.4		SRM integration into UNICORE														
1.5		Review and final document														
2.0	**Replacing httpg with https**															
2.1		Prototype one client one SE														
2.2		All EMI SE's and clients														
2.3		Migration scenario and design														
3.0	**Posix Access (NFS4.1)**															
3.1		Setup of performance test system														
3.2		Performance test results (CHEP)														
3.3		Production file:// StoRM and dCache														
3.4		file:// for all SE in production														
4.0	**https/WebDav**															
4.1		http(s)														
4.2		WebDav														
5.0	**Catalog synchronization**															
5.1		Agreed design														
5.2		Prototype (Prove of concept)														
5.3		LFC plus one SE													PM 18	
5.4		LFC plus all EMI SE"s														
6.0	**File catalogue access for UNICORE**															
6.1		Implementation of file catalogue integration														
6.2		Evaluation of file catalogue intergration													PM 15	
7.0	**GLUE 2.0**															
7.1		Common interpretation of GLUE 2.0														
7.2		Publishing 1.3 with 2.0														
7.3		Publishing 2.0														
8.0	**Data client library consolidation**															
8.1		Development														
8.2		Migration														
9.0	**Integration of ARGUS**															
9.1		Blacklist integration (at least one SE)														
9.2		Blacklist integration (all SE's)														
9.3		Integration of more functionality if required														
10.0	**Storage Accounting**															
10.1		Definition of storage accounting record														
10.2		Add support in SE/FTS														
11.0	**Monitoring**															
11.1		Definition of a common Monitoring Interface														
11.2		Implementing Sensors in SE's and FTS														
12.0	**Manageability**															
12.1		CLI and Web Interfaces														
13.0	**Evalution : Message Passing**															
13.1		Possible use cases														

4.3.1 SRM specification consolidation and integration into UNICORE

4.3.1.1 Start discussion with the OGF SRM community group [PM 3]

A discussion on how to progress with the idea of a remote storage resource manager protocol beyond SRM 2.2 has already been taken place at OGF in May in Chicago. The responsible group within OGF agreed with the goals of EMI to stabilize and consolidate the currently available specification based on feedback from the available implementations.

4.3.1.2 Collection of community feedback [PM 9]

As the SRM has been adopted by groups outside of the scope of EMI, EMI data will collect user feedback to find the flaws in the current specification and problems with the client and server implementations.

4.3.1.3 Initial document [PM 12]

A document to supplement the SRM v2.2 specification will be provided to OGF, clarifying all issues found so far. Moreover an “SRM 2.2 developers guide” is envisioned describing common use cases and known pitfalls.

4.3.1.4 SRM integration into UNICORE [PM 24]

Based on the previous work, UNICORE data will either integrate an SRM 2.2 client by implementing the specification or by adopting existing solutions. SRM plus 4.3.3 (POSIX data access) enables UNICORE to seamlessly interoperate with EMI storage elements.

4.3.1.5 Review and finalize document. [PM 24]

To goal is to have OGF accepting the EMI document as an input to the SRM community efforts and its specification process.

4.3.2 Replacing httpg with SSL/X509 (https) for the SRM

4.3.2.1 Creating one prototype SE and one client [PM 12]

One EMI storage element and one client implementation will be chosen to provide a prototype SRM using SSL/X509 instead of the Globus httpg security protocol. Other SRM software providers are contacted and a joint plan is made to migrate between PM 24 to PM 36.

4.3.2.2 All EMI Storage Elements plus EMI clients [PM 24]

Based on the experience with the above prototype, all EMI SE’s and clients should support https alternatively to httpg. In parallel, the development progress of non-EMI storage elements and clients is tracked.

4.3.2.3 Migration scenario design for infrastructures [PM 36]

The migration plan is implemented.

4.3.3 POSIX data access (native or NFS 4.1)

4.3.3.1 Setting up a system for performance and reliability testing [PM 5]

Reasonably large test Compute and Storage Elements are setup to evaluate reliability and performance of the native POSIX and NFS 4.1 implementations.

4.3.3.2 Report of first performance results at CHEP [PM 10]

Reports on the results of 4.3.3.1 will be reported at the CHEP’10 conference.

4.3.3.3 Production POSIX access for StoRM and dCache. [PM 12]

As StoRM is based on POSIX file systems (GFPS, Lustre), this milestone it fulfilled for StoRM with day zero. DCache will provide a production NFS 4.1 implementation and DPM is planning for a prototype with PM 12.

4.3.3.4 All EMI storage elements provide POSIX access [PM 24]

All EMI storage elements will provide genuine POSIX file system through mounted file systems.

4.3.4 Web access, http(s) and WebDav

4.3.4.1 Storage Elements provide access via http(s) for downloading data. [PM 12]

dCache and DPM will provide read access to data via http(s). StoRM will make this decision dependent on user requirements.

4.3.4.2 Storage Elements provide partial or full access to data via WebDav [PM 24]

dCache will provide protected WebDav access to data. StoRM and DPM will make this decision dependent on user requirements.

4.3.5 Catalogue synchronization

4.3.5.1 Agreed design [PM 8]

A decision on how to tackle the catalogue name-space synchronization problem will be made and agreed on between the EMI catalogue and Storage Element providers. The agreement should account for requirements by non EMI-catalogue providers, e.g. the LHC experiments.

4.3.5.2 Prototype as 'proof of concept' [PM 12]

A prototype of the above agreement will be implemented.

4.3.5.3 LFC plus one storage element provide full synchronization functionality [PM 18]

Dependent on the results of the above prototype, the LFC and at least one SE will provide the synchronization feature in production.

4.3.5.4 All EMI SE's provide synchronization functionality [PM 24]

The name space of the LFC and all EMI SE's will be able to stay synchronized.

4.3.6 File catalogue access for UNICORE

4.3.6.1 Implementation of file catalogue integration [PM 12]

A prototype of the file catalogue access will be implemented.

4.3.6.2 Evaluation of file catalogue integration [PM 15]

The prototype will be evaluated and tested and the final version will be published.

4.3.7 GLUE 2.0

4.3.7.1 Common agreement on the interpretation of the GLUE 2.0 schema. [PM 6]

Although GLUE 2.0 is already well defined, all partners, which will have to implement the standard, will have to agree on a common interpretation of the specification, and possibly extend it aligning with the OGF PGI group and their extensions if necessary.

4.3.7.2 Publishing GLUE 1.3 data as with GLUE 2.0 schema [PM 12]

The currently available information in the various components will be published, using the GLUE 2.0 schema in addition to GLUE 1.3. Addition information required by the GLUE 2.0 specification might still be missing. Clients are required to cope with the partially available information.

4.3.7.3 EMI data components fully GLUE 2.0 compatible [PM 24]

All required information is published using GLUE 2.0.

4.3.8 Data client library consolidation

4.3.8.1 Development of a consolidation between the ARC and gLite data clients. [PM 24]

A design will be presented on how to merge the ARC and gLite data access libraries and a plan on how to migrate.

4.3.8.2 Migration of the new data client component. [PM 36]

Full migration to the merged data access libraries will be finalized.

4.3.9 Integration of ARGUS

4.3.9.1 ARGUS Blacklist integration of at least one EMI Storage Element. [PM 12]

At least one EMI storage element will have integrated the EMI ARGUS blacklisting mechanism.

4.3.9.2 ARGUS Blacklist integration of all EMI Storage Elements. [PM 24]

All EMI storage elements will have the EMI ARGUS blacklisting mechanism integrated.

4.3.9.3 Integration of further ARGUS functionality into EMI SE's, if required. [PM 36]

Each storage element implementation may integrate more features, the ARGUS software is offering. Replacing internal file system or catalogue authentication with the ARGUS authentication is not envisioned.

4.3.10 Storage Accounting

4.3.10.1 Definition of a storage accounting record. [PM 8]

A storage accounting record is defined or a standardized schema like the OGF URs might be extended), reflecting practical, financial and legal requirements of storage location, usage and space and data flow.

4.3.10.2 Add support of storage accounting in FTS and EMI storage elements. [PM 36]

FTS and the EMI storage elements will provide information according to the agreed record in 4.3.10.1.

4.3.11 Monitoring

4.3.11.1 Collaborating with the EMI infrastructure group to define an interface for EMI data monitoring. [PM 12]

EMI data will support the EMI infrastructure group to define an interface for commonly monitor availability, problems and performance of EMI data components.

4.3.11.2 Implementing server side sensors accordingly. [PM 24]

The agreed interface will be implemented in EMI storage elements as well as in FTS and the LFC.

4.3.12 Manageability

4.3.12.1 CLI and Web Interfaces, where required. [PM 24]

Up to now, the development of most of the EMI data components has been focused on reliability and performance. Where necessary, the maintainers of the components, in collaboration with their users will evaluate the necessity of improved command line interfaces (CLIs) or Web based interfaces. The progress is tracked by EMI data but not enforced.

4.3.13 Evaluation: Message passing in EMI-data

4.3.13.1 Collecting possible use cases for message passing in EMI data. [PM 7]

EMI infrastructure will provide a common mechanism for message passing for all EMI components. EMI-data will provide input to the infrastructure group to help selecting an appropriate mechanism and convenient interfaces.

5. CONCLUSIONS

This report provides a solid data area work plan for the first year of the EMI project. The activities of the Product Teams involved in the data technical area are generally categorised into two complementary areas: harmonisation and evolution. Data components from ARC, gLite, UNICORE and dCache will work together 'in harmony' to create 'single' larger functional unit. With the advanced data functionalities of the involved middleware stacks, the resulting EMI data compound component has the potential to replace expensive commercial products. Orthogonal to this work is evolving and improving the different data components to address the data requirements from EMI clients. These requirements involve further developments in the areas of monitoring, accounting, maintainability and sustainability.

This report, even with its advanced and mature content, is not yet final. It is finalised now for submission to the EC before the end of 2010. A revision is expected in the first months of 2011 to cover recommendations from SA1 and SA2 to further improve this document.