EUROPEAN MIDDLEWARE INITIATIVE

DJRA1.4.3 - INFRASTRUCTURE AREA WORK PLAN AND STATUS REPORT

EU DELIVERABLE: D5.4.3

Document identifier:	EMI-DJRA1.4.3-1277585-Infrastructure_Area_Work_Plan_M24-v1.0.doc
Date:	30/04/2012
Activity:	JRA1
Lead Partner:	CERN
Document status:	Final
Document link:	https://cdsweb.cern.ch/record/1277585

Abstract:
This deliverable contains the detailed work plan of the Infrastructure Services technical area compliant with the overall EMI Technical Development Plan. The plan is released early in the project life and updated every year including a status report on the achievements of the past 12 months compared to the planned objectives.

I. DELIVERY SLIP

	Name	Partner/Activity	Date
From	Laurence Field	CERN/JRA1	19/04/2012
Reviewed by	Riccardo Zappi, Jedrzej Rybicki	INFN/JRA1 JUELICH/NA3	23/05/2012
Approved by	PEB	-	27/05/2012

II. DOCUMENT LOG

Issue	Date	Comment	Author/Partner
1	19/04/2012	Initial Draft	Laurence Field/CERN
2	16/05/2012	v0.3 for review	Laurence Field/CERN
3	27/05/2012	v1.0 PEB approved version	PEB

III. DOCUMENT CHANGE RECORD

Issue	Item	Reason for Change
1		
2		

IV. 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-organization not affecting the content and meaning of the document can be applied by the authors without peer review. Other changes must be submitted for 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.

V. GLOSSARY

Acronym	Long Name
APEL	Accounting Processor for Event Logs
API	Application Programming Interface
ARC	Asynchronous Resource Connector
CAR	Computing Accounting Record
CLI	Command Line Interface
DCI	Distributed Computing Infrastructure
DGAS	Distributed Grid Accounting System
EGI	European Grid Initiative
EMI	European Middleware Initiative
ERIS	EMI Resource Information Service
GLUE	Grid Laboratory Uniform Environment

LDAP	Lightweight Directory Access Protocol
LDIF	LDAP Data Interchange Format
NGI	National Grid Initiative
PEB	Project Executive Board
PTB	Project Technical Board
SAGA	Simple API for Grid Application
STAR	Storage Accounting Record
VO	Virtual Organization

The complete EMI glossary is available at https://twiki.cern.ch/twiki/bin/view/EMI/EmiGlossary .

VI. COPYRIGHT NOTICE

TABLE OF CONTENTS

1. INTRODUCTION

1.1. EXECUTIVE SUMMARY

This work plan document gives the status of the Infrastructure Area work after the second year of the EMI project and concrete work plans for the third year. These plans exclude day-to-day maintenance and focus on major changes that are required to meet the strategic goals of the EMI project. Section 2 provides a report on progress within the second year and the work plan for year three is given in Section 3.

The status report provides an update on the work done to achieve the objectives as described in DNA1.3.2 Technical Development Plan [R1] and DJRA1.4.2 Infrastructure Area Work Plan and Status Report [R2]. These status reports are organized by objectives, as given in the updated Technical Development Plan DNA1.3.3 [R3], and the status at the time of writing is given.

All of the objectives planned for the second year were met. There were a number of major highlights that are worth mentioning here.

A design document for the EMI Registry was produced [R4] and has been endorsed by the EMI Project Technical Board. The EMI Registry, which enables service discovery, represents a major new product for EMI and as such good progress with this is encouraging. The Service Discovery use case has been identified by the general user community as an important aspect of the Grid information system for which the existing solutions are falling short. A specific system for Service Discovery, the EMI Registry, aims to address the limitations.

Information system infrastructure components all now support the GLUE 2.0[R17] data model. GLUE 2.0 support is an important objective for EMI and the support GLUE 2.0 in such core components is critical for success. As of May 2012 52.09% (199/382) of Grid sites on the EGI have updated to the EMI version of the site-level BDII and are publishing GLUE 2.0 information. GLUE 2.0 publication can be used as metric to measure the effective rollout of EMI products.

Messaging is another major new area for EMI and a great deal of work has been done in this field. Messaging guidelines are now available for developers and a number of investigations have been carried out that evaluate the potential benefits of adopting a messaging system within EMI.

The introduction of the EMI Resource Information Services (ERIS) into the EMI 2.0 release as a common component for obtaining information directly from services is an important step for harmonization. It consolidates the existing resource-level products and ensure interoperability through the agreement on a common information model and interface.

With the majority of objectives achieved during the second year, only a few objectives remain for the final year of the project.

1.2. PURPOSE AND SCOPE

This document reports on the progress with respective to the objectives for Infrastructure Area over the second year and presents the work plan for year three. The work plan is an extended technical description of the high-level objectives that are defined in the overall EMI Technical Development Plan.

1.3. DOCUMENT ORGANIZATION

An executive summary of the document can be found in Section 1. Section 2 reports on the progress with respect to the objectives for Infrastructure Area over the second year and the work plan, covering year three, is presented in Section 3.

2. INFRASTRUCTURE AREA STATUS REPORT

This section reports on the progress during the second year with respect to the technical objectives as defined in the Technical Development Plan (DNA1.3.2) [R1] for the Infrastructure Area. Each subsection describes the status with respect to a specific objective. A few comments on the general progress in the Infrastructure Area are provided first.

A number of developments during the past year have had an impact on some of the objectives in the Infrastructure Area. Two points made at the 1st project review affected the cloud and virtualization aspects. The reviewers recommend the appointment of a Strategic Director to the project so they can build bridges between the various user groups as well as deploying a more proactive partnering approach with the other projects in the DCI ecosystem. The result has been a new networking activity, NA3 - Sustainability and Long-Term Strategies, which aims to establish a clear EMI market positioning with respect to cloud computing and virtualization. It will consider the security aspects and show how grid technologies from EMI can be used to address these issues.

The impact on the Infrastructure Area is that cloud and virtualization is seen more of part of the strategic vision for EMI rather than a concrete work item. During the ECB session at the All Hands Meeting in Padova [R16], it was agreed that there is a lack of clarity in the cloud and virtualization plans as consequence of a lack of clarity in the requirements provided by EGI. It was proposed that the current EMI cloud activities are put on hold while NA3 re-assesses the requirements.

The result of consultations with EGI has re-focused the objectives with respect to Service Monitoring and Management. In the original DoW [R5], it was envisioned that EMI would investigate interfaces in this area, along with the use of messaging technology, and propose a common interface for Service Monitoring and Management that could be adopted by all EMI services. The feedback from EGI and system administrators is that while there is general agreement that such a interface would be on everyone's wish list, in practice smaller more concrete objectives would have greater impact.

A survey of Grid sites [R6], which was jointly conducted with EGI, showed that there is no common interface used for service monitoring and management in existing data centers. The provision of another proprietary interface would only serve to increase the entropy in this area. Such an interface should be defined in within the wider context of standards in data centers. This is something that is out-of-scope for EMI. Another contributing factor is the re-tooling of data centers in order to adopt the virtualization paradigm and the advent of cloud computing. Due to these factors, EMI should instead watch this area closely for any emerging standards for Service Monitoring and Management. Such an activity should also fall under the responsibility of NA3.

The feedback from the survey suggested that there were a number of concrete steps that EMI could take to improve Service Monitoring and Management. As a number of different fabric management tools are used to provision EMI services, one recommendation is that EMI software should conform to the operating system guidelines for the platforms which EMI supports. This objective is already in the EMI Technical Plan as cross-area objective 6, *adhere to operating system standards for service operation and control regarding configuration, log and temporary file location and service start/status/stop (DNA1.3.3 ref X6.).*

For Service Monitoring, EGI requested that each service should provide a Nagios probe which can be used to measure the availability. Again, this objective is already in the EMI Technical Plan as cross-area objective 4, *provide and support monitoring probes for EMI services (e.g. Nagios) (DNA1.3.3 ref X4.).*.

An outcome of the EMI Vision Meeting in December 2011 was removal of the DGAS client from the EMI stack as the DGAS client is only used by two NGI. APEL client is now the sole accounting client in the EMI stack. It will be ensured that DGAS accounting server (not an EMI product) will be able to accept information from the EMI Accounting Client via interoperability testing. This decision simplifies the harmonization in this area.

2.1. MESSAGING INVESTIGATION; ACCOUNTING

An investigation into the use of messaging for accounting has been done by the APEL product team. A full report on the use of messaging for accounting [R7] has been provided. The use of messaging considerably simplifies the accounting system. As a result, work to implement messaging in accounting will continue. A specific objective, *implement accounting record publishing via messaging*, has already been defined and the status is reported below.

2.2. MESSAGING INVESTIGATION; SERVICE MONITORING AND MANAGEMENT

The result of consultations with EGI has effectively closed this objective. In the original DoW, it was envisioned that EMI would investigate the use of messaging technology for Service Monitoring and Management. A survey of Grid sites [6] was jointly conducted with EGI to investigate the requirements in this area. The feedback from system administrators is that while there is general agreement that such a solution is on everyone's wish-list, this would only make sense within the wider context of standards in data centers, which is out-of-scope for EMI. In practice smaller more concrete objectives would have greater impact for Service Monitoring and Management.

The result of this initial investigation is that this effort should be directed into other areas where EMI can have a greater impact. The Messaging Product Team already has experience related to Service Monitoring via Messaging and a specific objective to provide service monitoring via message could be defined if there is a specific requirement.

2.3. MESSAGING GUIDELINES

The initial version of the messaging guidelines [R10] was disseminated internally in February 2011 and externally at the EGI User Forum in April 2011. Feedback provided has been incorporated into the document and it was considered final in March 2012. Although the document is considered final, it will be updated when necessary, to address new developments in the area, and versioned accordingly.

2.4. IMPLEMENT ACCOUNTING RECORD PUBLISHING VIA MESSAGING

Both the Compute Accounting Record (CAR)[R13] and Storage Accounting Record (StAR)[R14] have been defined and agreed within EMI. The central accounting repository used by EGI was updated in March 2012 to accept new CAR messages from clients and summaries from other systems. The existing EMI Accounting Client has been updated to send these new CARs using messaging and it is expected that this will be released as an updated to the EMI 2 release on a time scale that is acceptable for EGI.

2.5. IMPLEMENT THE COMMON EMI REGISTRY

Following on from the design document [R4], a development plan [R8] was defined with the objective of providing an implementation of the EMI Registry in time for the EMI 2 release. A prototype of the EMI Registry was available for the 3rd EMI All Hands Meeting and is available in the EMI 2 release. The implementation should be considered experimental. The aim of the release is to gain initial experience for users in order to provide a production quality release in time for EMI 3.

2.6. GLUE 2 INFORMATION MODEL SUPPORT

All components from the Infrastructure Area that were required to support GLUE 2.0 in time for the EMI 2 release were identified and included in the EMI development tracker [R9]. All these components have been updated to support GLUE 2.0 and are available in the EMI 2 release.

2.7. EMI RESOURCE INFORMATION SERVICE

The EMI Resource Information Service (ERIS) Task Force was created to provide a technical proposal on resource-level information service. The task force considered the options and on balance recommended that the ERIS should provide an LDAPv3 interface to GLUE 2.0 information. Information providers, in the classic sense, will extract information from the under laying Grid service and produce GLUE 2.0 information in the LDIF format. These two together represent the external and internal interfaces for the ERIS. It is envisioned that this approach will be a minimal-cost solution for EMI and will have a low-impact on existing infrastructures. The proposal was endorsed by the Project Technical Board in September 2011 and an ERIS, based on minor modifications to existing EMI software, was release in EMI 2.

2.8. INFRASTRUCTURE AREA SIMPLIFICATION PLAN

An Infrastructure Area simplification plan has been defined [R11] and was presented at the 3rd EMI All Hands Meeting. The plan covers two topics; accounting and information systems.

Within EMI simplification of accounting will be achieved in three phases

1. Agreement on the information models.
2. Publication of records using EMI Messaging
3. Harmonization of the publisher

Phases one and two are already defined in the EMI Infrastructure Area work plan and have been achieved. For phase three, at EMI Vision Meeting in December 2011 a decision was made to remove the DGAS client from EMI stack leaving the APEL client as the sole accounting client within EMI. This Accounting Client will publish the agreed accounting records via the messaging system. This is the accounting interface which EMI will support and has been agreed with EGI. Any further feature requests or requirements will be considered for future releases of the Accounting Client. It must be ensured that the protocol used by the Accounting Client is interoperable with existing Accounting Servers.

For the information system, the adoption of GLUE 2.0, the development of the EMI registry and the agreement for the EMI Resource Information Service are already defined in the EMI Infrastructure Area work plan. The decision on harmonization of information system clients was delayed until after the EMI 1 release so that user feedback can be sort on the products. At the information system workshop hosted by EGI on 1st December 2011 in Amsterdam, the usage of the information system clients was presented [R12]. From an operations perspective, the lcg-info clients were reported to be widely used by NGI staff, user support and operations and well as native LDAP queries. ARC sites reported considerable usage of direct LDAP queries while ARC lib was not reported as used. The UNICORE CLI is heavily used by sites where UNICORE is deployed. No NGIs reported the use of the SAGA Service Discovery API.

A user survey directed at VO managers and members reported that almost all the users who responded to the survey ranked lcg-info as the most important information system client. The ARC lib and the UNICORE CLI were considered important by only 4% of the users and the SAGA Service Discovery API/CLI was considered important by 11% of the users. In terms of uses cases, 69% used the clients to discover services provided by a specific middleware stack, 23% used the clients to discover services across different middleware and 80% of users discover services supporting a specific VO.

One outcome of this workshop which is worth mentioning is the importance of GLUE 2.0 with respect to service integration and interoperability from an operations perspective.

2.9. IMPLEMENTATION OF THE INFRASTRUCTURE AREA SIMPLIFICATION PLAN

The objectives related to accounting have all been achieved. Both the Compute Accounting Records and Storage Accounting Records have been agreed within EMI. The work to publish these records via messaging has been done and will be release sometime during the third year as an update to EMI 2.

For information system, the adoption of GLUE 2.0, the development of the EMI registry and the agreement for the EMI Resource Information Service were already defined within the EMI Infrastructure Area work plan and been achieved. The issue of multiple information system clients remains but is not considered a priority due to the low maintenance overhead of these components.

2.10. SERVICE MANAGEMENT

The result of consultations with EGI has effectively closed this objective. In the original DoW, it was envisioned that EMI would investigate the use of messaging technology for Service Management. A survey of Grid sites [R6] was jointly conducted with EGI to investigate the requirements in this area. The feedback from system administrators is that while there is general agreement that such a solution is on everyone's wish list, this would only make sense within the wider context of standards in data centers, which is out-of-scope for EMI. In practice smaller more concrete objectives would have greater impact for Service Management.

The result of this initial investigation is that effort should be re-directed to other areas where EMI can have a greater impact. Cross-area objective 6 - *adhere to operating system standards for service operation and control regarding configuration, log and temporary file location and service start/status/stop, will be more beneficial (DNA1.3.3 ref X6.).*

2.11. MONITORING INTERFACES

The result of consultations with EGI has effectively closed this objective. In the original DoW, it was envisioned that EMI would investigate the use of messaging technology for Service Monitoring. A survey of Grid sites [R6] was jointly conducted with EGI to investigate the requirements in this area. The feedback from system administrators is that while there is general agreement that such a solution is on everyone's wish list, this would only make sense within the wider context of standards in data centers, which is out-of-scope for EMI. In practice smaller more concrete objectives would have greater impact for Service Monitoring.

The result of this initial investigation is that effort should be re-directed to other areas where EMI can have a greater impact.

2.12. EMI CLOUD STRATEGY AND ARCHITECTURE

Although work has been done by the Cloud and Virtualization Task force to define a cloud strategy and architecture, the result has been inconclusive. This objective has been moved to the new NA3 and put on hold while the requirements are reassessed.

Specific objectives for EMI Product Teams to implemented StAR have been added to the Data Area work plan. A workshop on this topic took place at the EGI Community Forum in March 2012. StAR publication is expected to be available in time for the EMI 3 release.

2.13. CLIENT TOOL USABILITY

No specific comments have been raised on the usability of clients from the Infrastructure Area. As such there were no specific issues to address.

2.14. OPERATING SYSTEM STANDARDS

This cross-area object has been renamed standard service operation and control. From the perspective of the Infrastructure Area, this objective also partially addresses the service management objective. An inventory of all EMI services was made and each was evaluated for conformance with standard service operation and control. All infrastructure area services released in EMI 2 conform to this standard, the Fedora Packaging Guidelines.

2.15. PLATFORM SUPPORT

The target platforms for EMI 2 were defined by the project to be Scientific Linux 5 & 6 and Debian 6 all on x86_64 architectures. All the infrastructure area service are available on Scientific Linux 6 with some core components also being available on Debian. Infrastructure products will be available on Debian as an EMI update when ready.

2.16. MONITORING PROBES

Nagios probes are available in EMI 2 for the infrastructure area services (ARC ARIS and EGIIS, BDII, and UNICORE Registry).

3. INFRASTRUCTURE AREA WORK PLAN

This section gives the work plans of the Infrastructure Area in order to attempt to achieve the technical objectives from DNA1.3.3 Technical Development Plan. The subsections below follow the technical objectives, giving a description of the objective, the benefits to EMI, sub tasks and risks. As the majority of the objectives have already been achieved during the second year, may of the subsequent objectives finalize on-going activities.

3.1. IMPLEMENT ACCOUNTING RECORD PUBLISHING VIA MESSAGING

This technical objective (DNA1.3.2 ref: I9) is provide or adapt the accounting publishers for compute and Data Area services to use the common messaging system. As the majority of the work has already been done, it just remains to integrate the result into the EMI release.

3.1.1 Subtasks

ID	Description	Responsible	Target Date
1	Integrate into the EMI release	APEL PT Leader	M28

3.1.2 Risks

There is minimal risk for this objective as the work has mainly been done.

3.1.3 JRA1 key performance indicators addressed by the objective

This may improve KJRA 1.1, KJRA1.2, KJRA1.3 and KJRA1.4.

3.2. IMPLEMENTATION OF THE INFRASTRUCTURE AREA SIMPLIFICATION PLAN

This technical objective (DNA1.3.3 ref: I10) is to implement the plan for simplification and reduction in the number of Infrastructure Area CLIs, libraries, internal components and services.

3.2.1 Subtasks

ID	Description	Responsible	Target Date
1	Oversee the adoption throughout EMI	Infrastructure Area Leader	M32

3.2.2 Risks

The main risk associated with this objective is the timely delivery of adoption by other product team in EMI.

3.2.3 JRA1 key performance indicators addressed by the objective

This may improve KJRA 1.1, KJRA1.2, KJRA1.3 and KJRA1.4.

3.3. PLATFORM SUPPORT

This technical objective (DNA1.3.2 ref: X7) is to port, release and support EMI components on identified platforms (full distribution on SL6 and Debian 6, UI on SL5/32 and latest Ubuntu).

3.3.1 Subtasks

ID	Description	Responsible	Target Date

1	Identify which components must be available on which platforms	Infrastructure Area Leader	M26
2	Ensure that each component is available on that platform for the release date.	JRA1 Leader	M32

3.3.2 Risks

The risks associated with this objective are difficult to ascertain. The problems to be overcome will only be known once an attempt is made to port the software to a specific platform.

3.3.3 JRA1 key performance indicators addressed by the objective

This does not address any KPIs.

3.4. EMIR ROLLOUT: SERVICE AND CONSUMER-SIDE INTEGRATION

This technical objective (DNA1.3.2 ref: X18) is ensure that all EMI services publish their existence to the EMI Registry. The pub

3.4.1 Subtasks

ID	Description	Responsible	Target Date
1	Provide a publication method	Infrastructure Area Leader	M28
2	Ensure that each EMI service publishes to EMIR.	JRA1 Leader	M30

3.4.2 Risks

The main risk associated with this objective is the timely delivery of the publication method. It must be ensured that it arrives with enough time for product teams to integrated and test this method in advance of the EMI 3 release.

3.4.3 JRA1 key performance indicators addressed by the objective

This does not address any KPIs.

3.5. CLEAR IDENTIFICATION OF THE MANDATORY CONFIGURATION VARIABLES

This technical objective (DNA1.3.2 ref: X19) is to identify the mandatory configuration variables.

3.5.1 Subtasks

ID	Description	Responsible	Target Date
1	Identify the mandatory configuration variables	Infrastructure Area Leader	M28

3.5.2 Risks

The risks associated with this objective are negligible.

3.5.3 JRA1 key performance indicators addressed by the objective

This does not address any KPIs.

4. CONCLUSIONS

This document has described the status of the work performed by the EMI Infrastructure Area in the second year of the project according to the objectives as outlined in DJRA1.4.2. All of the second year objectives were met. The concrete work plan for the third year was provided, based on objectives defined in DNA1.3.3.

5. REFERENCES

R1	DNA1.3.2 - Technical Development Plan http://cdsweb.cern.ch/record/1277543
R2	DJRA1.4.2 Infrastructure Area Work Plan and Status Report http://cdsweb.cern.ch/record/1277583
R3	DNA1.3.3 - Technical Development Plan http://cdsweb.cern.ch/record/1277544?
R4	EMI Registry Design http://twiki.cern.ch/twiki/pub/EMI/EMIRegistry/EMIRegistryDesign-v0.4.doc
R5	EMI Description of Work (Public DoW) http://twiki.cern.ch/twiki/pub/EMI/EmiDocuments/EMI-Part_B_20100624-PUBLIC.pdf
R6	Survey of Grid sites http://twiki.cern.ch/twiki/pub/EMI/EmiJra1T5TaskForceServiceManagement/TF_service_management.doc
R7	Accounting Report http://twiki.cern.ch/twiki/pub/EMI/APELClient/APEL-Messaging-v2.2.pdf
R8	EMI Registry Development Plan http://twiki.cern.ch/twiki/pub/EMI/EMIRegistry/Registry-Dev-Plan-v01.pdf
R9	EMI Development Tracker http://savannah.cern.ch/projects/emi-dev
R10	EMI Messaging Guidelines http://twiki.cern.ch/twiki/bin/view/EMI/EMIMessagingGuidelines
R11	Infrastructure Area Simplification Plan http://twiki.cern.ch/twiki/bin/view/EMI/SimplificationPlan
R12	Usage of the information system clients https://www.egi.eu/indico/conferenceTimeTable.py?confId=654
R13	Computing Accounting Record https://twiki.cern.ch/twiki/pub/EMI/ComputeAccounting/CAR-EMI-tech-doc-1.0.doc
R14	Storage Accounting Record https://twiki.cern.ch/twiki/pub/EMI/StorageAccounting/StAR-EMI-tech-doc-v7.doc
R15	Fedora Packaging Guidelines http://fedoraproject.org/wiki/Packaging:Guidelines
R16	3rd EMI All Hands Meeting https://indico.cern.ch/conferenceDisplay.py?confId=147484
R17	GLUE 2.0 Information Model http://www.ogf.org/documents/GFD.147.pdf