EUROPEAN MIDDLEWARE INITIATIVE

EMI STRATEGY 0.10

EMI DOCUMENT

Document identifier:	**EMI-DOC-WP6-EMI-Strategy-v0-10.doc**
Date:	**20/05/2013**
Activity:	**NA3**
Lead Partner:	**JUELICH**
Document status:	**FINAL**
Document link:	

Abstract:
The EMI project is a consortium of a wide variety of institutions that deliver software products for distributed systems. These products are used in e-Infrastructures, international research projects, and in commercial environments. This document outlines a strategic path of the EMI project in general and its consortium members in particular focussing on two simple messages: Firstly what are strategic recommendations for the future of middleware in Europe. Secondly, how these recommendations can be turned into concrete actions for future middleware collaboration initiatives.

Document Log

Issue	Date	Comment	Author / Partner
0.01	10/09/2012	TOC after NA3 Telcon on 2012-09-05	Morris Riedel / JUELICH
0.02	22/10/2012	TOC After NA3 Call for Comments	Morris Riedel / JUELICH, Oxana Smirnova / LUND
0.03	28/11/2012	Assignments of work items and debrief of CB of AHM meeting comments	Morris Riedel / JUELICH, Peter Stefan / NIIF
0.10	20/05/2013	Integration of all different NA3 discussed parts and material as well as PEB comments (simple messages, smaller report, ~ 10 pages)	Morris Riedel / JUELICH

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

The EMI project is a consortium of a wide variety of institutions that deliver software products for distributed systems. These products are used in e-Infrastructures, international research projects, and in commercial environments.

This document outlines a strategic path of the EMI project in general and its consortium members in particular by focussing on the following two simple messages:

(1) What EU Middleware Needs to Provide To Stay Competitive

The following recommendations are given to middleware providers within the EMI project in order to stay competitive to commercial solutions or community-based technologies:

Recommendation 1: Fostering European Middleware Technologies Leadership

Recommendation 2: Enable Lightweight Access to Federated Technologies

Recommendation 3: Position Middleware in the Context of Big Data and Clouds

Recommendation 4: Perform Application Enabling During Middleware Development

Recommendation 5: Provide Reusable Products with Open Standard Interfaces

(2) HOW EU Middleware Engages In Suggested Recommendations

The aforementioned recommendations are addressed by describing the following concrete action items and how middleware providers within the EMI project plan to engage in them:

Action 1: Form Open Middleware Collaboration Group with Complementary Strength

Action 3: Implement Federated Identity Concepts and Technologies

Action 4: Develop collaboratively Web-based Solutions with User Communities

Action 5: Augment Middleware Products with Big Data Processing Techniques

Action 6: Enhance Middleware with Data Curation and Management Functionality

Action 7: Observe and Adapt to Community Deployments around Emerging Clouds

Action 8: Adopt Application Case Study and Use Case Oriented Developments

Action 9: Form Development Teams that Focus around One Specific Problem

Action 10: Drive and Adopt Open Standard Interface Specifications

This document identified five distinct high-level recommendations that are crucial to take into account by middleware providers in Europe towards 2020. There are already many plans and approaches driven by the EMI project to implement these high-level recommendations leading to the suggested ten actions alongside their initial plans for implementation by middleware providers in Europe.

To remain Europeans leadership in Europe, the community around middleware must act, work, and establish plans together facing new challenges in the so-called 'big data waves' and required integration use cases with a growing number of scientific community-based solutions.

The EMI project started several concrete actions to implement the high-level recommendations of this report, but seeks to have fruitful collaboration with other middleware experts in the field. Hence, this report is only the start of a strategic plan that needs to be carried forward by firstly getting updates by scientific community representatives in the next years and secondly by new partners that engage with members of the EMI consortium into the middleware challenges towards 2020.

2. WHAT EU MIDDLEWARE NEEDS TO PROVIDE TO STAY COMPETITIVE

The European middleware technologies are successfully used worldwide in a wide variety of scientific communities and research disciplines. The EMI project was instrumental in forming a consortium that brought together the key experts in the field creating a collaboration that is capable of keeping Europeans leadership in the field of middleware development, maintenance, and deployment.

We foresee in the future that a lightweight collaboration is still needed to continue the success of the EMI project among its consortium members. This includes the creation of a collaboration model that enables the geographically dispersed middleware development teams to act and implement requirements under a slightly coordinated umbrella. In order to stay competitive such a collaboration model need to establish a way to align with 'open source community' products, trends, and procedures. We summarize this recommendation as follows:

Recommendation 1: Fostering European Middleware Technologies Leadership

The access to middleware products with certificates offers strong security, but we have also experienced that it represents a hindrance for new communities to take advantage of middleware technologies. This is in particular the case when lightweight access is required such as with scientific gateways or when new security paradigms gaining momentum in the community such as federated identity. We further expect in the future that the federation of existing community solutions with middleware products becomes increasingly imported that means that middleware providers need to offer hooks (e.g. plug-in mechanisms) and develop middleware in the most flexible ways. We summarize this recommendation as follows:

Recommendation 2: Enable Lightweight Access to Federated Technologies

Observing trends in the community is a very important aspect of strategic plans and we have identified by participation in many workshops that the aspect of 'big data' will be a key driver of projects and initiatives in the next decade. The term 'big data' is broadly defined and middleware providers in Europe that have already strong data management components need to position their middleware in the market of data components that not only emerge as community-specific solutions, but also become available from more commercial domains. We summarize as follows:

Recommendation 3: Position Middleware in the Context of Big Data and Clouds

Developing middleware products <u>for</u> the scientific communities has been the major activities in the past while in the future we foresee another growing market of <u>developing jointly solutions</u> together with the scientific communities. That means that strong middleware backend components will be part of (partly refactored) community solutions and this can only be done by applying the process concept of 'application enabling'. This concept stands for the fact of really having close relationships with user communities and to discuss iteratively how middleware developments should look like in the past. This process thus emphasizes on collaboration with multi-disciplinary user communities, but most notably the scientific applications that should be used in conjunction with middleware products. We summarize this important recommendation as follows:

Recommendation 4: Perform Application Enabling During Middleware Development

The EMI project changed the way how middleware development, maintenance, and deployment was done before. The change away from atomic middleware units (i.e. ARC, gLite, UNICORE, dCache) towards a joint release model (EMI 1,2,3, etc.) based on small focussed EMI products lead to many advantages. Most notably, the marketing, description, release, and availability of these EMI products encouraged the re-usability of EMI products (e.g. in dCore commercial products). At the same time, EMI significantly increased the adoption of open standards during its development phases meaning that EMI products can be more seamlessly re-used by other standard-compliant technologies. We therefore recommend that this way of providing middleware in Europe continues that we summarize as follows:

Recommendation 5: Provide Reusable Products with Open Standard Interfaces

3. HOW EU MIDDLEWARE ENGAGES IN SUGGESTED RECOMMENDATIONS

Before the EMI project, separate Grid middleware projects developed similar-but-different solutions to the same problems, notably in the security area. In order to focus on a wide spectrum of complementary solutions, the achievements of EMI was to factor out some of these and replace them with common components, striving for backwards compatibility. This not only improved interoperability, but also made the software development and maintenance processes more efficient, freeing up resources for implementing valuable functions.

In order to avoid again duplication of effort by continue the collaboration in a more open manner meaning that also other middleware-related groups can join. In this sense, the "social capital" in the forms of trust and the capability to collaborate across boundaries towards common goals need to remain. The middleware providers in Europe will thus form a European Middleware Collaboration Group (EMCG) that is open for non EMI middleware specialists, geographically dispersed, and only slightly coordinated. The group will focus on middleware itself in general and the important interactions (e.g. integration tests) among them in particular. Also, this report as strategic planning should be continuously updated within this group as a 'living document' These activities are summarized as follows:

> **Action 1: Form Open Middleware Collaboration Group with Complementary Strength**

Action 1 addresses the following recommendations:

- *Recommendation 1: Fostering European Middleware Technologies Leadership*

The middleware products in Europe are competitive when it comes to requirements that originated from the wide variety of research disciplines but need to keep the pace in the future to remain its European leadership. As more and more open source community technologies are available, the middleware providers in Europe will align with this model and establish a forum that is currently named as 'ScienceSoft'. This forum will provide a market place for middleware products and beyond thus having a seamless collaboration model with solutions of other middleware providers or scientific communities. This will provide a significantly important foundation for several future projects, initiatives, and activities that may only involve a fraction of partners in ScienceSoft, but are particularly focussed to solve a specific challenge in research infrastructures. At the same time, the forum will be instrumental to keep contacts and to align development procedures more with the general open source community movements while increasing the visibility of not only the products itself, but also the developers that are responsible for the product developments. We summarize these concrete activities as follows:

> **Action 2: Establish Forum to Align with Open Source Community Practices**

Action 2 addresses the following recommendations:

- *Recommendation 1: Fostering European Middleware Technologies Leadership*

In order to provide scientific end-user communities with solutions that can be seamlessly integrated in their own solutions, the EMI middleware products have to implement federated concepts. This stands mostly for simple yet flexible security, but also for open interfaces that enable the federation of community solutions with strong EMI products. These activities are summarized as follows:

> **Action 3: Implement Federated Identity Concepts and Technologies**

Action 3 addresses the following recommendations:

- *Recommendation 2: Enable Lightweight Access to Federated Technologies*

Lightweight Web-based access to large powerful servers with deployed middleware is considered to become the most used way of how scientific communities will exploit middleware technologies. Middleware providers need to implement thus together with experts in the field simple Web-based solutions such as Scientific Gateways, but also more Apps and Gadgets tuned for mobile devices. The latter becomes significantly interesting when 'citizen scientists' will be entering the scientific domain (at least for providing observation data). There are other players in Europe that focus on client-side and Web-based access and EMI consortium members will collaborate with those as part of the European Middleware Collaboration Group. We summarize as follows:

Action 4: Develop collaboratively Web-based Solutions with User Communities

---

Action 4 addresses the following recommendations:

- *Recommendation 2: Enable Lightweight Access to Federated Technologies*

The handling of 'big data' will become a key pillar of activities of the European Middleware Collaboration Group. These activities are basically twofold reflecting on different but closely related needs. Firstly, existing processing components of the middleware products need to be augmented with de-facto standard 'big data' processing techniques. Existing strong processing solutions of middleware based on batch-processing and Grid distribution need to support additional modern distributed processing approaches such as Map-Reduce or optimized versions of it such as iterative Map-Reduce. This also includes the requirement to work well with machine-learning and data-mining algorithms that will be increasingly important in the future with exceptional high volumes of data (i.e. Exabytes). We can summarize these activities as follows:

Action 5: Augment Middleware Products with Big Data Processing Techniques

---

Action 5 addresses the following recommendations:

- *Recommendation 3: Position Middleware in the Context of Big Data and Clouds*

Complementary to the previous aspects on active processing, the second aspect of 'big data' can be categorized as data curation and management activities. The middleware components of Europe are already strong in serving the needs of scientists for large volumes of data. Nevertheless, there are several important other functionalities that gain momentum in the scientific communities, especially in the context of data curation. Examples include the support for policy-based management (global, local, infrastructure-oriented) and extensibility of data management products in order to enable community-specific functions as part of already established data management plans in the scientific communities. Another example of the specific support for data curation and management is the enhancement of middleware products to support Persistent Identifier (PIDs) or new forms of data management approaches such as those inspired by the NoSQL movement (e.g. Mongo DB, etc.). We summarize as follows:

Action 6: Enhance Middleware with Data Curation and Management Functionality

---

Action 6 addresses the following recommendations:

- *Recommendation 3: Position Middleware in the Context of Big Data and Clouds*

The EMI project already conducted a study and published material on how EMI products can be positioned alongside private and public cloud offerings. Nevertheless, following the digital agenda strive 'every citizen cloud-enabled by 2020', the movement around cloud computing is ever increasing. We therefore see the demand for further activities to observe and potentially adopt to cloud deployments and required features since several scientific communities also find this model of deployment increasingly important. At the same time, the economic crisis in Europe may lead to increasing motivations in cloud hostings as well thus once again emphasizes that activities around cloud computing are still very relevant to the middleware community. In fact, this community trend represents a good example standing for a more general consideration of community trends. This means that also in the future it will be important to observe and adapt to future potential community deployments that are currently often done via Grids, might merge to clouds, or will even adopt other not yet known models in the future. We therefore summarize as follows:

Action 7: Observe and Adapt to Community Deployments around Emerging Clouds

Action 7 addresses the following recommendations:

- *Recommendation 3: Position Middleware in the Context of Big Data and Clouds*

The EMI products offered an overview of its use cases for a wide variety of scientific domains where all the scientific applications are used in a similar fashion but having still their unique overall requirements and thus detailed architecture. Focussing on the requirements of the European Strategy Forum on Research Infrastructures (ESFRI) alone will be not sufficient as their roadmap and activities are specifically focussed on large investments in creating a research infrastructure. The activities that middleware provider need to perform is to break these requirements into specific work on real application algorithms, code, data management practices and so forth. More and more solutions from these communities emerge as their own products (e.g. CLARIN and WebLicht Web services) and the middleware providers in Europe need to perform concrete case studies during development in order to provide solutions that meet two requirements: (1) being fully supported and adopted by end-user communities, and (2) provide the known stability and functionality strength of EU middleware. The middleware providers in Europe as part of the European Middleware Collaboration Group will invite therefore user community experts into the process of developing middleware and will perform use case oriented developments in the future as in the past, but with more emphasize on creating application case studies in parallel to developments. We summarize these activities as follows:

Action 8: Adopt Application Case Study and Use Case Oriented Developments

Action 8 addresses the following recommendations:

- *Recommendation 4: Perform Application Enabling During Middleware Development*

The re-usability of middleware products in a wide variety of scientific application use cases is key to increase the usage of European middleware products on the one hand and on the other it enables a path to sustainability for products where research disciplines and scientific communities have a clear interest to keep developments alive. In order to avoid 'big atomic middleware stacks' in the future, the European Middleware Collaboration Group (EMCG) will continue with the established model of product teams leaving it to products themselves to decide which granularity of a product makes sense. Nevertheless, the slight coordination of the group enables the identification of development teams with similar goals that focus on the same problem and thus paves the way for joint developments oriented towards solving one specific problem with one or two specific products rather to implement the same functionalities in all European middleware technologies in parallel. We summarize these activities as follows:

Action 9: Form Development Teams that Focus around One Specific Problem

Action 9 addresses the following recommendations:

- *Recommendation 5: Provide Reusable Products with Open Standard Interfaces*

The EMI project increased the adoption of open standard interfaces and schemas across its products. In addition, members of the EMI consortium have driven standards (e.g. UR for data and compute) by providing their distinctive knowledge in the field as part of standardization groups. Open standard specifications will continue to play an important role in the future to encourage trust in end-user communities that often fear to be bound to one middleware technology (i.e. vendor-lock). An important part of the work of the European Middleware Group Collaboration will be therefore the participation in standardization groups (e.g. within the open grid forum, research data alliance, organization for the advancement of structured information standards, etc.). The newly formed group will form sub-groups around standards across different products in order to understand their adoption potential and benefits to scientific user communities while not losing sight of defining interoperability criteria among them. These activities are summarized as follows:

Action 10: Drive and Adopt Open Standard Interface Specifications

Action 10 addresses the following recommendations:

- *Recommendation 5: Provide Reusable Products with Open Standard Interfaces*