MIDDLEWARE DEVELOPMENT AND INNOVATION ALLIANCE (MEDIA)

Proposal for a long-term, open, lightweight collaboration on distributed middleware technology

v. 0.4

25/03/2013
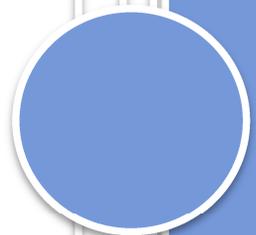
EXECUTIVE SUMMARY

The EMI (European Middleware Initiative) project is a collaboration among four of the major European providers of middleware for grid and distributed computing and data networks, namely ARC, dCache, gLite and UNICORE. The project is currently organized as a three-year collaboration partially funded by the European Commission and comprises 26 partners from 18 countries, 16 in Europe and 2 in the Asia-Pacific region.

After the initial period of research and consolidation that took place in grid infrastructures at the beginning of the years 2000, the growing usage of distributed computing and data resources by scientific communities and individual researchers required an effort of stabilization, simplification and standardization in the use of the associated software tools.

The EMI project has contributed to the realization of this vision by addressing a number of problems that prevented users from easily accessing and using the existing computing infrastructures, like usability, interoperability and standardization. Furthermore, by providing technical coordination across the major European middleware provided, EMI has allowed to establish single access point to most of the software development activities to the benefit of infrastructures like EGI and WLCG and the user communities they support.

The objectives of EMI were specifically to maintain and evolve a comprehensive set the middleware services used in European grid infrastructure such as EGI, WLCG and LSGC. The technical collaboration among EMI partners and with third-party software developers on standardization, interoperability, common functionality and forward-looking strategies is fundamental in keeping the middleware up-to-date and able to satisfy the evolving requirements of the research infrastructures. This includes not only work on the evolution of existing services, but also on the integration with new emerging technologies like clouds or flexible federations of security and information services.

With the end of the EMI project the fruitful collaboration among middleware developers and between developers and users (infrastructure managers and research communities) should not stop. The need has been clearly expressed for an open and inclusive mechanism to bring together technical experts, discuss of roadmaps, common development, selection and adoption of standards.

This document describes a proposal to form such a collaboration and outlines the possible scope, functions and operational mechanisms. The proposal focuses on the collaboration aspects of such an initiative and explicitly assumes that all practical aspects related to software development and management are within the responsibility of the development teams.

MOTIVATION

The EMI project has promoted and enforced the establishment of Product Teams (PT) as independent units of development and maintenance for individual middleware services of small groups of related services. Most of the PTs are located in single Institutes or shared across a very small number of partners. The PTs contribute software to the EMI distribution, which is used to group compatible sets of versioned services on specific platforms.

The PT-based model provides a more flexible approach to software development than a monolithic structure with centralized control. It is also better aligned with standard open source practices. However it also introduce a number of collaboration and communication challenges, since the different PTs need to be aware of the work carried out by the other PTs, handle any dependency constraint and discuss of common functionality or interoperability requirements.

The coordination and shared planning activities, the management of the EMI distribution and the underlying technical infrastructure like policies, repositories, continuous integration tools, web pages, announcements, meeting organization, etc. have all been functions provided by the EMI project.

The EMI project is funded for three years from May 2010 to April 2013. At the end of the project the Product Teams are expected to continue independently the maintenance and development activities of the middleware products. However, no follow-up project is foreseen to continue the coordination of common activities. It is therefore necessary to investigate and establish a new model for continuing the technical collaboration activities of the Product Teams, preserving the common work done until now and opening it up to new technologies and new services.

Based on input from the EMI PTs and other stakeholders in the research infrastructures, motivations have been identified to continue the collaboration after the EMI project in a lightweight and open form. The strongest motivation is to continue the collaboration on roadmaps definition, common functionality, interoperability and standardization. This is essentially based on having a reference organization where shared interests can be discussed, ideas proposed and checked and duplications of work avoided by sharing and reusing solutions.

Furthermore, the existence of a single point for establishing relationships with infrastructures and users is considered of interest to both providers and users. The interest in such a single point is not in providing representation, since most PTs have their own ways of engaging with their communities, but rather collecting and providing information about activities and plans in a single reference place and having a manageable set of knowledgeable experts with an overall understanding of the main technical features and constraints and proper contacts into the individual PTs for further discussion.

The availability of a mechanism to support the continuous and pro-active sharing of information across PTs is also considered a necessity to preserve interoperability, increase standardization and evolve the middleware services and define high-level technical roadmaps that can be openly discussed and shared.

SCOPE

MeDIA is a collaboration on software development and innovation. Its members must be developing and maintaining software products. However, specific activities needs to be open to users or infrastructure managers to ensure the proper flow of information and discuss about requirements and their implementation.

The software in the scope of MeDIA activities is composed of services for enabling distributed computing and data eScience infrastructures. Any existing or future technology that allows developers to provide efficient services is in scope, from traditional batch and grid systems to the emerging cloud systems and beyond.

The typical functional domains MeDIA wants to work on are security services, information services, data and compute services providing the shared core layer which eInfrastructures are built upon.

Higher-level services, applications, domain-specific frameworks or tools are not directly in scope of MeDIA. However, specific collaboration activities with developers of applications and high-level services can be foreseen to discuss about requirements, interfaces and functionality.

The main commitment and interest of MeDIA are the academic and scientific research infrastructures. Technology sharing and adoption within industrial applications are welcome and members from commercial providers are accepted provided they fall within the scope of MeDIA activities.

ACTIVITIES

MeDIA provides a forum for coordinating innovation and development of middleware services across the research infrastructures based on the members' interests and priorities. Participation to the activities is voluntary and bottom-up. It relies on an active sharing of information, proposals and ideas from members to other members. Although periodic, high-level monitoring of ongoing activities is foreseen, there is no top-down enforcement of requirements, priorities or technologies.

Requirements Coordination

Requirements within the scope of MeDIA must be of general interest across different products or be related to common functionality and interfaces or specifications. MeDIA can provide coordination in collecting and prioritizing requirements.

Collection

Requirements can be collected in two ways. Since MeDIA envisages to provide a common entry point for general requests from communities or organizations like EGI or WLCG, MeDIA can receive requirements that should be further propagated and discussed within the developers' community. However, since all MeDIA members have their own ways of collecting requirements from their users, MeDIA can provide a place where requirements of common interest can be shared and discussed among the developers. An updated set of requirements and their status is essential to develop and maintain roadmaps, therefore a common tracker can be provided where members can provide information and pointers to the implementation status.

Prioritization

Once a list of requirements is maintained, the requirements are prioritized with an open mechanism where all MeDIA members and the user communities can expressed their endorsement and priority. This is a periodic continuous process coordinated by MeDIA that allows to identify the high-priority requirements of most interest to most users and developers. The prioritization allows to define common roadmaps and organize common development activities. Of course priorities are not binding, but they provide transparent indications of where the developers can put their effort.

Roadmaps

MeDIA assumes that every development team has its own roadmaps to innovation. The roadmap may often affects areas of common interest for example related to security, interfaces or (re)use of low-level libraries. MeDIA provides a place where roadmaps can be discussed and synchronized and where shared development can be agreed.

Standardization and Technical Agreements

The EMI project achieved many important results in the design and implementation of common technical agreements in several areas. This worked has shown the importance of coordination and iterative practical implementation and validation. MeDIA provides this coordination mechanism through the definition of dedicated task forces, but it is not an alternative to standardization bodies like OGF. On the contrary it complements OGF

standardization activities by providing a practical mechanism for selecting and implementing existing usable standards and proving input to ongoing pre-standardization activities.

Common development

Although every product team is fully responsible for its products, there are at least two situations where a coordination of the implementation is highly beneficial. The first case is when the same functionality is required in many products. For example this is required when adopting common interfaces or specifications, migrating to or adopting common libraries, or implementing the same functionality in response to a common requirement (a clear example within EMI was the development of Nagios probes for all EMI services or the development of a common authentication library). Another relevant area is the porting of services to new operating systems, where the availability of components must be synchronized and experiences shared.

The second important area of common development is the shared design and implementation of new services across different teams. This is particularly necessary when a service or library is of general interest or when not enough effort is available within a single team. MeDIA can provide a place where shared teams or communities can be created and organized.

Technology surveys

MeDIA can also provide a forum for sharing and planning technology surveys and provide comparisons of products. The survey can be run by individual members or organized across multiple teams.

Forum for knowledge sharing (push model)

A fundamental aspects of information sharing in MeDIA is that members must actively provide information about the status of their work, their requirements, technical evaluations, etc. The benefit of an active sharing is that everybody can benefit from up-to-date information. A possible mechanism for sharing information is the use of the SciencePAD collaboration platform that includes (according to current plans) tools for information sharing, announcements, discussion forums and social-networking-style tools.

In addition to the technical coordination activities described above, a number of more engineering-oriented activities can benefit from an open mechanism to share information, although it is expected that the Product Teams in MeDIA have the necessary skills and resources to handle their software development process. The following areas have a potential for positive impact.

Cross-Product Integration, Certification and Testbeds

MeDIA members are expected to be able to test their services on their local resources. However, shared certification activities are critical to assess the quality and maturity of distributed software. For example, the availability of a "registry" of service end-points or development testbeds where integration, interoperability or standard compliance activities

can be performed across Products Teams could ease the certification tasks and improve the reliability of the services. The expected mechanism to implement cross-product activities is through task forces dedicated to specific, time-limited integration or certification activities.

Software Releases

Services have very often dependencies on other services or libraries and require a certain amount of explicit coordination to be released in a compatible way. The main foreseen mechanism to synchronize releases is to use the standard methods provided by mainstream operating systems communities like Fedora/EPEL or Debian. However, not all products may be available in the Fedora/EPEL or Debian testing repositories at the same time. A mechanism to communicate releases or discuss synchronization issues may be required. This function can be provided using a dedicated communication or announcement channel within MeDIA.

Software Repositories (EPEL, EMI/MeDIA, UMD, etc.)

The recommended mechanism to publish software packages for all MeDIA members is to use standard operating system repositories like Fedora/EPEL and Debian. However, there are several cases where additional repositories are needed. The following cases have been so far identified:

1) Packages not available in Fedora/EPEL, Debian, Maven Central or other community repositories
2) Packages in early technology preview state for internal development or certification
3) Packages with known incompatibilities at the official version level that have to be hold back until fixes are provided
4) Dedicated community distributions

If necessary, MeDIA could provide a dedicated repository to address cases (1) and (2) for packages maintained by its members and possibly any required third-party dependencies not otherwise available. However, repository functionality provided by EGI (UMD) or other community-specific repository should be used within each community. Cases (3) and (4) have to be taken care of by administrators of community or infrastructure repositories, but a MeDIA repository could be used as an upstream or reference repository.

Maintenance and Support

Routine maintenance of middleware products does not require large coordination efforts. Each Product Team has by now established procedures to assess software issues, fix them and release updates and patches.

However, there is still a general need for users and infrastructure operators to be able to know where to submit support requests, where to find help, discover who is responsible for each of the products they need and so.

The current user support mechanism based on GGUS as provided by EGI is so far the recommended solution to keep the requests and the tracking in a single place. However, it may not be available for use outside EGI or may not be used by all software providers in the MeDIA community. At the very least, MeDIA can provide a mechanism to collect and provide

from a single place all necessary information for the products developed by its members with links to their official user support channels.

It can be envisaged that a collaboration platform like SciencePAD (also bootstrapped by EMI, but of more general scope) can be used to provide this type of information registries.

RELATIONSHIPS

MeDIA needs to establish clear relationships and communication channels with infrastructures and user communities. The role of MeDIA with respect to other projects and initiatives must be defined.

It is expected that collaborations have to be established first and foremost with multi-domain infrastructures like EGI and OSG, research infrastructures like WLCG or those provided by ESFRI projects and with standardization coordination bodies like OGF. MeDIA positions itself as a mechanism to provide visibility to a community of software developers and as a channel to distribute information within that community through the shared effort of its members. It is not realistic as this point in time for MeDIA to represent the entire community and take decisions on its behalf. However, MeDIA can greatly help in simplifying the communication of information from external projects to the community members and assist in addressing common issues and coordinate common work, like the establishment of task forces dedicated to solving problems identified by the user communities. MeDIA members attending conferences, workshops, meetings, etc. will be asked to provide summaries of the discussions in areas of common interest to the entire community or to specific subsets like the task forces. Conversely technical experts from MeDIA can be invited to take part to coordination meetings to provide information about ongoing initiatives and assist in reaching out to the right development teams.

It is in any case expected that individual Product Teams have their own mechanism to deal with specific support requests and develop the functionality required by their user communities and that detailed technical discussions are managed directly by the PT members and not by MeDIA coordinators.

ORGANIZATION

Legal Status

MeDIA is a Community Collaboration without legal entity status. The members provide contributions in the form of time and resources as they can. The minimum levels of contributions are defined in the following sections. It is not excluded that in time and if required and desirable, different forms of collaboration such as a legal non-profit foundation can be established.

Membership

Membership in MeDIA is open to development teams (Product Teams) developing or maintaining middleware products or other software directly using the middleware products. It is expected that the teams in MeDIA operate as part of the official activities of an Institute, project, collaboration or company. A PT must be officially represented by a Team Leader, but any person officially working within the PT can contribute to MeDIA activities. The PT Leader is responsible to define who makes part of the Team and make them known to the MeDIA community via the collaboration tools to be provided by MeDIA.

Different levels of membership can be foreseen.

Full members have access to all tools and information and can participate in all activities. . The status of Full Membership requires active participation to MeDIA activities. Full members who become inactive over an extended period of time become Observers. Full Members can elect the Coordinator and the Technical Board members and can be candidates to be elected.

Observer members have access to subsets of the information and activities depending on their contributions and interests. Observer Members cannot elect the Coordinator and the Technical Board members and cannot be candidates to be elected.

Membership is free, but subject to acceptance of MeDIA policies on participation and sharing of information. Full membership is reserved to developers or maintainers of middleware services and components. Users, developers of high-level services or applications and infrastructure providers can participate as Observer members.

Governance

MeDIA has an explicit objective of being lightweight and not bureaucratic. However, a thin layer of governance and mechanisms to take the necessary decisions are needed. The following structure is foreseen

Coordinator: the MeDIA Coordinator supervises the overall procedures within MeDIA, organizes major events and can represent the MeDIA community in external events where information needs to be passed across. The Coordinator also leads the work of the Technical Board. A deputy Coordinator can be nominated to assist the Coordinator and share the load

of participating in the events. The Coordinator is elected by the community every 12 months from a shortlist of candidates.

Technical Board: it is composed of a limited number of technical experts from the community elected every 12 months from a shortlist of candidates. The TB members represent the different technical areas within scope of MeDIA and specific roles within the TB must be defined to ensure this (data, compute, security, etc.) The TB meets periodically to discuss membership requests, assess membership requirements, discuss the formation of task forces, and propose and organize events. TB members are expected to be active and engaged in their role. If an TB member is not active enough after his/her election or wishes to withdraw from the role, the other members can call for a replacement election before the end of the mandate.

Full Members: are individual Product Teams represented by their Team Leader. The Team Leaders must have full authority to take decisions sign out for commitments on behalf of their PT.

Observer Members: are projects or initiatives interested in being informed of the technical activities of MeDIA even if they do not directly develop software or their software is out of scope for MeDIA.

Task Forces: the main mechanism to discuss and address specific topics of interests and carry out any agreed technical activity. All activities from, for example, repository management to standards definitions or software development, are executed as part of a task force. Each task force has an appointed leader duration, scope, expected outcome, etc. Task Forces can be proposed by any Full Member of MeDIA and have to be approved or rejected by the Technical Board.

Procedures

Membership application: Any provider of software fulfilling the membership criteria for Full or Observer status can apply to become a member. Membership requests are received by the Technical Board that verifies the criteria and formally accepts the new member. Membership approval requires a majority vote of the Technical Board members. Rejection of requests must be clearly motivated, since the overarching goal of MeDIA is openness and inclusiveness. The Team Leader of the new Full Member must publish and keep up-to-date the names of its team members using the MeDIA collaboration tools. A Full Member is expected to take actively part in the MeDIA discussions and activities within the scope of their area of interest. A Full Member can become an Observer Member if the team is not active and committed. The change of status can be proposed by the Member or by the Technical Board.

Election of Coordinator

The MeDIA Coordinator is elected from voluntary candidatures from any team member of Full Members. Candidatures are invited once a year, the Coordinator is elected by majority vote of all registered Full Member at the date of the vote. The Coordinator has a mandate for

one year from the 1st of the month following the election and can be re-elected for further mandates.

Election of Technical Board Members

The MeDIA Technical Board members are elected from voluntary candidatures from any team member of Full Members. Candidatures are invited once a year for the different TB roles. The roles represent areas of technical expertise, like security, data management, computing, etc.[1] The EC members are elected by majority vote of all registered Full Member at the date of the vote. The EC has a mandate for one year from the 1st of the month following the election and individual members can be re-elected for further mandates. In case of resignation or inactivity of an EC member, the Coordinator and the other members can call for a secondary vote to replace the defaulting member for the remaining duration of EC mandate.

Creation of Task Forces

Task forces are the main mechanism to implement the MeDIA community decisions. Task forces can be proposed at any time by any Full Member leader or by the Coordinator the Technical Board members. Task forces proposals are assessed by the Coordinator, the EC members and the proposing member (for community peoposals). If the proposal fit all agreed criteria of leadership, scope, duration, milestones, final outcome, etc., they are approved and published to the whole community for information. They enter immediately into force.

Costs and Effort

MeDIA is an open collaboration based on voluntary participation and contributions from its members. In its present form, no membership fees are considered, but members are required to actively participate to the community activities and events, in particular to staff the Task Forces and share information in an active, bottom-up way.

The Coordinator and Technical Board members are required to dedicate a fraction of their working time to evaluate membership requests and task force proposals, to organize and share information with the community and from the community to external projects and infrastructures, to organize events and to discuss high-level technological directions.

All community members are expected to actively engage in the activities, promote the MeDIA in their working environment and share information with the other community members.

Any communication and collaboration requirement (meetings, conference calls, mailing lists, etc.) is defined within the scope of each task force. A general community event to present the main activities, discuss of technical strategies and elect the Coordinator and the Technical Board members is expected to take place once per year.

---

[1] The roles are not yet defined at the time of writing this draft and will have to be defined in the follow-up public discussions about MeDIA

The costs and effort required to perform the MeDIA technical activities and who should bear them must be defined as part of the Task Force proposals. Whenever a Task Force is proposed, an estimate of the required effort and resources must be provided and they must be agreed and allocated among the proposed Task Force members before the Task Force can be endorsed.

WORKING ENVIRONMENT

The MeDIA will provide a set of collaboration and communication tools to facilitate the sharing of information among members and any required coordination activity. Beside standard mailing lists, Facebook or Twitter, it is envisaged that MeDIA would make use of the software collaboration platform being set up by SciencePAD to publish and share information about development teams, people, software, task forces, collaborations, etc. Each Task Force is also free to define their own internal communication mechanism, as long as the outcome of their activities is made available to the community.