

EMI Data, the unified European Data Management Middleware

In May 2010, one of the most challenging European initiatives in the context of software provision (EMI) was launched, targeting a universal Grid Middleware Distribution across the major players in this field, namely gLite, ARC, UNICORE and dCache. The goal is to offer an easy-to-use, reliable and well maintained software stack from which large infrastructures like EGI, PRACE and other Distributed Computing Infrastructures (DCIs) may select the most appropriate components for their particular needs. A second major objective is to simplify interoperability between historically 'different' computing approaches, such as HPC and Cluster Grid Computing. This presentation will focus on the plans, challenges and the progress of the European Middleware Initiative in terms of data management. The authors will present the EMI approach to unify and simplify the security and authorization layer in storage by implementing standard, well-known authentication mechanisms, as well as interfacing a common authorization layer (ARGUS) across the supported storage elements. Moreover, in order to allow seamless access of data across ARC, UNICORE and gLite and to attract new communities to use the EMI distribution, standard data control and data access protocols are implemented or improved where needed. This includes the Storage Resource Manager protocol (SRM), currently the most important ingredient in the WLCG storage model, as well as the network data access protocols, NFS4.1 and WebDav. Besides many other activities, data management in EMI will improve interoperability of file catalogues and will address synchronization between the content of catalogues and the underlying storage systems.

Primary authors : Dr. FUHRMANN, Patrick (DESY)