

Outcome of the EuroVO-DCA WP5 Activities

C. Vuerli ^(1,2), *G. Taffoni* ⁽³⁾, *M. Sponza* ⁽¹⁾,
F. Pasian ^(1,2)

(1) INAF – Astronomical Observatory of Trieste

(2) INAF – SI (Information System)

(3) Sincrotrone Trieste SCpA

- **Astronomical Databases in Grid**
- **The DCA (Data Center Alliance) Project**
- **WP5 and its Outcome**
- **What after WP5: The EGEE-III DB WG and AIDA**

- The VObs consists of a collection of data centers each with unique collections of **astronomical data**, **software systems** and **processing capabilities**
- All these three entities that characterize the VObs can be handled in a homogeneous and uniform way by means of the Grid and its integrated tools and services
 - Astronomical data (archives, catalogues, etc.) federated in the VObs are typically stored in database repositories that can be in turn integrated in the GIS
 - Software systems can be organized in software repositories as well. Software repositories can be handled like data repositories
 - In Grid VObs users find the computing power/capabilities they need to process the huge amount of VObs data
- All this can be carried out in a uniform working environment

- **Two possible approaches to make the Grid and the VObs interoperable**
 - Full integration: the user access the VObs through the Grid; it is mainly a Grid user with his/her personal certificate; authentication/authorization mechanisms are those of the Grid; the user operates in the Grid environment in the usual way; VObs resources (mainly data and software repositories) are fully integrated in the Grid; VObs data and software repositories are resources of the Grid
 - Light integration: the user access the VObs and operates in the VObs environment through its available tools and services; Grid resources (typically computational resources and processing capabilities) are requested and used when necessary

- **When organizing and processing large amounts of data in Astronomy, databases play a central role**
- **Initiatives aimed at setting up national and international Virtual Observatories are a clear evidence of this**
- **Because of the great importance of databases in astronomy, the success in porting astronomical applications in Grid is in tight correlation with the level of integration between the Grid and database resources**
- **For a full integration between the VObs and the Grid a strong level of integration between the Grid and databases is mandatory**

- **The EU FP6 funded EuroVO-DCA was an alliance of national data centre communities to assist European Data Centers to take up VObs standards, share best practice for data providers, consolidate operational requirements for VObs-enabled tools and systems and enable the identification and promotion of scientific requirements from programs of strategic national interest that require VObs technologies and services**

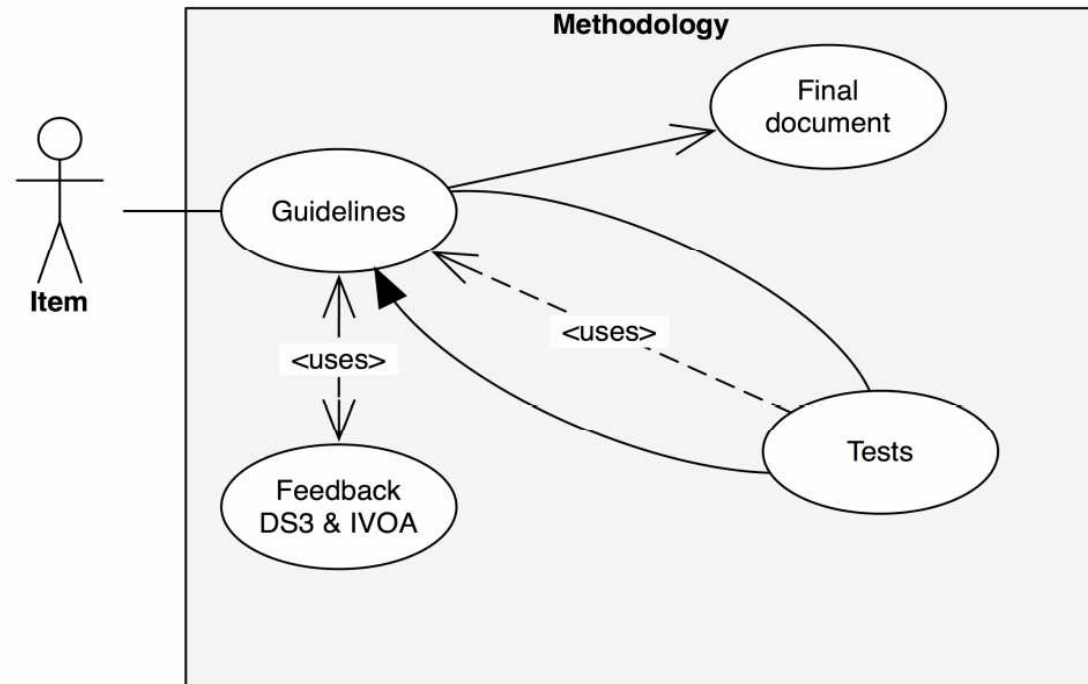
- **One of the objectives of EuroVO-DCA was:**
 - seek coordination with national and international projects for computational Grids → **WP5**
- **EuroVO-DCA WP5**
 - Coordinated by INAF (G. Taffoni)
 - Activity Plan
 - Knowledge acquisition
 - Interactions/Interoperation
 - Coordination
 - Suggestions and guidelines
 - Dissemination & Re-dissemination

- **Knowledge Acquisition**

- **Census of EU grid initiatives is a living document, available at the DCA wiki page [Should be picked up by EuroVO-AIDA]**
- **Grid standards on Authentication and Authorization, Data Management and Job Management**
 - Correlated with the IVOA standards

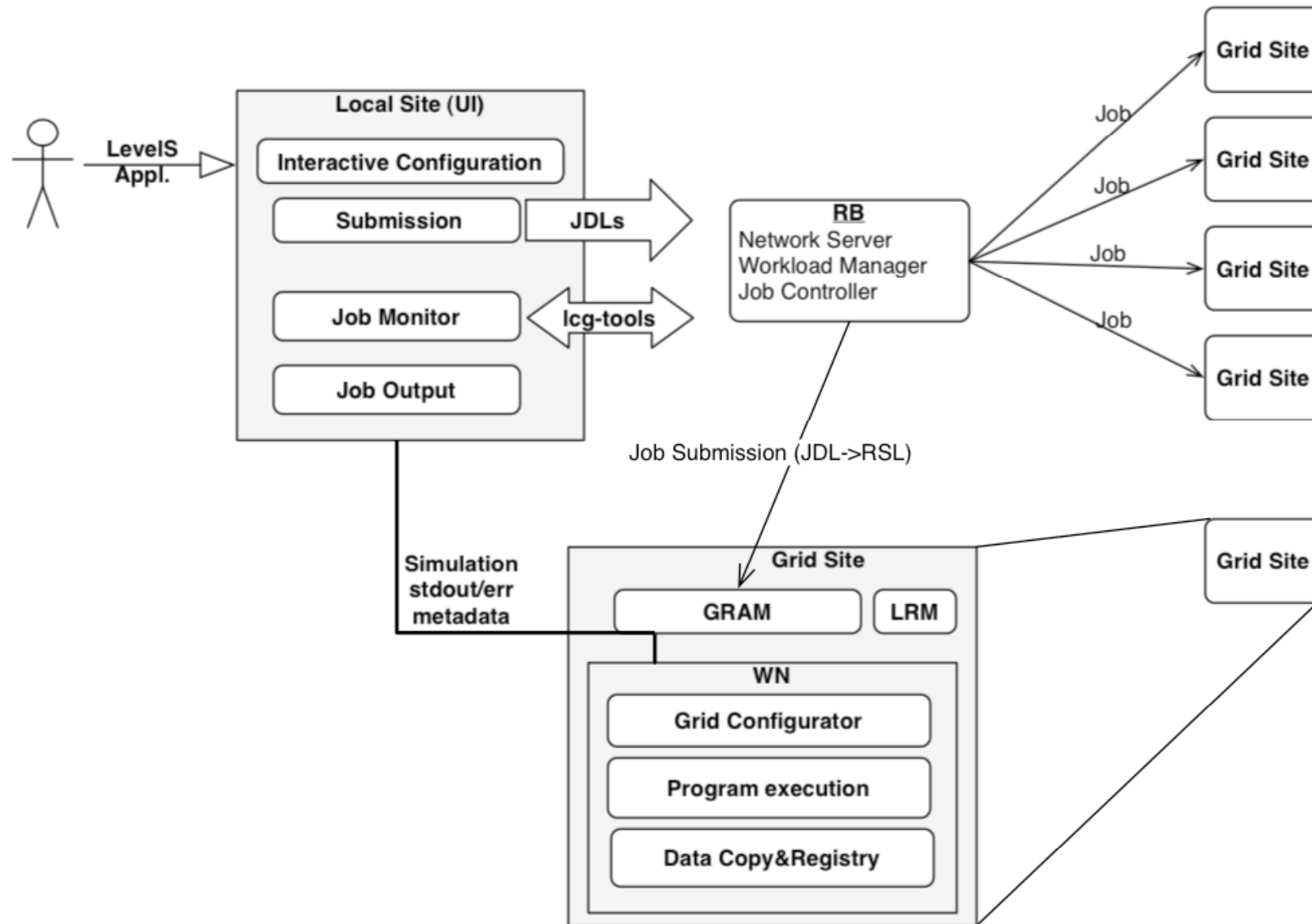
Methodology

- Documents are not enough
- We identified some DCs wanting to use the Grid
- We identified two use cases:
 - data oriented
 - theory oriented
- Experience drove our work
- Trial-and-Error → Prototypes



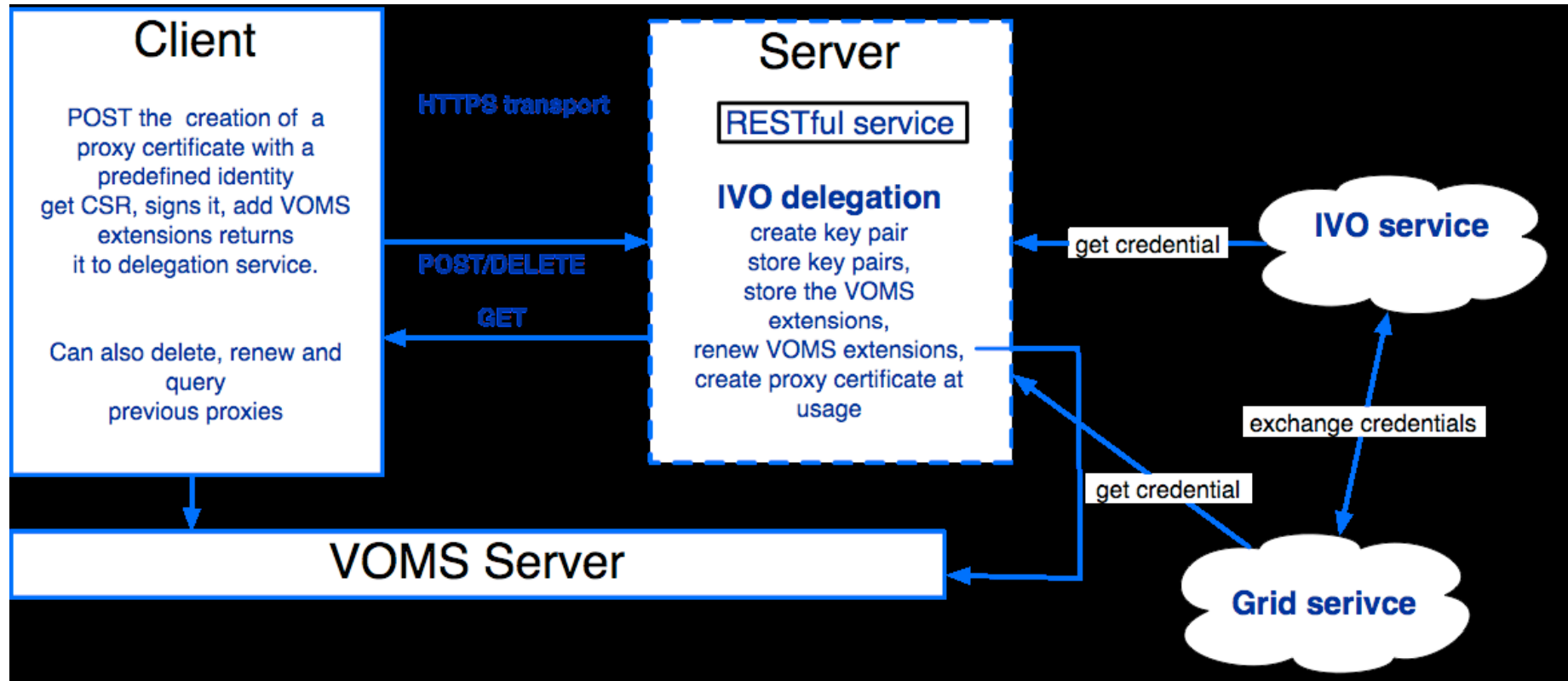
- **Two use-cases identified:**
 - **Data oriented: GSC-II (Guide Star Catalog-II)**
 - An astronomical database generated at STScI starting from scans of 9541 Schmidt photographic plates coming from surveys of the Palomar and UK Telescopes.
 - The GSC-II has been split in slices covering different regions of the sky. These slices have been stored and processed in different Grid nodes.
 - **Theory oriented: BaSTI (A Bag for Stellar Tracks and Isochrones)**
 - BaSTI is currently one of the most updated database of stellar models all around the world and it is a formidable tool for population synthesis investigations. In the next future, it will provide also theoretical predictions about integrated colors and magnitudes as well as integrated spectra for complex stellar populations

- **Steps in common**
 - **Access to data repositories to get already produced real/simulated data**
 - **Production of new simulated data and their storage in data repositories**
 - **Data Management: visualization, data mining, further processing, etc.**
 - **To manage extracted data, software already in place can be used (access to data repositories)**
 - **It shall be possible to deploy new software tools/packages in Grid and make them available in software repositories**

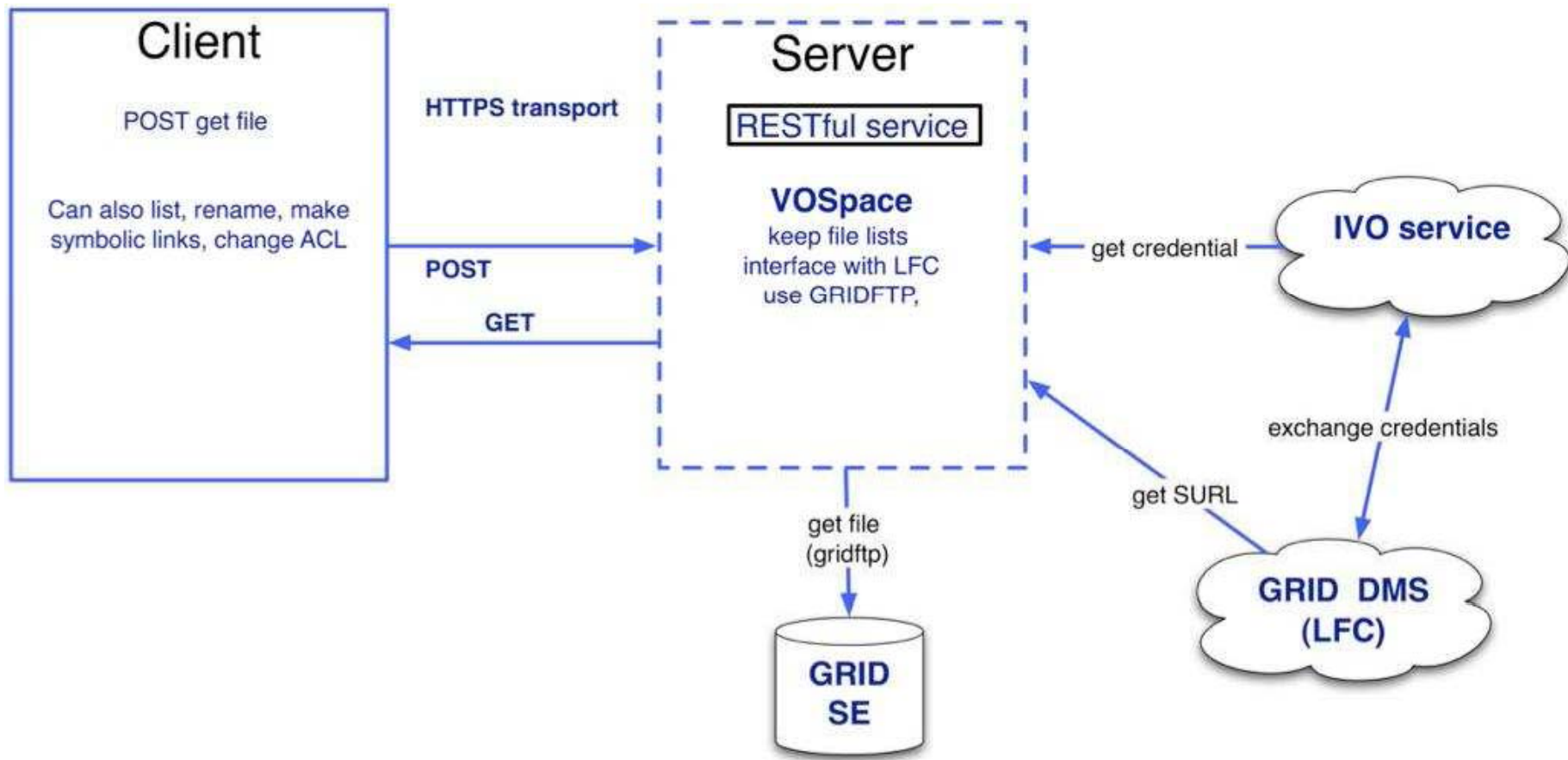


- **Authentication and Authorization**
- **Grid data management: data access for/to DCs**
- **Grid job management and workflow tools: job access/exchange for DCs**
- All these topics require now further work within projects and initiatives that can inherit what produced within WP5

- **VOMS: Virtual Organization membership service**
- **IVOA delegation service is not available (just prototype)**
- **We add the VOMS plug-in**



VOSpace – LFC implementation



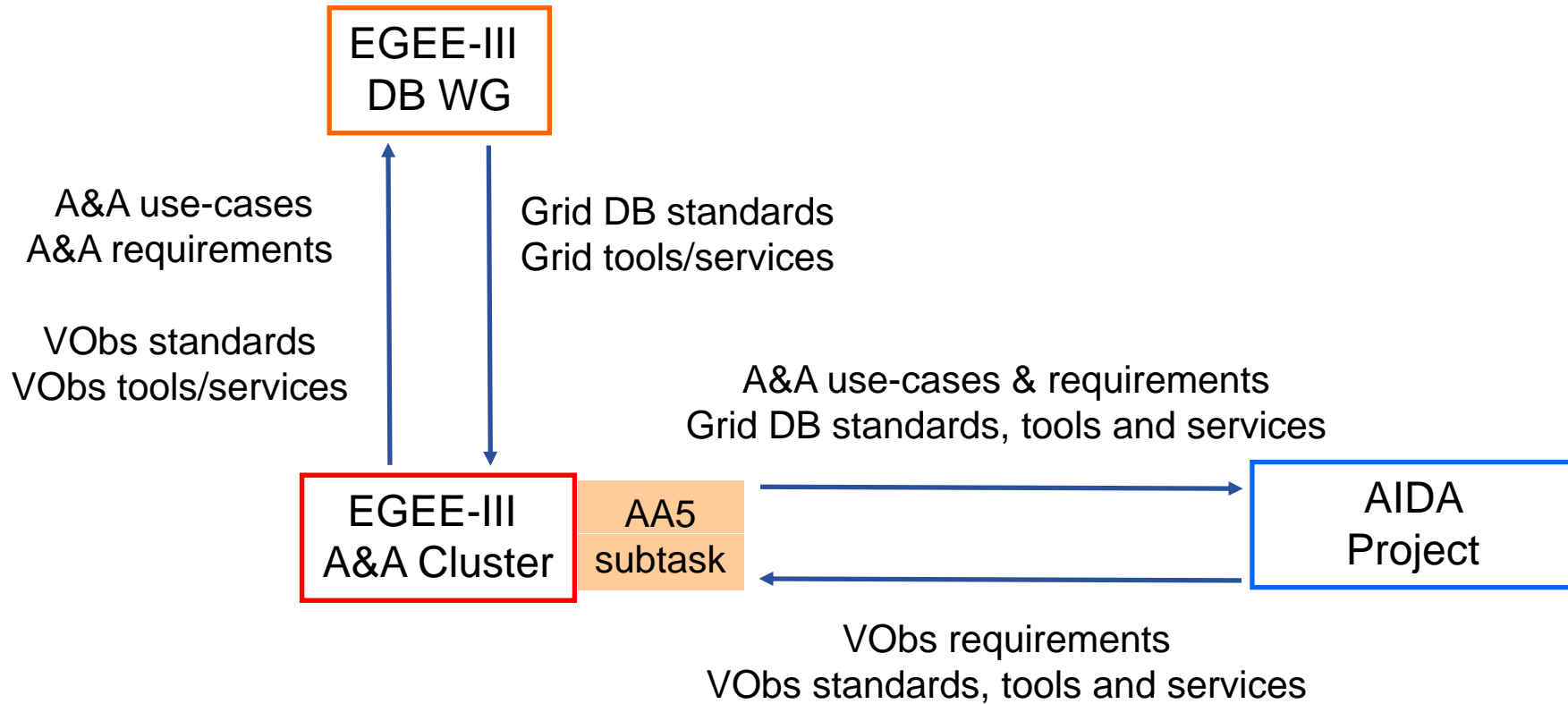
- **Interaction/Interoperability**
 - DCA VOrg in EGEE: “dca.eurovo.org”
 - “Astro” cluster included in EGEE-III project
 - Participation in Grid conferences and forums (in particular OGF, EGEE)
 - OGF interoperability group (Astro-RG, work in progress in coordination with IVOA)
- **Coordination**
 - **Coordination with WP4**
 - back-to-back workshop
 - direct collaboration with Theory teams
 - **Coordination with VO-TECH DS3**
 - **Coordination with IVOA GWS group**
 - **Coordination with NA3 EGEE for training**
 - **Coordination between National VObs projects and Grid projects**

- **Dissemination**
 - Wiki page for links and documents
 - EGEE NA4 and NA3
 - Local and Campus Grids
 - Point DCs to workshops and training events (by Grid projects)
 - Direct support and collaboration
- **Participation to OGF and EGEE Forums/Conferences**
- **Cross-Fertilization**
 - Earth Science Community (Frascati, December 2006)
 - ICTP workshop on Instruments and Sensors in Grid (Trieste, April 2007)
 - Biomed Grid school (Varenna, May 2007)
- **Various schools and courses organized**

- **Objectives**
 - **Survey requirements of existing applications and identify the core use cases for database access from the grid infrastructure**
 - **Survey the currently-used techniques and existing tools to access databases from the grid infrastructure**
 - **Make recommendations of particular tools for each identified use case, or identify gaps that require development**
 - **Make recommendations for changes in the gLite services or service/site configurations to facilitate use of database access tools on the grid infrastructure.**

- **The project is supported by EU in the framework of the FP7 e-Infrastructure Scientific Research Repositories initiative**
- **Project Overview**
 - **EuroVO-AIDA aims at unifying the digital data collections of European astronomy, integrating their access mechanisms with evolving e-technologies, and enhancing the science extracted from these datasets. The EuroVO-AIDA project is proposed to lead the transition of EuroVO into an operational phase**
- **EuroVO-AIDA integrates the technology, networking and service activities of EuroVO into a fully functioning structure**

- Interactions between VObs and the Grid through three main actors



Thank you