Euro-VO HPC needs

S. Derriere – CDS, Observatoire Astronomique de Strasbourg
sebastien.derriere@astro.unistra.fr

F. Genova (CDS), F. Pasian (INAF)

The astronomical Virtual Observatory

- Large volumes of distributed and heterogeneous data :

    - Images, Data cubes

    - Catalogues

    - Multi-wavelength : optical, radio interferometry

    - Computer simulations, models

- Need for interoperability

- Virtual Observatory = VObs or VO

Vobs Registry authorities

Metadata in XML + OAI-PHM

International Virtual Observatory Alliance : IVOA

- International coordination (like W3C)
  - WGs
  - Standards

At the European Level Euro-VO

- VOTECH (FP6, 2005-2009) : Technology Center, design study for technical preparation

- Euro-VO DCA Data Center Alliance (FP6 e-infrastructure Communication network development, 2006-2008) :

- Euro-VO AIDA Astronomical Infrastructure for Data Access (FP7 e-infrastructure Scientific Research Repositories, 2008-2010)

- Euro-VO ICE International Cooperation Empowerment (FP7, 2010-2012) : coordination and take-up

- Euro-VO CoSADIE Collaborative and Sustainable Astronomical Data Infrastructure for Europe (FP7 Coordination Action, 2012-2014)
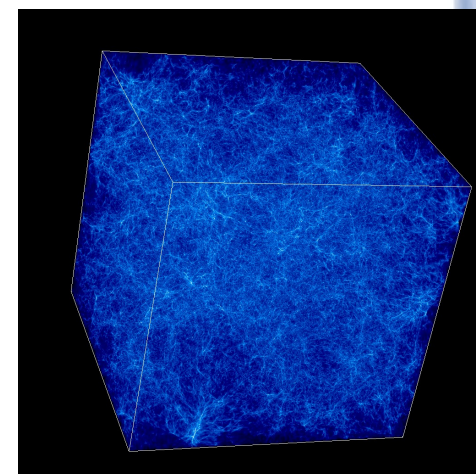
European resources

Astronomical data

- Catalogues, lists of sources (up to few 10^9)
  - Small files : « long tail », Simple Store
  - Large Surveys : several GB, up to TB

- Images
  - Individual : MB, GB (few Gpixel cameras)
  - Sky surveys : several TB/night, PB/year

- Radio interferometry
  - LOFAR : ~100Gb/s of raw data
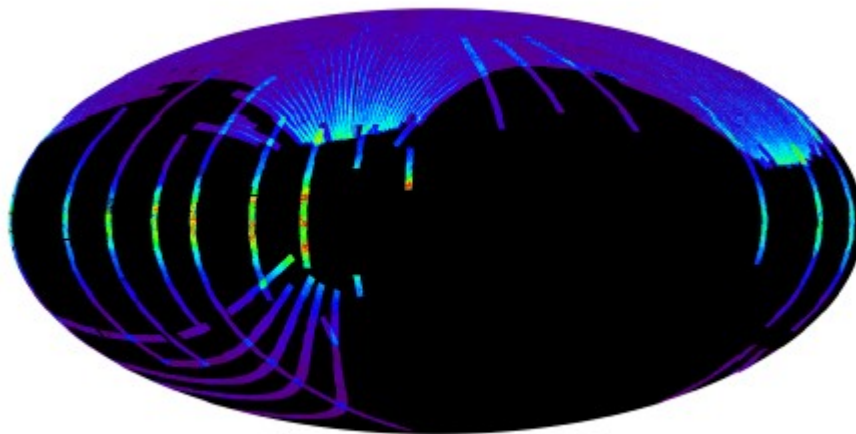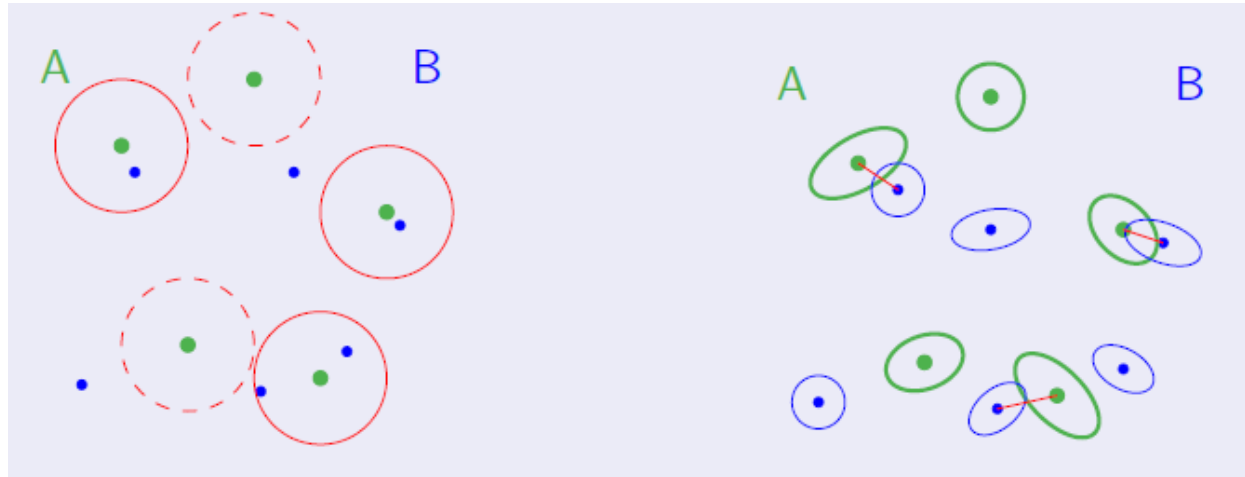 - ALMA, SKA : even larger !

- And much more...

Computing needs

- **Large simulations (e.g. Cosmology, 10^6 hours)**
    - Also produces massive data sets
- **Multi-wavelength astronomy : cross-match...**
- **Pipelines, workflows**
    - Reprocessing, source extraction...
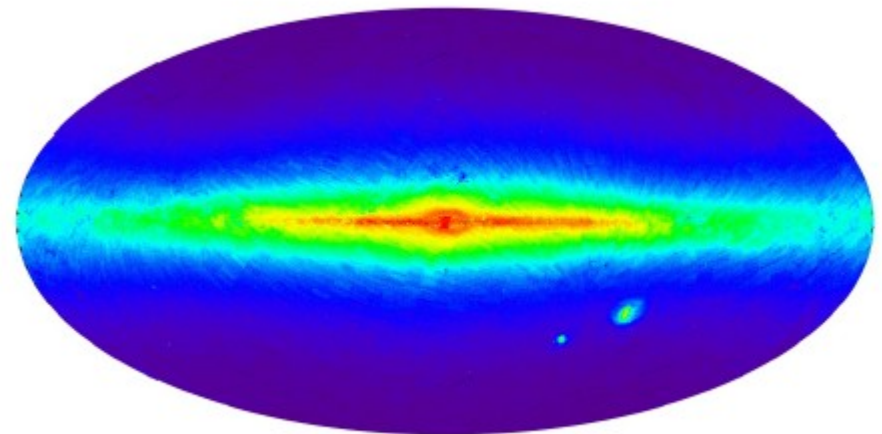- **Distributed computation (each facility processing and distributing data)**

Cross-match

SDSS DR7 (~357 000 000 sources)

2MASS (~470 000 000 sources)

Cross-match

Data Management

- Usually in astronomy : short proprietary period, then open access to data

- IVOA (Grid & WebServices WG) has defined Single Sign-On profile authentication mechanisms

Grid and Web Services WG

- The IVOA has a Grid and WS working group, which has defined some standards :

    - Universal Worker Service : asynchronous execution of jobs

    - VOSpace : interface to distributed storage

    - Support interfaces : SOAP and REST Web Services requirements

Grid, cloud and the VO

- The Euro-VO e-infrastructure transparently benefits of the use of the European scientific network (GÉANT)

- CoSADIE WP 3.4 : Interface Data Centres with Supercomputing-based grids (PRACE), cluster-based grids (EGI.eu) and clouds

    – Areas of commonality and complementarity between VObs and distributed computing infrastructures (DCI)

    – DCIs - VObs interoperability, adapt VObs standards on top of DCIs middleware

Astro - Community Group

- Astro-CG charter approved by Grid Forum Steering Group in August 2012

- Framework where entities dealing with e-Infrastructures and operating within the A&A community meet, share common issues, collaborate looking for solutions

- Definition of services and interfaces enabling the integration of e-Infrastructures and of their related resources used by astronomers for their scientific production

Astro-CG

- Contains a liaison group with IVOA : interoperability between distributed computing e-infrastructures (DCIs) and data e-infrastructure (DDIs)

- Main problem at present : feedback from the community is missing almost at all

    – Poor understanding of user's needs

    – Poor scientific/technological offer

    – Lack of training and dissemination

    – Lack of funds

Euro-VO, Astro-CG and EUDAT

- Many similarities between VObs and EUDAT
  - IVOA Semantics WG vs Metadata
  - VOSpace vs Simple Store
  - Grid and WS vs Data Staging & HPC
- Apply EUDAT services to astronomical data
  - Identify needed interfaces with VObs standards
- Share experience from existing VObs implementations and take-up from community

Thank you !

Questions ?