A few experiences with VOSpace

André Schaaff, Cyril Pestel

CDS

Topics

- **Brief and informal comments about the use (client and server side) of VOSpace at CDS**

- **Based on different experiences : Aladin, ...**

- **What we have learned**

Developments, implementations and experiences done in the frame of the European VOTECH project

VOSpace at CDS

- **VOSpace**

    - **Web Services : Tomcat 5.5.27 & Apache Axis2 & Rampart 1.3**

    - **IVOA VOSpace 1.1 RC3**

    - **H2 (pure Java) database for the metadata storage**

- **iRODS 2.01 (integrated Rule-based Data System)**

    - **Developed by SDSC, open source, easy to move or resize the resources allocated to the VOSpace**

- **Access both through VOSpace interface (CDS and public) or directly to iRODS (Jargon API for CDS internal use (e.g. CDS Portal, cf. Thomas's presentation))**

- **Hardware**

    - **2 quad core servers with 12 TB data in RAID 5 and 2 TB system in RAID 1**

    - **DRBD and Heartbeat mechanism for the High Availability**
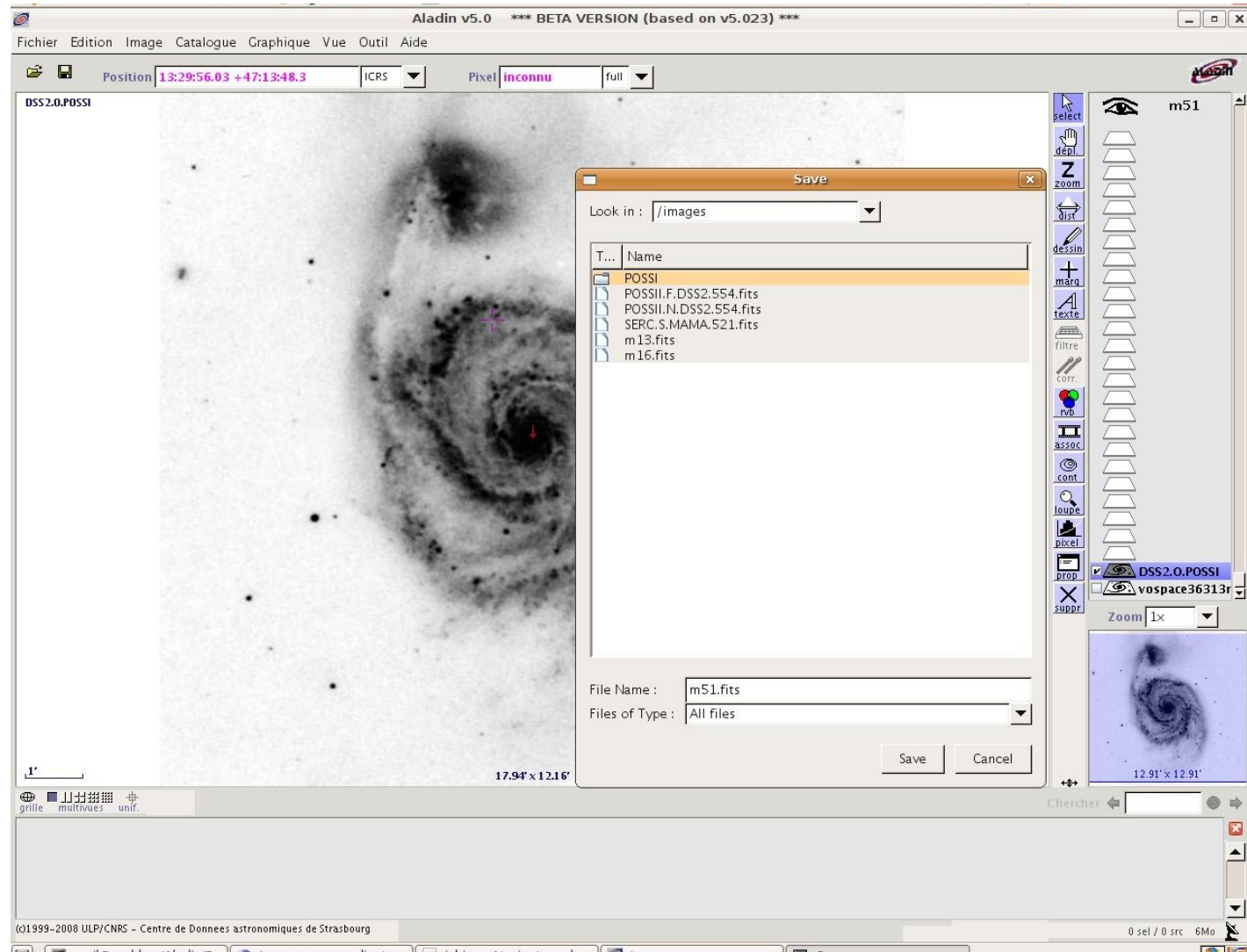
VOSpace access

■ **How to access the VOSpace ?**

- ■ **With a tool like Aladin**

    - ■ Implement the WSDL ?

        - ➡ Axis libraries in Aladin... (Not done)

    - ■ Or develop a plug-in ?

        - ➡ Not in the core tool, the plug-in (and the axis overhead) can be installed only if needed (Done)

    - ■ Or through another tool with the Plastic/SAMP mechanism (Done)

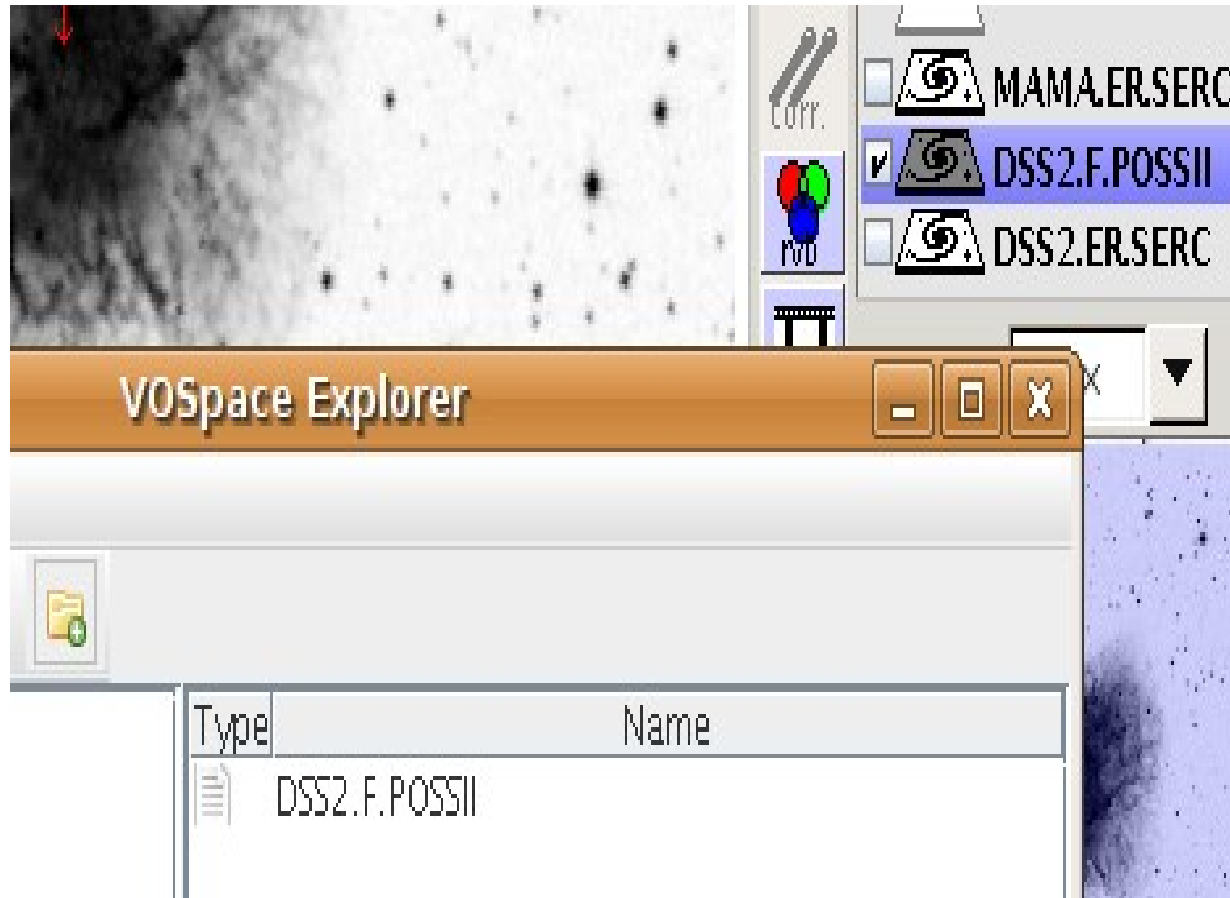   - ■ Drag and drop with a file explorer having VOSpace capabilities (Done)

    - ■ ...

VOSpace access in Aladin through a plug-in

- **A simple VOSpace file chooser implemented as a plug-in**

VOSpace access through Plastic

■ **VOSpace Explorer makes the interaction possible through Plastic (/SAMP) or Drag and Drop with Aladin**

In case of Aladin

■ **Different kinds of manipulations**

■ **With the VOSpace plug-in**

■ As a remote USB key

➡ Large working space

➡ Access to the data from everywhere

➡ Possible easy sharing with other people

➡ But just interactive load/save from/to VOSpace

■ **Through Plastic**

■ Use of a workflow tool with VOSpace and Plastic interface implementations to make non interactive treatment on FITS image and to send the result (stored on VOSpace) to Aladin (possible to other VO tools enabling Plastic)

■ ...

VOSpace implementation Up to scale test

- **Experience in the frame of EuroVO AÏDA with ~300,000 FITS images (~ 1 TB)**

  - **Creation of 6 directories with ~10000 subdirectories each**

  - **5 images per subdirectory**

  - **~ 3 MB per image**

- **VOSpace feeding**

  - **Very fast at the beginning but performances have gradually fall down**

  - **Index problems in the metadata DB**

    - Not visible in case of current use (e.g. temporary storage of a few files) => importance of deep tests

  - **iRODS tools (for example the Web client) very efficient even for large datasets**

    - Useful (avoids VOSpace interface overhead) to give an access to the native data storage (VOSpace capabilities) if possible

Conclusion

- **VOSpace access**
  - Integration (VOSpace < v 2.0) in a tool means SOAP libraries to include in the distribution
  - Possible use through other tools via PLASTIC/SAMP
  - VOSpace access is used both in workflows and by interactive tools and it is not possible to have in this case too much time overheads in the metadata access (example of 10000 subdirectories)

- **Always necessary to make tests with large datasets**
  - It is not possible to say to the user : "You are limited to 1000 files otherwise the system could become unstable..."

- **VOSpace REST version welcome to simplify the integration of VOSpace access into tools**
  - We will implement it asap