

Victoria Interop, 17-21 May 2010

# VOSpace 2.0 implementation (status) at CDS

André Schaaff, Vincent Meslard



# Today's topics

- Previous work
- VOSpace 2.0 now !
- More technical details
- Service architecture
- Comments
- Security
- Roadmap
- References and collaboration
- Conclusion

## Previous work

- VOSpace 1.\*
  - Implementation based on
    - Axis for the SOAP part
    - iRODS as the storage server side
  - Development of tools both over VOSpace and iRODS (VOSpace Explorer, Aladin FileChooser plugin, ...)
  - Link with work about workflows
- Comments
  - SOAP was not really adapted
  - For example it was slow to show the content (in case of a large number of files) of a container node (~ directory)
  - Need of external libraries like Axis

## Previous work (2)

- iRODS experience
  - In use since 2 years for a few experiments and for the CDS future Portal
  - No problem to evolve from the 1.0 to the last 2.3 version
  - Easy to use in Java developments through the Jargon API
- ...

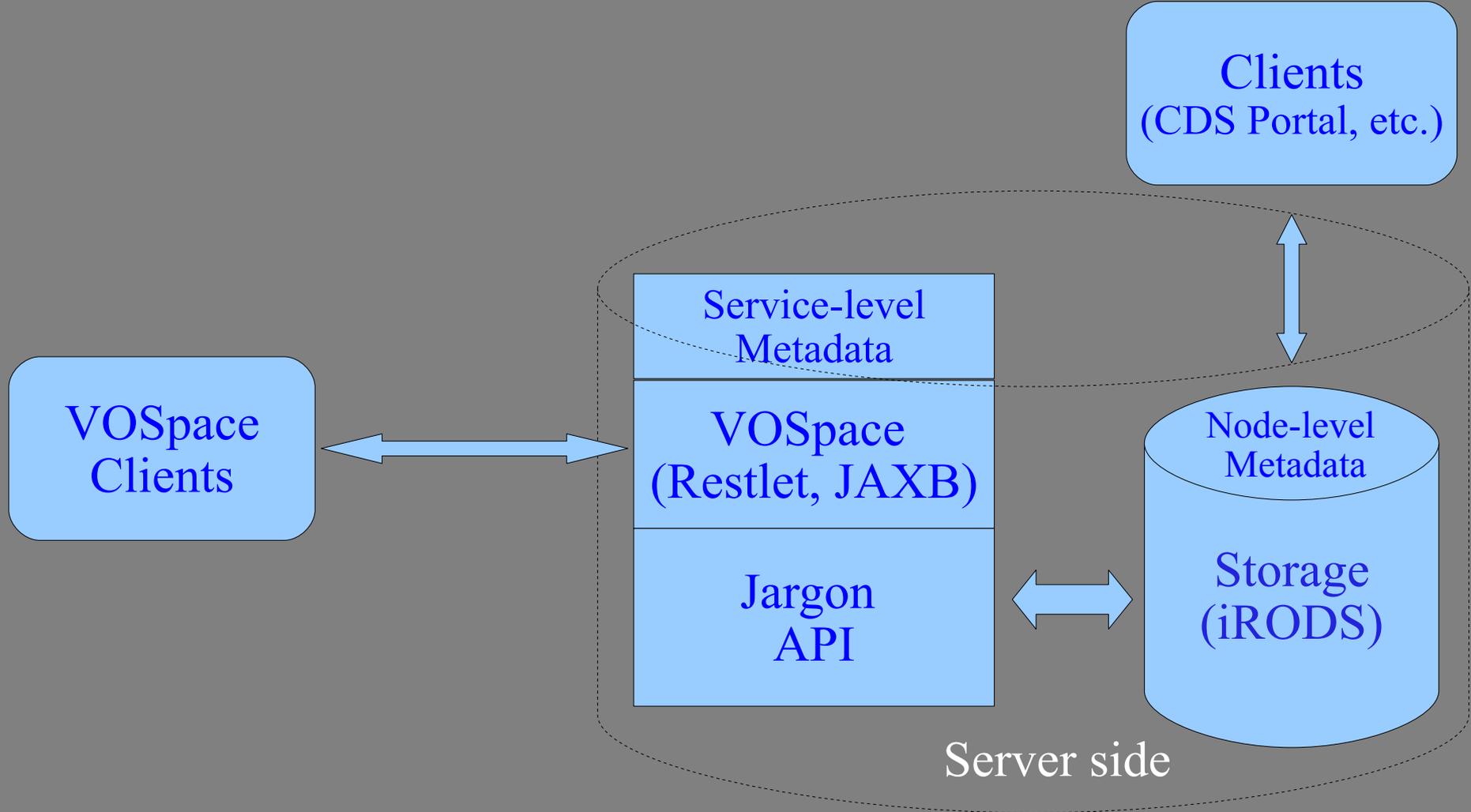
## VOSpace 2.0 now !

- Development from scratch (excepted iRODS and experience from 1.\* implementation)
- We have simplified the management of the metadata
  - In VOSpace 1.\* we had a database for VOSpace metadata and another for iRODS metadata (iCAT “iRODS CATalog”)
  - iRODS 2.3 (released the 12 March 2010) provides new features like the extension of iCAT and Jargon (Java API to access iRODS) allows the setting of a list of (Attribute-Value-Unit) items for each file
  - In the VOSpace 2.0 implementation we use this features to add the VOSpace “node level” metadata to iRODS metadata
- The “service level” metadata are stored in a simple configuration file

## More technical details

- Based on Restlet framework and JAXB
  - Implementation as generic as possible
    - to take into account last minute changes and future evolutions
    - not to close to iRODS
- iRODS is used as main back-end providing some easiness
  - No need of another metadata layer, metadata storage is done by iRODS
  - A true java abstraction through Jargon (developed by the iRODS' team)
  - All the possibilities and extents enabled by iRODS (replication, authorization, etc.)

# Service architecture



Server side is fully duplicated with a heartbeat mechanism

## Comments

- Modification of the VOSpace 2.0 schema to be valid and thus to generate the XML binding Java classes
- StructuredNodeType (example : VOTable file) and Searches (find nodes which match a set of specified values) not yet implemented
- Even if nodes metadata is stored in iRODS it is easy to switch to a different storage system
- Even if the implementation is near the end it doesn't mean that everything works → period of deep test

# Security

- Authentication, Authorization (SecurityFacade from AstroGrid) tested at CDS (Collaboration with Guy Rixon)
- Will be retested with the final implementation → decide which security policy to apply

# Roadmap

- Complete the implementation
  - Search features
  - StructuredDataNode implementation for a short list of formats
  - Evaluate the performances for large data containers (at least a few thousand nodes)
- Decide which access policy to implement
- After full implementation : ready to provide the source code and to give some support to people who would like to implement VOSpace 2.0, with or without iRODS

# VOSpace (& iRODS), references & collaboration

- References
  - 3 presentations in other communities done in 2009
    - In December 2009 at Queen Mary University of London, Interoperability of Digital Repositories meeting
    - In June 2009 at a CNES workshop about data preservation
    - In February 2009 at a joint IN2P3 – DICE iRODS workshop
  - iRODS main page illustration points to the VOSpace-iRODS article
    - <https://www.irods.org/>
- Collaboration is going on with people from DICE (Reagan Moore's team) and IN2P3 (Jean-Yves Nief) to follow the iRODS evolutions

## Conclusion

- Full implementation expected by the end of June
- Open sources and ready to give a support to implement a VOSpace 2.0