

# CDS WP6 report

about VOTable parsing, REST Best practice and VOSpace/iRODS

André Schaaff, Vincent Meslard



# Topics

- VOTable parsing : SAVOT 3.0
- REST Best practice (with Norman Gray)
- VOSpace (& iRODS)
  - Lessons learned, ongoing work, collaborations and references
- SAMP implementation and dissemination (following presentation)

# VOTable : SAVOT 3.0

## ■ Current status

- Beta version developed after the May Interop
- VOTable 1.2 recommendation process is not fully completed
- A stable version will be released just after the 1.2 version publishing
- This version takes also into account a list of comments expressed by D. Tody about the NVO needs

## ■ New documentation and list of evolutions available on CDS resources corner when released

# REST Best practice

- Discussed at last Interop in Strasbourg
  - Roadmap deadline fixed for the November Interop in Garching
  - As we need also implementation use cases, VOSpace 2.0 prototyping will be a good one
  - M. Graham in July : “NASA Earth Science Data Systems Standards Process Guide wants to integrate this work after the recommendation process”

# VOSpace (& iRODS), lessons learned

- Previous implementation (at last AIDA meeting)
  - VOSpace 1.1
    - Apache : Tomcat 5.5.27 & Axis2 & Rampart 1.3
    - IVOA VOSpace 1.1 RC3
    - H2 (pure Java) database for the metadata storage
  - iRODS (integrated Rule-based Data System)
    - Developed by SDSC - DICE, open source, easy to move or resize the resources allocated to the VOSpace
    - Version 2.01
  - 2 servers with 6TB data each, iRODS and VOSpace metadata synchronized on the 2 servers but data replication managed by iRODS

# VOSpace (& iRODS), lessons learned (2)

- Experience in the frame of EuroVO AIDA with ~300,000 FITS images (~ 1 TB)
  - Creation of 6 directories with ~10000 subdirectories each
  - 5 images per subdirectory
  - ~ 3 MB per image
  - The experience generates a small amount of data (text files)

# VOSpace (& iRODS), lessons learned (3)

## ■ VOSpace feeding

- Slow, corrections made (importance of deep tests)
- But performances were still too bad to work efficiently on this experience
- We decided to use directly iRODS without VOSpace 1.1
  - Easy to feed iRODS with TB of data
  - Efficient tools to access the data
    - ▶ In applications through the Jargon API
    - ▶ With Web browsers
    - ▶ Useful (avoids VOSpace interface overhead) to give an access to the native data storage (VOSpace capabilities) if possible
  - Important to be able to manage large datasets with VOSpace 2.0

# VOSpace (& iRODS), ongoing work

## ■ Current implementation

### ■ VOSpace 2.0 prototyping

- VOSpace 1.1 maintained but perhaps not after 2.0 production release

### ■ iRODS 2.1 released in July 2009 (upgrade highly recommended by the iRODS staff)

### ■ All the data and metadata are now synchronized through DRBD, the content of the 2 servers is now identical with a mechanism of heartbeat

- Easy to maintain a high availability
- Focus the work on the VOSpace 2.0 implementation



# VOSpace (& iRODS), collaboration & references

## ■ Collaborations

- Discussions started with Dave Morris for 2 use cases in UK about the use of VOSpace/iRODS, will be planned during the hackathon
- Thanks to Jean-Yves Nief (IN2P3) for the useful discussions and informations about iRODS

## ■ References

- iRODS main page illustration :

[https://www.irods.org/index.php/IRODS:Data\\_Grids,\\_Digital\\_Libraries,\\_Persistent\\_Archives,\\_and\\_Real-time\\_Data\\_Systems](https://www.irods.org/index.php/IRODS:Data_Grids,_Digital_Libraries,_Persistent_Archives,_and_Real-time_Data_Systems)

- Presentations

- In June 2009 at a CNES workshop about data preservation
- In February 2009 at a joint IN2P3 – DICE iRODS workshop

# Conclusion

- Presentations at the next Interop in Garching
- Visibility of the work in other communities
- Discussions during the hackathon