

Enabling a robust VOspace based on iRODS

André Schaaff, Cyril Pestel

**Observatoire astronomique de Strasbourg
CDS**



Plan

■ Context

- The CDS
- Data and data centres in the astronomical community
- The Virtual Observatory (VO) and its technical challenge
- Some VO standards (... , VOSpace)

■ VOSpace and iRODS

■ Illustrations

■ Conclusion

The CDS

- **Centre de Données astronomiques de Strasbourg (since 1971)**
 - **Team of astronomers, engineers and librarians**
 - **Main services**
 - VizieR is a catalogue (>7 000) access service, large catalogues can have more than 10^9 entries
 - Simbad provides basic data, cross-identifications, bibliography and measurements for (> 4 300 000) astronomical objects outside the solar system
 - Aladin, a sky atlas with interoperability capabilities
 - **Services are widely used by the community (~100 000 hits per day for VizieR)**
- **Main French partner of the Virtual Observatory project**

Astronomical data

■ **Observation** : “detection of a signal, carried out by someone at a particular point and a particular time, with a certain instrument for a particular purpose”, *Carlos Jaschek, “Data in Astronomy”, 1989*

- signal : radio, image, spectrum, ...
- someone : NASA, ESO, ESA, universities and institutes, ...
- particular point : northern/southern hemisphere, ...
- particular time : epoch (J2000, B1950, ...)
- instrument : telescopes (optical, radio), satellites, interferometers, etc.
- purpose : cartography, magnitude, distance from Earth, chemical composition, etc.

Astronomical data (2)

■ From raw data to publications

■ The observations

- Raw data
- Observation registries
- Calibrations and auxiliary data

■ Reduced data

- Catalogues (physical units) (example : VizieR on line service)
- Databases

■ Publications

- Papers
- Documentation, “grey substance” literature, etc.

Data centres

- **Management of data of spatial missions and ground observatories**
 - Data are mainly where the expertise is
 - Massive data processing
 - ...
- **Cooperation is a “tradition” and is easy**
- **Small community**
- **Go further with the concept of Virtual Observatory**

To the Virtual Observatory

- **Access to the digitised sky, using archived and interconnected observations (especially large surveys of the whole sky, observed at different wavelength)**
 - **Inventory of the data available at the international scale**
 - **Coherent set of archives, surveys, services, and reference dictionaries**
 - **Standardized data access modes, Interoperability**
- **Scientific challenges**
 - **Understand the structures of the Universe at a large scale**
 - **Formation and evolution of our Galaxy (and others...)**
 - **Rare object discovery (black matter, extrasolar planets...)**
- **Educative and cultural dimension, outreach**

IVOA

- **International Virtual Observatory Alliance, started in 2000**
- **Consortium of national and transnational Virtual Observatories**
- **Different Working/Interest groups**
 - **Semantics, Grid and Web Services, Data Model, Data Access Layer, VO Query Language, Applications, Theory, ...**
- **2 meetings / year, active mailing lists, ...**
- **Standardisation work**
 - **Notes, Working drafts, Proposed recommendations, ...**
 - **~like W3C**

IVOA (2)



VO technical challenge

- **Long time storage of petabytes of data**
 - **Mostly archives**
 - **High availability (used in interactive services, cross-matching between data, etc.)**
 - **Easy to find**
 - **...**

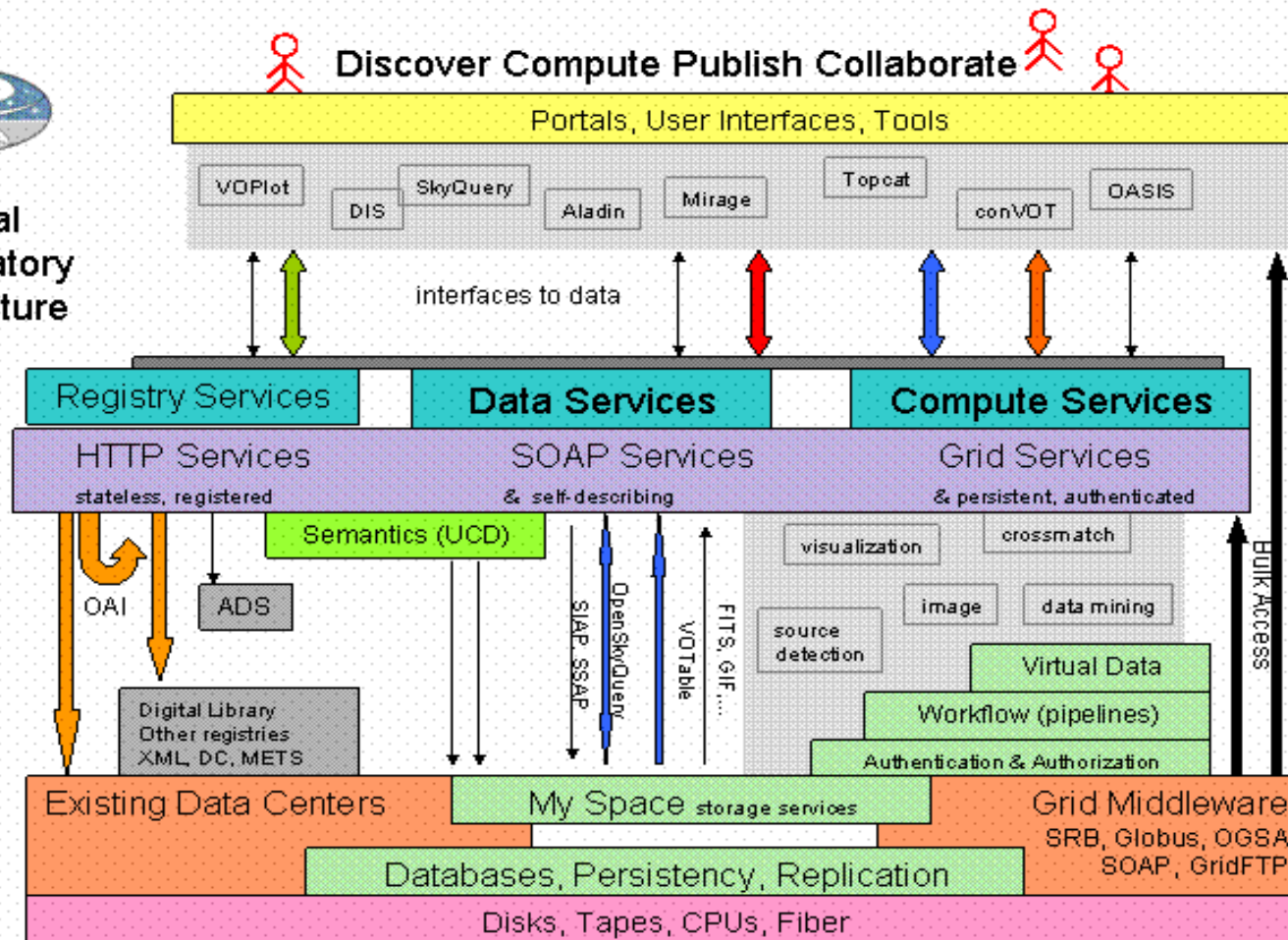
- **Interoperability between astronomical services**

- **Computation power**
 - **Needed for simulations, ...**
 - **Local clusters not sufficient, use of grids like EGEE**

VO Architecture



Virtual
Observatory
Architecture



Some IVOA results

■ VO Registry

- Interfaces to publish, query, and harvest
 - Allows people to publish to a registry by filling a Web form in a Web portal
 - Not unique and centralized, each registry harvests each other to know the new dataset and services added to other VO-registries
 - Compliant with digital library standards (Open Archive Initiative) for metadata harvesting and metadata schema
 - Contains VO resources identified by a universal identifier, starting with ivo://
- In the future, a VO registry may also accept queries in different languages

Some IVOA results (2)

■ The UCDs (Unified Content Descriptors)

- A standardized vocabulary used to describe astronomical quantities and related concepts (in VizieR 1500 UCDs are enough for 100000 columns)
 - "phot.mag;em.IR.K" means a photometric magnitude in infra red between 2 and 3 microns
- No formal representation structure, with syntax and semantics, describing the relationships and dependencies between the words, and it is not possible to perform automated reasoning on UCDs

■ Definition of ontologies...

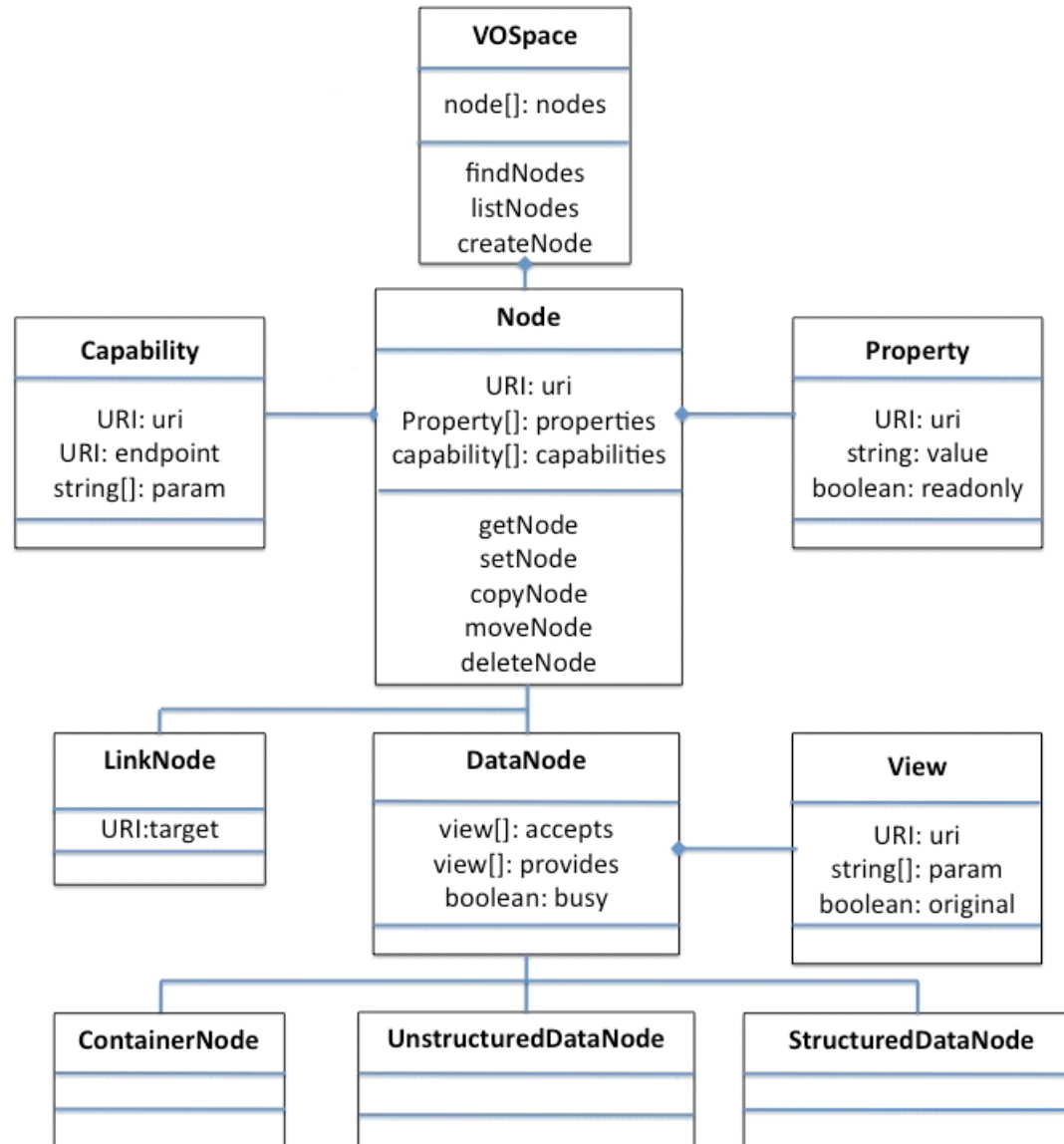
...	
src.ellipticity	Source ellipticity
src.impactParam	Impact parameter
src.morph	Morphology structure
src.morph.param	Morphological parameter
src.orbital	Orbital parameters
src.orbital.eccentricity	Orbit eccentricity
src.orbital.inclination	Orbit inclination
...	

Some IVOA results (3)

- **VOSpace is the IVOA interface to distributed storage. It specifies how VO agents and applications can use network attached data stores to persist and exchange data in a standard way. A VOSpace web service is an access point for a distributed storage network. Through this access point, a client can:**
 - **add or delete data objects**
 - **manipulate metadata for the data objects**
 - **obtain URIs through which the content of the data objects can be accessed**

- ***VOSpace does not define how the data is stored or transferred, only the control messages to gain access.* Thus, the VOSpace interface can readily be added to an existing storage system**

VOSpace schema



Lyon, iRODS workshop 2-5 February 2009

André Schaaff, Cyril Pestel

Enabling a robust VOSpace based on iRODS

VOSpace and iRODS

■ First step : experiment iRODS

- Development of an Aladin (a sky atlas which is also a VO portal) plugin giving an access to the iRODS implementation through Jargon

■ Second step

- Implementation of the VOSpace interface over iRODS
- Use of iRODS in the new CDS portal

■ Third step : creation of VOSpace client tools

- A VOSpace Explorer
- A VOSpace file chooser

■ Last step : release for real life (VOSpace and CDS portal)

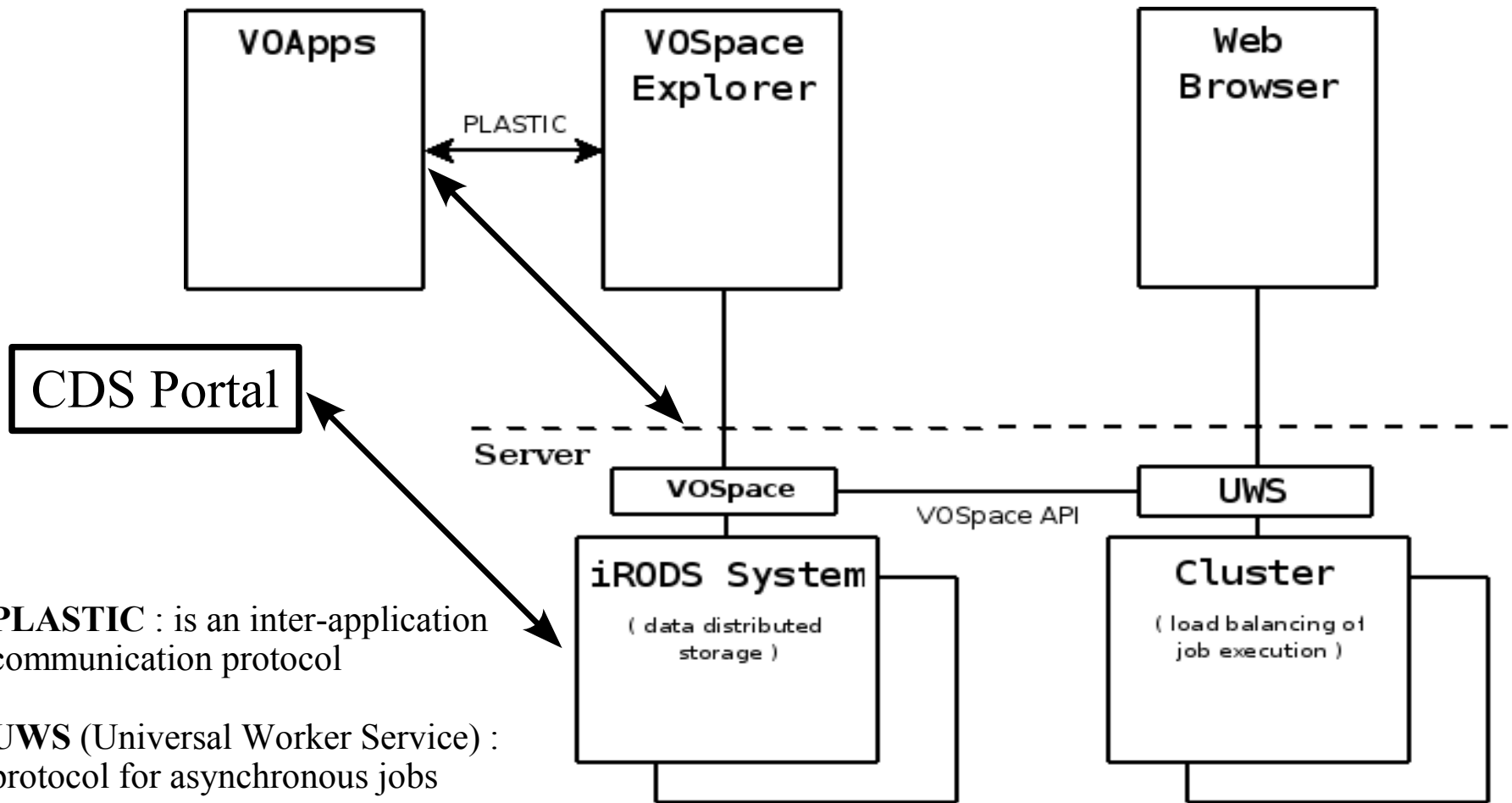
VOSpace and iRODS (2)

- **Use of different iRODS versions**
 - **iRODS 1.0 for the first prototype, iRODS 1.1 for the second and iRODS 2.0.1 for the final release**
 - **Jargon API from 1.* to 2.***

- **VOSpace**
 - **Web Service : Axis2 & Tomcat**

- **iRODS at CDS : 2 quad core servers with 12 TB for the production release → small configuration to evaluate the production needs (not easy to fix “à priori”)**

VOSpace-iRODS architecture



PLASTIC : is an inter-application communication protocol

UWS (Universal Worker Service) : protocol for asynchronous jobs

VOApps : Aladin, Topcat, VizIvo, VOSpec, etc.

Illustration

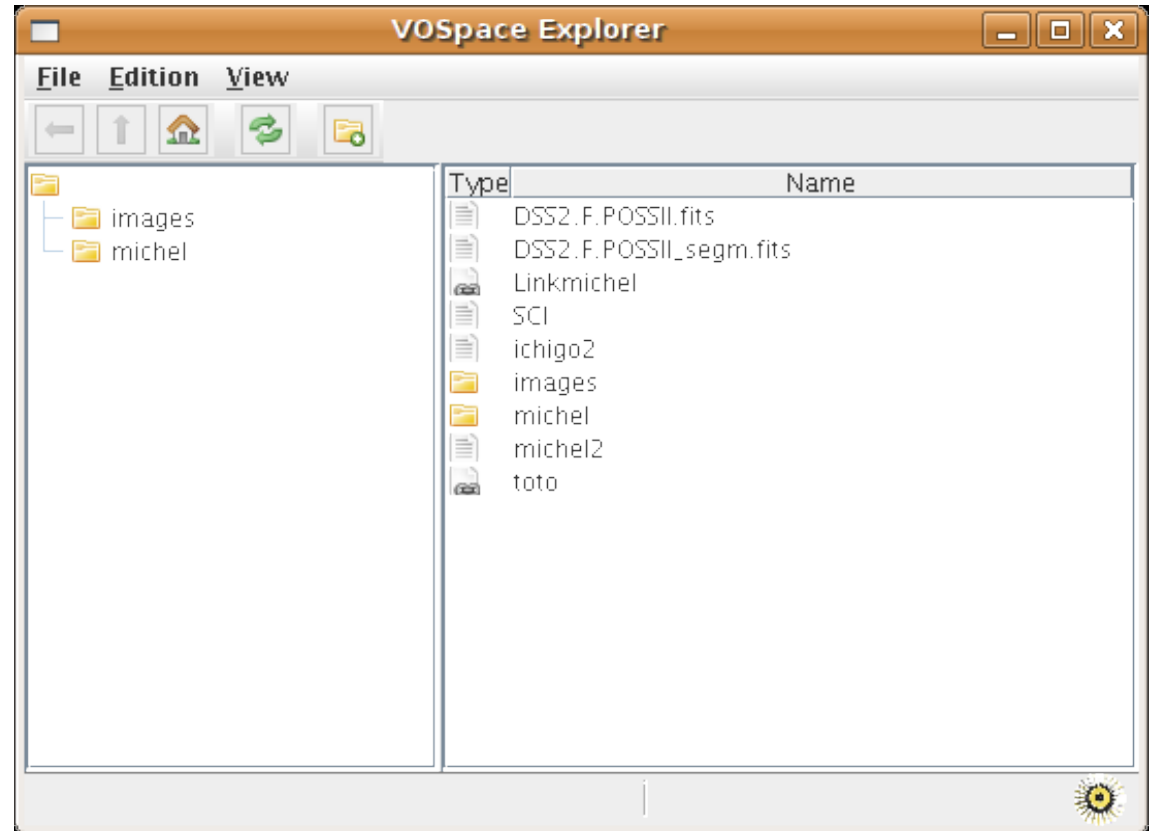
VOSpace tools

iRODS provides a robust storage system
in back of VOSpace

Developed by Cyril Pestel

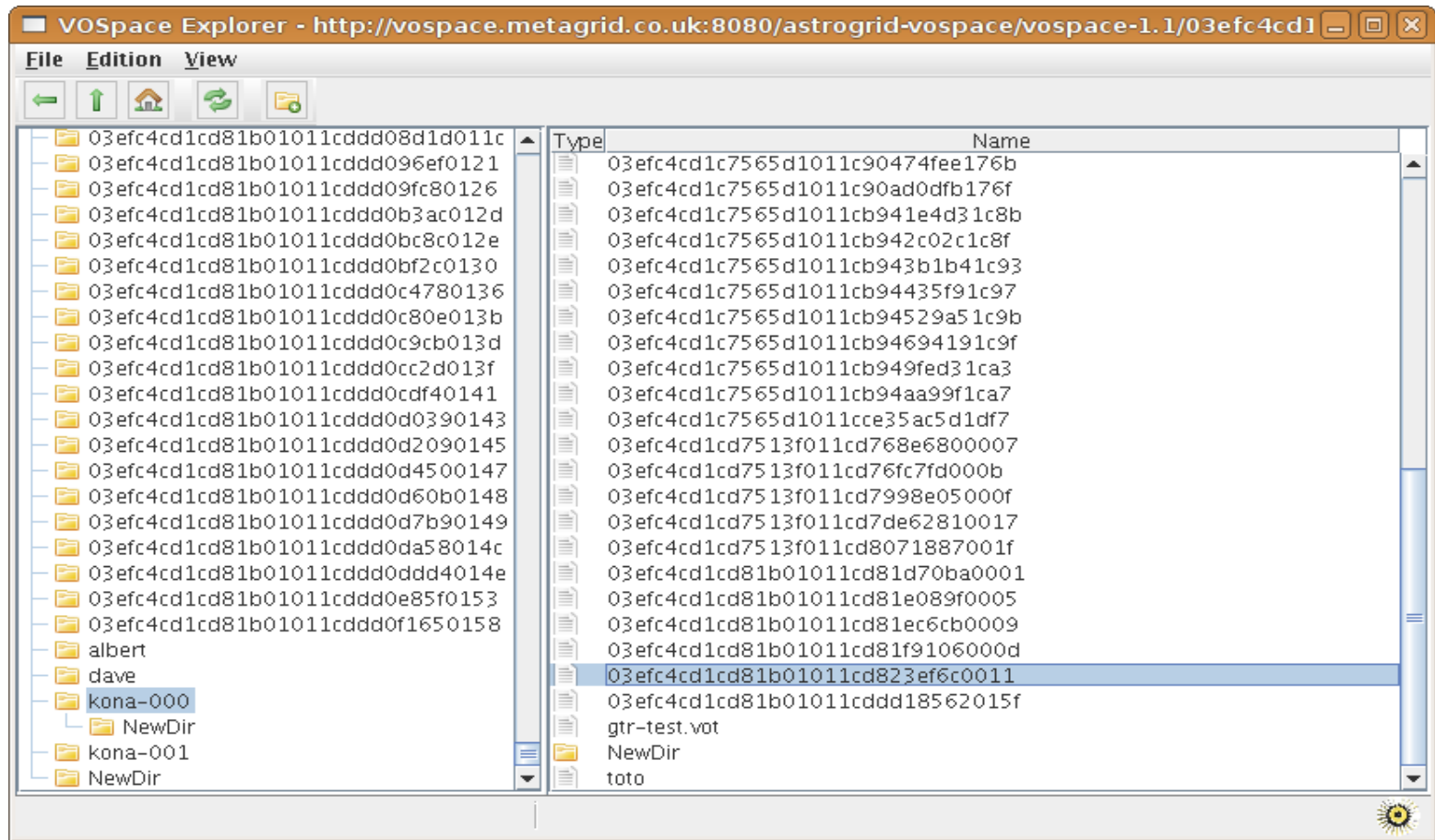
VOSpace Explorer

- **Development of a VOSpace Explorer in Java**
- **If a VO tool supports drag and drop it is possible to interact through this way with the explorer**
- **PLASTIC has been added**



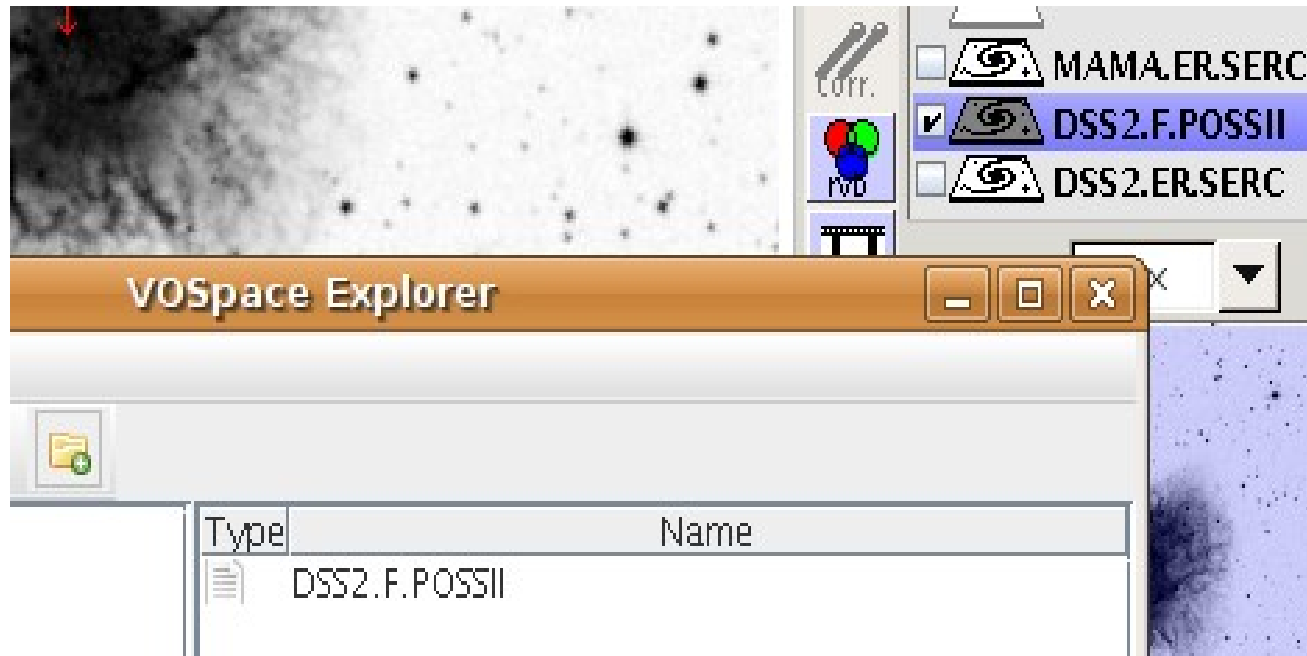
VOSpace Explorer (2)

- Access to other VOSpace, (ex. : Astrogrid's VOSpace, Dave Morris)

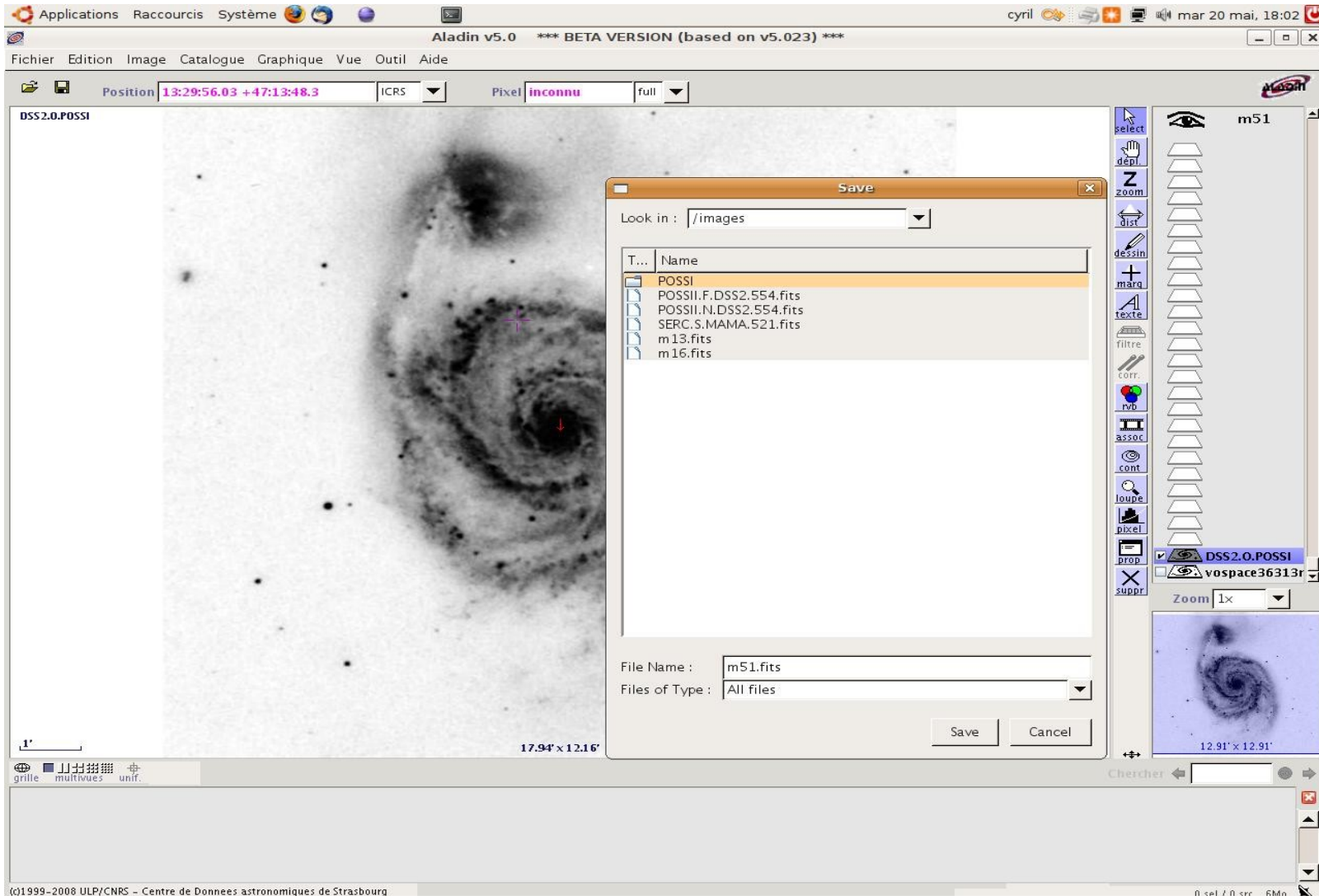


VOSpace Explorer (3)

- PLASTIC use between Aladin and the VOSpace Explorer

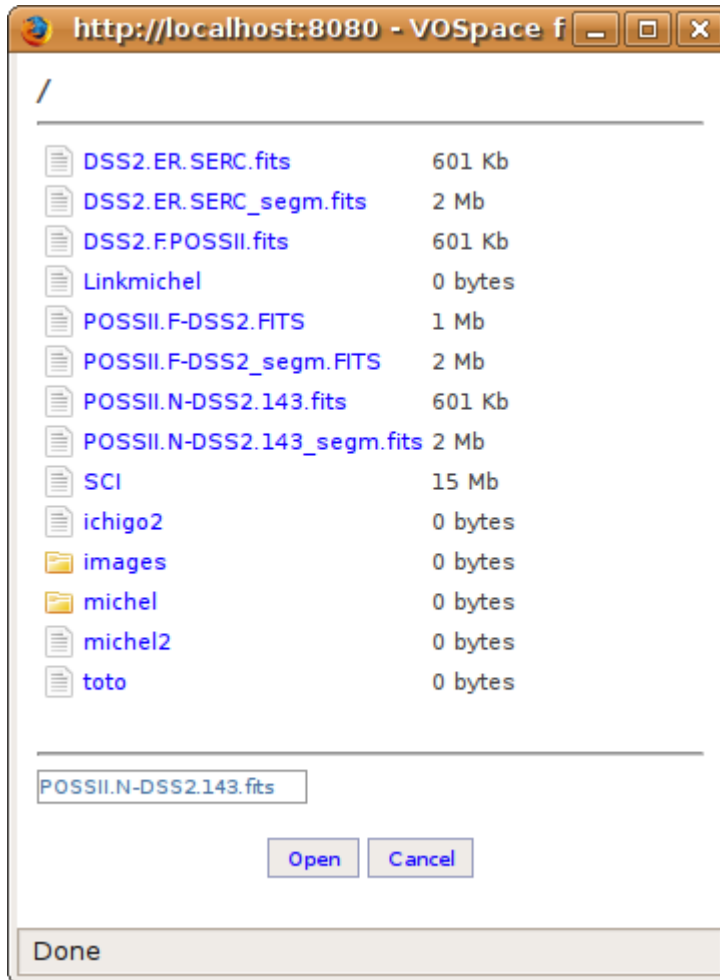


File chooser used in Aladin



FileChooser as a servlet

■ Used in CDS UWS framework



SEE
Submit

Welcome to the **Universal Worker Service** proposed by the **CDS**. It allows the management of asynchronous execution of jobs on a service. The UWS pattern was inspired by AstroGrid's **Common Execution Architecture**. This implementation is based on the **REST paradigm**. You can obtain more details by reading this [document](#) or at the [IVOA Grid&Web services](#) group in [Asynchronous activities proposal](#).

This UWS service proposed the execution of tools coming from **AIDA**. It contains some MatLab (visualization HSV, detectLSB, ...) or C/C++ programs (MARSIAA segmentation, ...). It has been developed in the frame of the MDA project and more recently it is part of the EuroVOTech project in the DS Infrastructure.

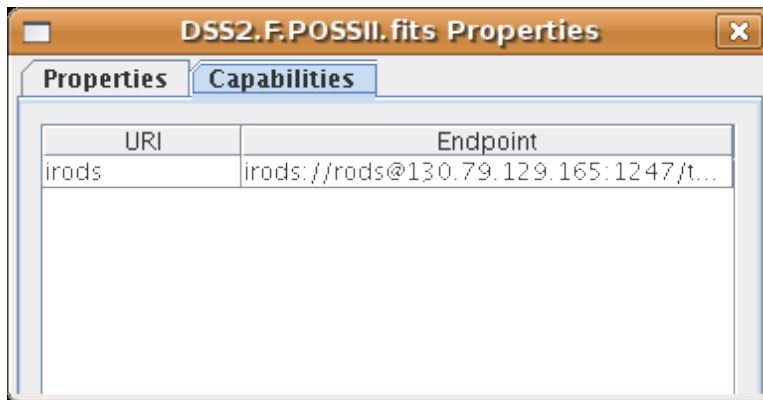
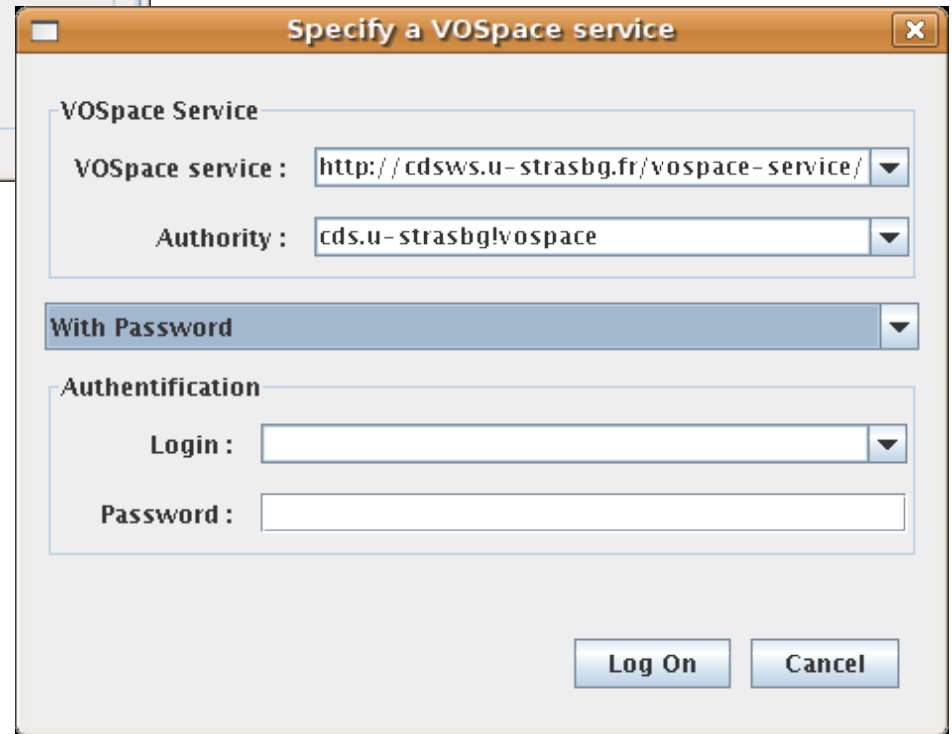
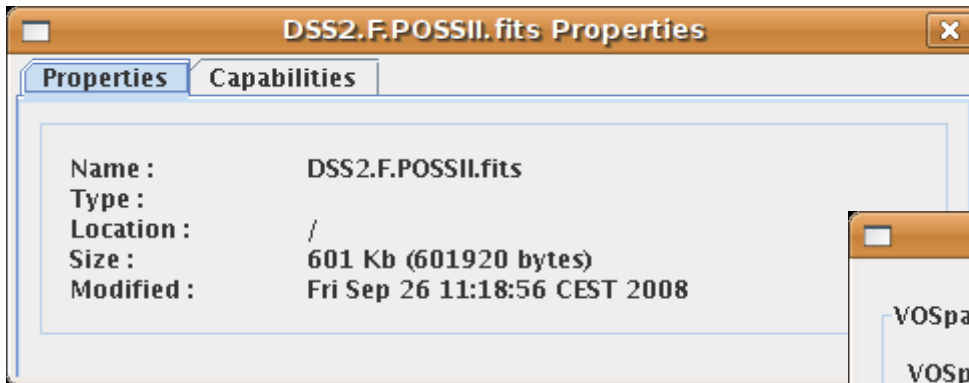
[show/hide form](#)

This is the list of jobs which has been created :

JobID	Phase	Results	Actions
200809241629471214	PENDING	results	
200809241629241013	COMPLETED	results	
20080924162807812	PENDING	results	
20080924162433410	COMPLETED	results	

Tool : segm1 (cluster)
class : 4
image : http://zedd.u-strasbg.fr/aida/vhsvinput/1/3
iter : 3
segm :

Properties, capabilities, security



TLS : with password ok, with certificate soon

Illustration (2)

CDS Portal

iRODS is used to store the user data generated during a session

Developed by Pascal Wassong



CDS Portal

Portal [My data](#)

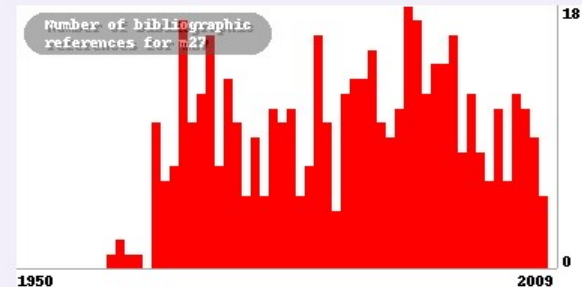
[Login](#) [Register](#) [Preferences](#)

Target:

J2000 position for m27: 19 59 36.340 +22 43 16.09

Object identifiers, measurements and bibliography for m27

- Object type: Planetary Nebula
- Spectral type: DA
- [More SIMBAD data for m27](#)
- [475 bibliographic references](#)
- [14 objects within 2'](#)
- [Display map around m27](#)



Images for m27

- [Display region in Aladin \(Web Start\)](#)

Aladin images

Survey	Band	Wavelength (µm)	Size	Epoch	Resolution	Download
2MASS	K	2.16	8.5' x 17.0'	1999-11-04	0.9" / pixel	FITS
2MASS	H	1.65	8.5' x 17.0'	1999-11-04	0.9" / pixel	FITS
2MASS	J	1.24	8.5' x 17.0'	1999-11-04	0.9" / pixel	FITS
POSSII	F	0.65	12.9' x 12.9'	1992-09-19	1.0" / pixel	FITS JPEG
POSSII	F	0.65	12.9' x 12.9'	1996-07-11	1.0" / pixel	FITS JPEG
POSSII	J	0.49	12.9' x 12.9'	1990-07-24	1.0" / pixel	FITS JPEG
POSSII	N	0.83	12.9' x 12.9'	1992-07-22	1.0" / pixel	FITS JPEG
POSSII	N	0.83	12.9' x 12.9'	1995-07-19	1.0" / pixel	FITS JPEG
POSSII	N	0.83	12.9' x 12.9'	1994-06-15	1.0" / pixel	FITS JPEG
POSSI	O	0.64	12.9' x 12.9'	1951-07-13	1.0" / pixel	FITS JPEG
POSSII	J	0.49	13.0' x 13.0'	1990-07-26	1.0" / pixel	FITS JPEG
POSSII	J	0.49	13.0' x 13.0'	1988-06-14	1.0" / pixel	FITS JPEG
POSSI	E	0.40	14.1' x 14.1'	1951-07-13	1.6" / pixel	FITS JPEG



[Display grayscale image](#)

Catalogues for m27

- 0 catalogues with 'm27' keyword

POSSII	N	0.83	12.9' x 12.9'	1992-07-22	1.0" / pixel	FITS JPEG
POSSII	N	0.83	12.9' x 12.9'	1995-07-19	1.0" / pixel	FITS JPEG
POSSII	N	0.83	12.9' x 12.9'	1994-06-15	1.0" / pixel	FITS JPEG
POSSI	O	0.64	12.9' x 12.9'	1951-07-13	1.0" / pixel	FITS JPEG
POSSII	J	0.49	13.0' x 13.0'	1990-07-26	1.0" / pixel	FITS JPEG
POSSII	J	0.49	13.0' x 13.0'	1988-06-14	1.0" / pixel	FITS JPEG
POSSI	E	0.40	14.1' x 14.1'	1951-07-13	1.6" / pixel	FITS JPEG



[Display grayscale image](#)

Vizie Catalogues for m27

- 0 catalogues with 'm27' keyword
- [73 catalogues around m27](#)

Name	Description	Local density	Wavelength	Popularity	Coverage map
I/297 Query	NOMAD Catalog (Zacharias+ 2005) [ReadMe]	53	optical,IR	85	
I/284 Query	The USNO-B1.0 Catalog (Monet+ 2003) [ReadMe]	51	optical	92	
I/305 Query	The Guide Star Catalog, Version 2.3.2 (GSC2.3) (STScI, 2006) [ReadMe]	49	optical	85	
I/304 Query	Carlsberg Meridian Catalog 14 (CMC14) (CMC, 2006) [ReadMe]	36	optical	78	
II/246 Query	2MASS All-Sky Catalog of Point Sources (Cutri+ 2003) [ReadMe]	34	IR	100	
I/267 Query	The APM-North Catalogue (McMahon+, 2000) [ReadMe]	20	optical	79	
J/A+A/469/1221 Query	Sydney observatory Galactic survey (SOGS) (Fresneau+, 2007) [ReadMe]	16	optical	69	
IX/10A Query	ROSAT All-Sky Bright Source Catalogue (1RXS) (Voges+ 1999) [ReadMe]	11	X-ray	89	
VI/110 Query	Final Merged Log of IUE Observations (NASA-ESA, 2000) [ReadMe]	10	UV	70	
B/hst Query	HST Archived Exposures Catalog (STScI, 2007) [ReadMe]	10	optical	76	



VizieR Search Page

[Tokyo, Japan](#) · [IUCAA, India](#) · [CADK, Canada](#) · [Cambridge, UK](#) · [CFA/Harvard, USA](#) · [UKIRT-Hawaii, USA](#) · [INASAN, Russia](#) · [Beijing Obs., China](#)

I/284 The USNO-B1.0 Catalog (Monet+ 2003) [Similar Catalogues](#) [ReadMe](#)

I.I/284/out The Whole-Sky USNO-B1.0 Catalog of 1,045,913,669 sources (1045913669 rows)
The USNO-B Catalog presents positions, proper motions, magnitudes in blue, red and infrared, as well as star/galaxy estimators for 1,045,913,669 objects derived from 3,648,832,040 separate observations. The data were taken from scans of 7,435 Schmidt plates taken from various sky surveys during the last 50 years. USNO-B1.0 catalog was created by Dave Monet and collaborators at <http://www.nofs.navy.mil/data/fchpix/>. Note that the star/galaxy estimators may be mixed up in dense regions.

Query Setup (usage)

Maximum Entries per table: Output layout: Output Order: + -

Query by Position on the Sky (Adapt Form to use a List of targets)

Target Name (resolved by [Simbad](#)) or Position: Target dimension: Radius or Box size

Position in Sexagesimal, or Decimal °

Output preferences for Position:

	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	r and x,y are the distance to the Target; Position is in the same coordinate system as Target.
Compute	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Sort by	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	

Query by Constraints applied on Columns

Show	Sort	Column	Clear	Constraint	Explain (UCD)
<input checked="" type="checkbox"/>	<input type="radio"/>	USNO-B1.0	<input type="text"/>	(char)	Designation of the object (Note 1) (meta.id:meta.main) (ID MAIN)
<input type="checkbox"/>	<input type="radio"/>	Tycho-2	<input type="text"/>	(char)	Designation in the Tycho-2 Catalog I/259 (meta.id) (ID ALTERNATIVE)
<input checked="" type="checkbox"/>	<input type="radio"/>	RAJ2000	<input type="text"/>	deg	Right Ascension at Eq=J2000, Ep=J2000 (Note 2) (pos.eq.ra:meta.main) (POS EQ RA MAIN)
<input checked="" type="checkbox"/>	<input type="radio"/>	DEJ2000	<input type="text"/>	deg	Declination at Eq=J2000, Ep=J2000 (Note 2) (pos.eq.dec:meta.main) (POS EQ DEC MAIN)
<input checked="" type="checkbox"/>	<input type="radio"/>	e_RAJ2000	<input type="text"/>	mas	Mean error on RAdeg*cos(DEdeg) at Epoch (stat.error:pos.eq.ra) (ERROR)
<input checked="" type="checkbox"/>	<input type="radio"/>	e_DEJ2000	<input type="text"/>	mas	Mean error on DEdeg at Epoch (stat.error:pos.eq.dec) (ERROR)
<input checked="" type="checkbox"/>	<input type="radio"/>	Epoch	<input type="text"/>	yr	Mean epoch of observation (Note 2) (time.epoch:obs) (TIME EPOCH)
<input checked="" type="checkbox"/>	<input type="radio"/>	pmRA	<input type="text"/>	mas/yr	Proper motion in RA (relative to YS4.0) (pos.pm:pos.eq.ra) (POS EQ PMRA)
<input checked="" type="checkbox"/>	<input type="radio"/>	pmDE	<input type="text"/>	mas/yr	Proper motion in DE (relative to YS4.0) (pos.pm:pos.eq.dec) (POS EQ PMDEC)
<input type="checkbox"/>	<input type="radio"/>	muPr	<input type="text" value="0.1"/>	0.1	(n) Total Proper Motion probability (Note 7) (stat.probability) (STAT PROBABILITY)
<input type="checkbox"/>	<input type="radio"/>	e_pmRA	<input type="text"/>	mas/yr	Mean error on pmRA (stat.error:pos.pm:pos.eq.ra) (ERROR)



Store VizieR data to CDS Portal

Target: m27
Catalogue: I/284/out
Radius:

Filename:
Comment:

Select account to use to save data:

Anonymous account
 Login
Username:
Password:
 Remember me



Show my data at the CDS Portal

[Portal](#)
[My data](#)
[Return to VizieR](#)

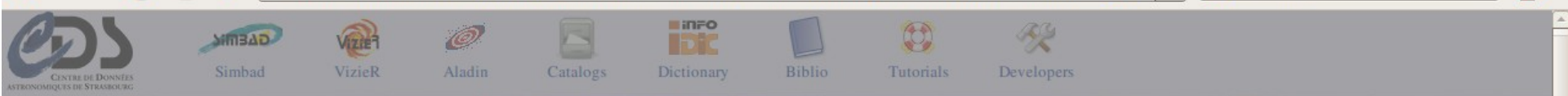
[Login](#)
[Register](#)
[Preferences](#)

Lists of sources

[Query Simbad](#)
[Upload](#)
[Delete](#)

Selected	Target	Catalogue	File	Creation date	Comment	Origin	Nb rows
<input type="checkbox"/>	m27	I/284/out	m27-I_284_out-090128-bis	mer 28 jan 2009 13:09:34 CET	<No comment>	VizieR	3
<input type="checkbox"/>	m27	I/284/out	m27-I_284_out-090128	mer 28 jan 2009 11:29:10 CET	<No comment>	VizieR	3

Page 1 of 1
 Displaying 1 - 2 of 2



Show my data at the CDS Portal

Return to VizieR

Lists of sources

Query Simbad Upload Delete

Selected	Target	Catalogue	File	Creation date	Comment	Origin	Nb rows
<input type="checkbox"/>	m27	I/284/out	m27-I_284_out-090128-bis		ent>	VizieR	3
<input type="checkbox"/>	m27	I/284/out	m27-I_284_out-090128		ent>	VizieR	3

Page 1 of 1

Displaying 1 - 2 of 2

Upload a VOTable

VOTable:

Target: (optional)

Comment: (optional)

[Close this window](#)

Conclusion of this work

- **iRODS is easy to implement and seems to be a good solution to ensure the robustness of a VOSpace**
- **This work is done in the frame of VOTECH project ending in June 2009, tools are available and maintenance will continue for the tools used at CDS (VOSpace-iRODS framework, VOSpace Explorer, CDS Portal)**
- **On iRODS Wiki : <http://www.irods.org/index.php/VOSpace>**
- **On DICE pages :**
http://www.diceresearch.org/DICE_Site/Uses/Entries/2008/11/5_iRODS_Opens_Virtual_Vistas_for_Astronomy.html
- **IVOA wiki : <http://www.ivoa.net>**
- **CDS website : <http://cdsweb.u-strasbg.fr>**