



Observatoire astronomique
de Strasbourg

TAPVizieR - interop Sao Paulo 2012

G. Landais, F. Ochsenbein

VizieR TAP implementation feedback



Observatoire astronomique
de Strasbourg

Regular VizieR contents :

- 10,000 catalogues
- 20,000 tables
- 300,000 columns
- Data volumes
 - ~60Gb in database (without large catalogues)
 - ~1Tb of compressed binary files (outside DBMS)





Observatoire astronomique
de Strasbourg

The technology used

Storage system	PostgreSQL database (3.5Tb)
Position indexation	H3C (custom PostgreSQL lib + NASA HEALPix lib)
Parser/ADQL translator	Java library (G.Mantelet (CDS))
TAP	Java library (G.Mantelet (CDS))
Convert coordinate system	AS4 (F.Ochsenbein)

New database dedicated to TAP for VizieR :

- Database size : ~3.5Tb
- XML output of the (*reduced*) TAP schema : ~ 3.5 Mb
 - FULL TAP schema ~86 Mb



Observatoire astronomique
de Strasbourg

TAPVizieR - interop Sao Paulo 2012

The TAP implementation

Implementation progress in VizieR TAP beta version

- Done
- Partially done
- To do

ADQL user-friendly web page : <http://tapvizier.u-strasbg.fr/adql/>

The TAP service entry point : <http://tapvizier.u-strasbg.fr/TAPVizieR/tap/>

- Creation of a dedicated PostgreSQL database
- Synchronization with the VizieR original database
- Simple Web Page
- Homogenization of the coordinate system in ICRS
add computed columns in tables
- The TAP interface is available
with some adjustments
- Upload



Observatoire astronomique de Strasbourg

TAPVizieR - interop Sao Paulo 2012

The TAP implementation

The TAPVizieR simple web page

First beta version: some explanations about the implementation of VizieR is available [here](#)

The TAPVizieR service provides VizieR tables using the ADQL (a SQL extension in Astronomy).

Type your ADQL Query in the bottom area or try an example [] or use the VizieR capabilities to construct your ADQL query []

Search tables [2mass] Go

Search by catalog, author's name, word(s) from title, position (resolved by [Sesame](#)), ...
e.g.: Veron, 2Mass, redshift, M31...
Note: The VizieR capability takes advantage of METAdata (described [here](#))

all | by wavelength | **by mission** | by astronomy | X

catalogues	description	tables
II/241	2MASS Catalog Intermediate Data Release (IPAC/UMass, 2000) wavelength: IR;	c2massi (162213354 rows) (positions) 2MASS Point Source Catalogue, 2MASS 2000 Second Incremental Release
II/246	2MASS All-Sky Catalog of Point Sources (Cutri+ 2003) wavelength: IR;	c2mass (470992970 rows) (positions) The Point Source catalogue of 470,992,970 sources: % Additions to 2MASS Intro (last modif: 27-Nov-2007) %----- Please f/bt vaWtock[http://www.ipac.caltech.edu/2mass/releases/allsky/doc/sec1_8b.html] (acknowledge the usage of the 2MASS All-Sky Survey); see also the W/word[http://www.ipac.caltech.edu/2mass/](2MASS- /frum/Vphase>3{fg/red3}quad Note that the magnitudes in red correspond to low quality (upper limits or very poor photometry)]/h
II/281	2MASS 6X Point Source Working Database / Catalog (Cutri+ 2005) wavelength: IR;	c2mass6x (24023702 rows) (positions) Sample of 2MASS-6X
VII/233	The 2MASS Extended sources (IPAC/UMass, 2003-2006) wavelength: IR ; astronomy: Galaxies ;	c7233xsc (1547599 rows) (positions) *The 2MASS Extended Catalog (2MASX)

Showing 1 to 4 of 4 entries

Selected tables

* You can not make query on more than two tables.
* Selected tables are automatically stored locally.

catalog	table
<input checked="" type="checkbox"/> II/311	wise (563921564 rows)

construct your query

columns and constraints
sky area
unit change of coordinates

Max records all limit 100 create query

```

1 -- output format : csv
2 SELECT TOP 100 wise.WISE, wise.RA32000, wise.DE32000, wise.eeMaj, wise.eeMin, wise.WImag, wise.e_WImag,
3 wise.W2mag, wise.e_W2mag, wise.W3mag, wise.e_W3mag, wise.W4mag, wise.e_W4mag, wise.Jmag, wise.Hmag,
4 wise.Kmag, wise.ccf, wise.ex, wise.var, wise.d2M
5 FROM wise

```

Query name wise Output format csv Run Quit View Reset

List of your TAP queries Refresh Abort Destroy Properties

name	phase	start	destruction	results
c1239hip_main	COMPLETED	Fri Sep 07 16:36:30 CEST 2012		result (csv)
c1239hip_main	COMPLETED	Fri Sep 07 16:53:33 CEST 2012		result (csv)
c2mass	COMPLETED	Mon Sep 10 16:34:45 CEST 2012		result (csv)
c2mass	COMPLETED	Mon Sep 10 16:35:13 CEST 2012		result (csv)
c2mass-hip	COMPLETED	Mon Sep 10 16:57:20 CEST 2012		result (csv)
c2mass-usnob1	COMPLETED	Fri Sep 07 17:04:02 CEST 2012		result (csv)
sds8	COMPLETED	Mon Sep 10 16:44:02 CEST 2012		result (csv)
sds8_cache	COMPLETED	Mon Sep 10 16:42:07 CEST 2012		result (csv)
tyc2	COMPLETED	Mon Sep 10 15:45:06 CEST 2012		result (ascii)
usnob1-2mass	COMPLETED	Fri Sep 07 17:28:29 CEST 2012		result (ascii)

Showing 1 to 10 of 10 entries

Search Tables using the VizieR METAdata

the list resulting of the search process

The ADQL text area

The Results of your queries (using asynchronous mode)



Observatoire astronomique
de Strasbourg

Main issues and questions during implementation

- TAP Scalability
- Coordinate systems in ADQL
- Naming tables and columns
 - User-readable vs native DB writing
- Additions
 - Individual table schema details
 - Coordinate conversions
 - Searching capabilities
 - Debugging



Observatoire astronomique
de Strasbourg

TAPVizieR - interop Sao Paulo 2012

The XML TAPSchema output

Scalability of the XML TAP schema for VizieR

- We deal with 20 000 tables, 300 000 columns !
- Volumetry needed :
 - Entire description containing tables+columns (→ TAP compliant) : 86Mb
 - without columns description : 3.5 Mb
- Most parsers will choke on too large a schema
- Proposed solution :
 - Provide table descriptions without columns details
 - Additional URL to retrieve column details for a table
 - e.g. <http://tapvizier.u-strasbg.fr/TAPVizieR/tap/tables/tablename>



Observatoire astronomique
de Strasbourg

Managing coordinate systems

- Ambiguities between the parameter used in the function and the « real » database coordinate system

```
SELECT .....
```

```
FROM tyc2
```

```
WHERE 1=CONTAINS(POINT("FK4", ra_icrs_,de_icrs_), CIRCLE("ICRS", 0,0,2/60.)
```

TAP - compliant

- The columns (ra_icrs_,de_icrs_) are in « FK4 »
- A change of coordinate system is required by the CONTAINS function because « FK4 » != « ICRS »

TAPVizieR

- The columns (ra_icrs_,de_icrs_) are in « ICRS »
- The coordinate system « FK4 » is changed in the coordinate system defined by METAdata: → « ICRS »

Note : The Crossmatch case

if columns of tables are in an other coordinate system, TAPVizieR will make a change of coordinate system



Observatoire astronomique
de Strasbourg

Table and column names mapping

VizieR uses

- User-friendly (logical) names : II/259/tyc2, VII/248, RA(ICRS), DE(ICRS)
- Internal DB names : tyc2, c7248vv06, ra_icrs_, de_icrs_

name	DBname
II/259/tyc2	ty2

name	DBname
RA(ICRS)	ra_icrs_
DE(ICRS)	de_icrs_

```
SELECT ra_icrs_,de_icrs_ FROM tyc2 ;
```

At present, VizieR TAP exposes the internal DB table and column names.

Better to use logical names to identify tables and columns ?



Observatoire astronomique
de Strasbourg

Some additions available for VizieR : **(not described in /capabilities)**

- A search url : </tap/search?query=word>
 - take advantage of VizieR METAdata (in particular position)
- A debug url : </tap/syntax?request=doQuery&query=SELECT....>

<TAP>

<INFO>generated by TAPVizieR</INFO>

<INFO name="aqdl" value="SELECT cmc14,raj2000,dej2000 FROM cmc14

WHERE 1=CONTAINS(POINT('FK4',raj2000,dej2000), CIRCLE('ICRS', 279, 38, 5/60.))"/>

<INFO name="warning" value="Wrong coordinate system ADQL says FK4 and Database says FK5[1]"/>

<INFO name="warning" value="Wrong coordinate system ADQL says FK4 and Database says FK5[1]"/>

<INFO name="warning" value="Change the Coordinate system of 'FK4' for FK5[1]"/>

</TAP>