



Observatoire astronomique
de Strasbourg

TAPVizieR - interop Sao Paulo 2012

G. Landais, F. Ochsenbein

TAP improvements from VizieR experience



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

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

TAPVizieR - interop Sao Paulo 2012

The XML TAPSchema output

The XML TAP schema returned by TAPVizieR

→ Filter XML output for a single table

→ The same XML schema used to describe columns or tables:

```
<tableset xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"  
xmlns:vod="http://www.ivoa.net/xml/VODDataService/v1.1" xsi:type="vod:TableSet">
```

→ A new URL dedicated to columns description

→ **not (yet?) a TAP standard**

→ How to write the URL for columns description ?

→ currently : <http://tapvizier.u-strasbg.fr/TAPVizieR/tap/columns?query=tablename>

→ proposed : *(Markus Demleitner)*

<http://tapvizier.u-strasbg.fr/TAPVizieR/tap/tables/tablename>



Observatoire astronomique
de Strasbourg

Managing coordinate system

- 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

TAPVizieR - interop Sao Paulo 2012

The coordinate systems

- Function to make a change of coordinate system in **OUTPUT**
 - To provide the same capability in TAPVizieR than in VizieR
 - No standard in ADQL

```
SELECT POINT("FK4", ra_icrs_, de_icrs_)
```

TAPVizieR

- « FK4 » in **OUTPUT**
- Enable change of coordinate system

TAP - compliant

- « FK4 » in **INPUT**
- Parameter is ignored ???



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	tyc2

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 ?

Need to homogenize TAPVizieR with using logical name (*M.Taylor, M.Allen*)



Observatoire astronomique
de Strasbourg

TAPVizieR - interop Sao Paulo 2012

Naming the tables & columns

2 different ways:

→ the application accepts logical name :

→ Using double quotes : **SELECT "RA(ICRS)", "DE(ICRS)" FROM "II/259/tyc2"**

→ keep a unique identifier homogenized with VizieR, but uses a tedious ADQL query

→ By changing the tablename (and the tablename only)

→ add information in the TAP XML schema :

→ Using the <DESCRIPTION> tag

→ With adding a new tag <dbname>

Return a logical name for table only

→ Replace '/' by '_' (*M.Taylor*) : example I_239_hip_main

→ no ambiguity with tables but impossible with columns
(because the case is sensitive in VizieR)

→ Cut the name by separating the catalog name of the table name

example : I_239.hip_main (*M.Demleitner*)

→ Problem : the '.' supposes the existence of a schema 'I_239'
(10,000 catalogs in VizieR=> 10,000 schemas!!)



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>



Observatoire astronomique
de Strasbourg

TAPVizieR - interop Sao Paulo 2012

annexes

Annexes



Observatoire astronomique
de Strasbourg

XML output : add tag <dbname>

```
<?xml version="1.0" encoding="UTF-8"?>
<tableset xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
          xmlns:vod="http://www.ivoa.net/xml/VODDataService/v1.1"
          xsi:type="vod:TableSet">
  <schema>
    <name>viz1</name>
    <table type="base_table">
      <name>II/259/tyc2</name>
      <dbname>tyc2</dbname>
      <description>*The Tycho-2 main catalogue</description>
      <utype></utype>
      .....
      <column std="true">
        <name>RA(ICRS)</name>
        <dbname>RA_ICRS_</dbname>
        <description>Observed Tycho-2 Right Ascension, ICRS</description>
        <unit>deg</unit>
        <utype>                </utype>
        <ucd>pos.eq.ra;meta.main</ucd>
        <dataType xsi:type="vod:TAPType">DOUBLE</dataType>
        <flag>indexed</flag>
        <flag>primary</flag>
      </column>
      .....
    </table>
  </schema>
</tableset>
```



Observatoire astronomique
de Strasbourg

XML output : use <description>

```
<?xml version="1.0" encoding="UTF-8"?>
<tableset xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
          xmlns:vod="http://www.ivoa.net/xml/VODDataService/v1.1"
          xsi:type="vod:TableSet">
  <schema>
    <name>viz1</name>
    <table type="base_table">
      <name>tyc2</name>
      <description>[II/259/tyc/main] *The Tycho-2 main catalogue</description>
      <utype></utype>
      .....
      <column std="true">
        <name>RA_ICRS_</name>
        <description>[RA(ICRS)] Observed Tycho-2 Right Ascension, ICRS</description>
        <unit>deg</unit>
        <utype>                </utype>
        <ucd>pos.eq.ra;meta.main</ucd>
        <dataType xsi:type="vod:TAPType">DOUBLE</dataType>
        <flag>indexed</flag>
        <flag>primary</flag>
      </column>
      .....
    </table>
  </schema>
</tableset>
```