

Les standards de l'OV

« TAP 1.0, Table Access Protocol »

Les standards de l'OV

Journée « L'Observatoire Virtuel pour les débutants »
Vendredi 17 mars 2017

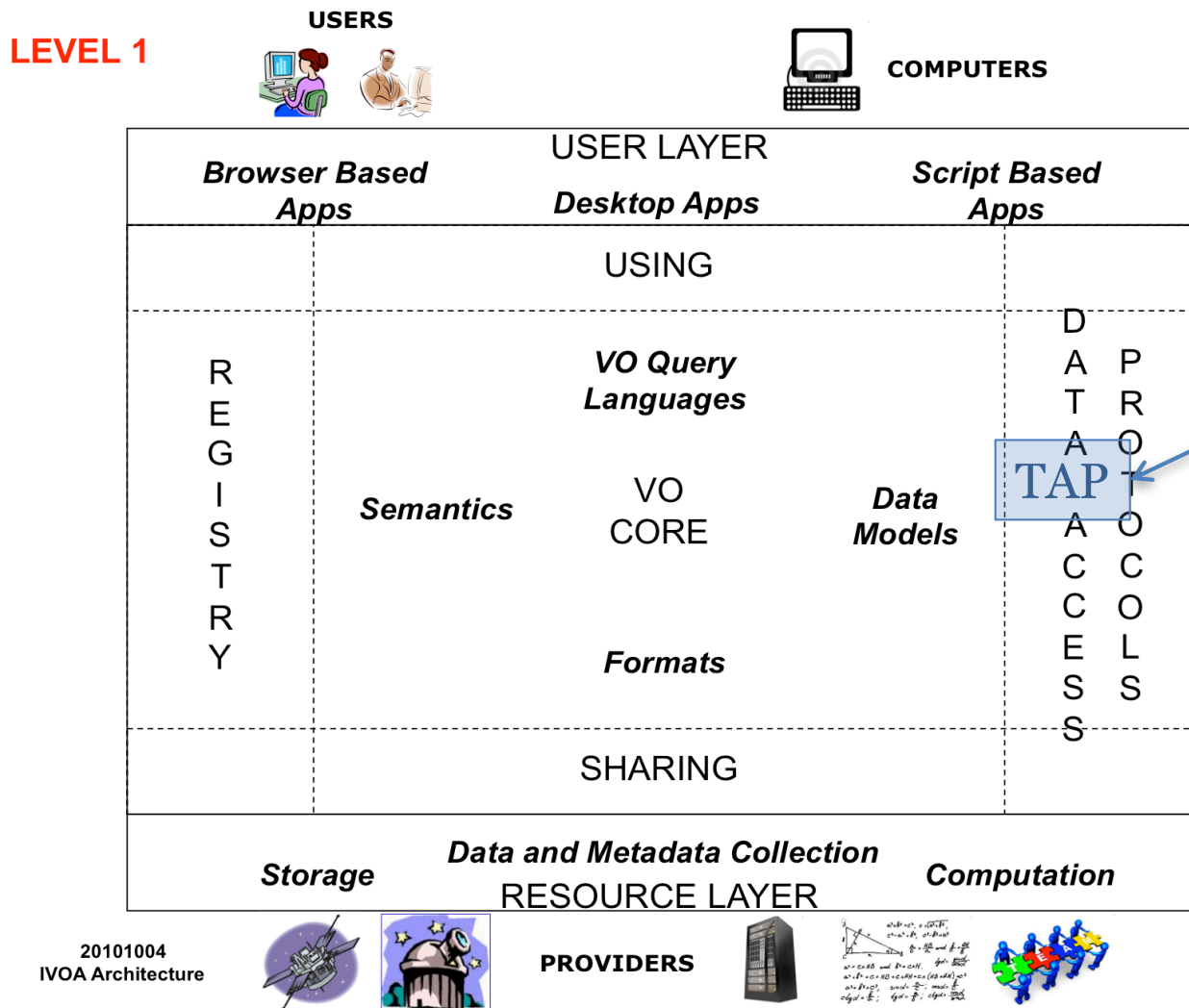


Avec le soutien de l'AS OV France



IAP, 17 mars 2017

TAP dans l'architecture IVOA





Introduction

- Pourquoi s'orienter vers TAP ?
 - Pour offrir une visibilité et un accès (VO) standardisé à ses données
 - On peut également choisir de rendre ses données accessibles via TAP (en mettant en ligne son propre service ou en faisant ingérer ses données par un autre service) afin de profiter des outils de traitements qui acceptent le format en entrée



TAP, c'est quoi ?

- Il existe des protocoles IVOA pour les images (SIAP), les spectres (SSAP), etc. donc spécialisés
- TAP est un protocole de service Web plus générique qui donne accès à des collections de données tabulaires issues de bases de données
- Basé sur REST (Representational State Transfer)



TAP, c'est quoi ? (2)

- Un service TAP accepte des requêtes concernant les données disponibles et, en réponse, retourne une table
 - En proposant différents langages pour soumettre une requête: **ADQL** à minima, SQL, etc.)
 - En proposant différents formats pour la table retournée: **VOTable** à minima, JSON, etc.
- Remarque:
 - Un service TAP ne proposera pas forcément toutes les données réellement présentes dans la base de données
 - Un utilisateur pourra charger temporairement ses propres données, ce qui lui permettra par exemple de les croiser avec les données du service



Un service TAP accepte donc des requêtes...

- ... de 3 types
 - « Data Queries »
 - Accès au « contenu astronomique » via l'un des langages de requête supporté par le service
 - « Metadata Queries »
 - Mêmes langages de requête que précédemment mais pour des tables particulières (cf. plus loin TAP_SCHEMA)
 - Pour récupérer par exemple des métadonnées concernant la disponibilité du service TAP (VOSI, VO Support Interfaces)



Un langage de requête en particulier...

- Le seul qui soit obligatoire: ADQL
 - Permet d'effectuer des requêtes sur une ou plusieurs tables (y compris de métadonnées) du service TAP
 - En général, on effectue une requête pour accéder aux métadonnées pour connaître les noms des tables et des colonnes puis on peut effectuer des requêtes sur les données
 - Les requêtes sont dépendantes du service sauf si l'on reste très générique
 - La plupart des caractéristiques de SQL
 - Support de fonctions de recherche spatiale
 - Etc.



Exemple de requête ADQL dans TAP VizieR

--xmatch_wise_hipparcos

-- xmatch between wise (WISE) and I/239/hip_main (hipparcos) around the lmc center within 1 deg

```
SELECT TOP 100 "II/311/wise".RAJ2000, "II/311/wise".DEJ2000, "I/239/hip_main".HIP, "I/239/hip_main"."RA(ICRS)",  
"I/239/hip_main"."DE(ICRS)", "II/311/wise".WISE
```

```
FROM "II/311/wise", "I/239/hip_main"
```

```
WHERE 1=CONTAINS(POINT('ICRS', "I/239/hip_main"."RA(ICRS)", "I/239/hip_main"."DE(ICRS)"), CIRCLE('ICRS',  
80.89417, -69.75611, 1.))
```

```
AND 1=CONTAINS(POINT('ICRS', "I/239/hip_main"."RA(ICRS)", "I/239/hip_main"."DE(ICRS)"),  
CIRCLE('ICRS', "II/311/wise".RAJ2000, "II/311/wise".DEJ2000, 10/3600.))
```

```
RAJ2000,DEJ2000,HIP,RA(ICRS),DE(ICRS),WISE
```

```
81.660548,-70.063505,25452,81.66074576,-70.06368899,"J052638.53-700348.6"  
82.708637,-69.524824,25822,82.70866058,-69.52483500,"J053050.07-693129.3"  
81.409941,-69.595237,25369,81.41017980,-69.59551844,"J052538.38-693542.8"  
79.443100,-69.849222,24694,79.44322532,-69.84923159,"J051746.34-695057.1"  
78.378178,-69.539904,24347,78.37823715,-69.53990311,"J051330.76-693223.6"  
81.645502,-68.868693,25448,81.64533902,-68.86870606,"J052634.92-685207.2"
```




Comment décrit-on un service TAP ?

- C'est plutôt souple, il n'est absolument pas obligatoire de décrire toutes les tables de la base de données, de dévoiler le schéma réel de celle-ci
- Le service doit fournir le « TAP_SCHEMA » qui décrit les tables exposées dans le service
 - TAP_SCHEMA = Ensemble de 5 tables qui contiennent une description des schémas, tables et colonnes. Elles sont stockées comme des tables de données et sont donc accessibles via les mêmes outils (ADQL par exemple)
- La description comporte des métadonnées (nom, type de données et unité) et des données sémantiques
- TAP_SCHEMA contient également une description de la façon dont les tables peuvent être jointées



- Ce que doit/peut proposer un service TAP

```

HTTP POST http://example.com/tap/sync
REQUEST=doQuery
LANG=ADQL
QUERY=SELECT * FROM magnitudes as m where m.r>=10 and m.r<=16
  
```

Feature	Resource	Resource support	Parameters	Parameter support
synchronous query execution	/sync	must		
asynchronous query execution	/async	must		
availability (VOSI)	/availability	should		
capabilities (VOSI)	/capabilities	must	REQUEST=getCapabilities	must
table metadata (VOSI)	/tables	should		
ADQL queries			REQUEST=doQuery LANG=ADQL	must
PQL queries			REQUEST=doQuery LANG=PQL	may
other query languages			REQUEST=doQuery LANG=<other>	may
table upload			UPLOAD	may
VOTable output			FORMAT	must
other formats			FORMAT	should
limiting output			MAXREC	must
logging			RUNID	should



Le résultat d'une requête

- Les formats de sortie conseillés, seul VOTable étant obligatoire

<code>table type</code>	<code>MIME type(s)</code>	<code>short form</code>
<code>VOTable</code>	<code>application/x-votable+xml</code> <code>text/xml</code>	<code>votable</code>
<code>comma separated values</code>	<code>text/csv</code>	<code>csv</code>
<code>tab separated values</code>	<code>text/tab-separated-values</code>	<code>tsv</code>
<code>FITS binary table</code>	<code>application/fits</code>	<code>fits</code>
<code>pretty-printed text</code>	<code>text/plain</code>	<code>text</code>
<code>pretty-printed Web page</code>	<code>text/html</code>	<code>html</code>

Des exemples

Dedicated Clients	Description
Vizier TAP interface http://tapvizier.u-strasbg.fr/adql/ CDS (Fr)	A complete web application dedicated to the Vizier TAP service. A table selector help to locate Vizier tables
Simbad TAP interface http://simbad.u-strasbg.fr/simbad/sim-tap CDS (Fr)	Complete web application dedicated to the SIMBAD TAP service. Query templates are provided.
CADC TAP interface http://www1.cadc-ccda.hia-ih.nrc-cnrc.gc.ca/tap/ CADC (Ca)	Simple page helping to discover the CADC TAP service with various examples.
GAVO TAP interface http://dc.zah.uni-heidelberg.de/system_/tap/run/tap/async GAVO (Ge)	Basic html form to submit ADQL jobs on the GAVO TAP node.

Les services TAP dans le Registry Euro-VO




The screenshot displays the EURO-VO Registry interface. At the top, the logo for EURO-VO AIDA (Astronomical Infrastructure for Data Access) is visible. Below the logo, there are navigation tabs for 'The Euro-VO projects: VOTECH, EuroVO-DCA, EuroVO-AIDA'. A left-hand sidebar contains a list of search resources, with 'Table Access (TAP)' highlighted in green. The main content area shows 'Search Results' for 'Table Access (TAP)'. Three results are listed:

- Two Micron All Sky Survey [TWOMASS_JVO]**
IVOA Identifier: ivo://jvo/twomass [CatalogService] [TableAccess]
[XML | EDIT | CLONE]
The Two Micron All Sky Survey (2MASS) project is designed to close the gap between our current technical capability and our knowledge of the near-infrared sky. In addition to providing a context for the interpretation of results obtained at infrared and other wavelengths, 2MASS is providing direct answers to immediate questions on the large-scale structure of the Milky Way and the Local Universe. The optimal use of the next generation of infrared space missions, such as HST/NICMOS, the Space Infrared Telescope Facility (SIRTF), and the Next Generation Space Telescope (NGST), as well as powerful ground-based facilities, such as Keck I, Keck II, and Gemini, require a new census with vastly improved sensitivity and astrometric accuracy than that previously available. To achieve these goals, 2MASS has uniformly scanned the entire sky in three near-infrared bands to detect and characterize point sources brighter than about 1 mJy in each band, with signal-to-noise ratio (SNR) greater than 10. Published by: JVO on the 2008-04-10T00:00:00.000Z and last updated on the 2014-05-29T03:00:00.000Z
- Subaru Suprime-Cam data service [SUBARU_SUP]**
IVOA Identifier: ivo://jvo/subaru/spcam [CatalogService] [TableAccess] [SimpleImageAccess] [SimpleImageAccess]
[CHECK | CHECK | XML | EDIT | CLONE]
The Subaru Prime Focus Camera (Suprime-Cam) is a mosaic of ten 2048 x 4096 CCDs, located at the prime focus of Subaru Telescope, which covers a 34' x 27' field of view with a pixel scale of 0.20". Published by: JVO on the 2008-04-10T00:00:00.000Z and last updated on the 2014-05-29T03:00:00.000Z
- Subaru HDS Spectrum data service [SubaruHDS]**
IVOA Identifier: ivo://jvo/subaru/hds [CatalogService] [TableAccess] [SimpleSpectralAccess]
[CHECK | XML | EDIT | CLONE]
This data service provides various data related to Subaru HDS. Published by: JVO on the 2008-04-10T00:00:00.000Z and last updated on the 2014-05-29T03:00:00.000Z
- The SWIRE Data Service [SWIRE]**
IVOA Identifier: ivo://jvo/swire [CatalogService] [TableAccess]
[XML | EDIT | CLONE]



Conclusion

- Développer son service TAP sans recourir à une librairie ou à un « framework » existant est tout à fait faisable mais prend du temps
- TAP étant basé sur REST, le développement d'un client TAP est par contre plus simple
- Les présentations à venir vous donneront:
 - Des exemples d'implémentation du standard à la fois côté serveur et client dans des « frameworks » et des librairies
 - L'occasion de voir des exemples de requêtes ADQL



TAP_SCHEMA, un peu de détail

TAP_SCHEMA.schemas

Column name	datatype	
schema_name	varchar	schema name, possibly qualified
description	varchar	brief description of schema
utype	varchar	UTYPE if schema corresponds to a data model

TAP_SCHEMA.tables

Column name	datatype	
schema_name	varchar	the schema name from TAP_SCHEMA.schemas
table_name	varchar	table name as it should be used in queries
table_type	varchar	one of: table, view
description	varchar	brief description of table
utype	varchar	UTYPE if table corresponds to a data model

TAP_SCHEMA, un peu de détail (2)

TAP_SCHEMA.columns

Column name	datatype	
table_name	varchar	table name from TAP_SCHEMA.tables
column_name	varchar	column name
description	varchar	brief description of column
unit	varchar	unit in VO standard format ⁵
ucd	varchar	UCD of column if any
utype	varchar	UTYPE of column if any
datatype	varchar	ADQL datatype as in section 2.5
size	integer	length of variable length datatypes ⁶
principal	integer	a principal column; 1 means true, 0 means false
indexed	integer	an indexed column; 1 means true, 0 means false
std	integer	a standard column; 1 means true, 0 means false

TAP_SCHEMA, un peu de détail (3)

TAP_SCHEMA.keys

Column name	datatype	
key_id	varchar	key identifier from the TAP_SCHEMA.keys
from_column	varchar	key column name in the <from_table>
target_column	varchar	key column in the <target_table>

TAP_SCHEMA.keys_columns

Column name	datatype	
key_id	varchar	unique key identifier
from_table	varchar	fully qualified table name
target_table	varchar	fully qualified table name
description	varchar	description of this key
utype	varchar	utype of this key

Cela fait donc bien 5 tables



Des exemples de services TAP (non exhaustif)

- <http://simbad.u-strasbg.fr/simbad/sim-tap>
- <http://tapvizier.u-strasbg.fr/adql/>
- Etc.

- Pour les outils, cf. les présentations à venir



Documents IVOA

- La recommandation IVOA du 27 mars 2010 (WG Data Access Layer)
 - <http://www.ivoa.net/documents/TAP/>
- Autres
 - Une recommandation « TAPRegExt » du 27 août 2012
 - <http://ivoa.net/documents/TAPRegExt/20120827/index.html>
 - Une note « TAP implementation Notes » du 13 décembre 2013
 - <http://ivoa.net/documents/Notes/TAPNotes/index.html>
 - Une recommandation « Registry Relational Schema » du 27 juin 2014
 - <http://ivoa.net/documents/RegTAP/20140627>