

TAP into TOPCAT

Mark Taylor (GAVO/Bristol)

IVOA Interop
Naples

18 May 2011

Outline

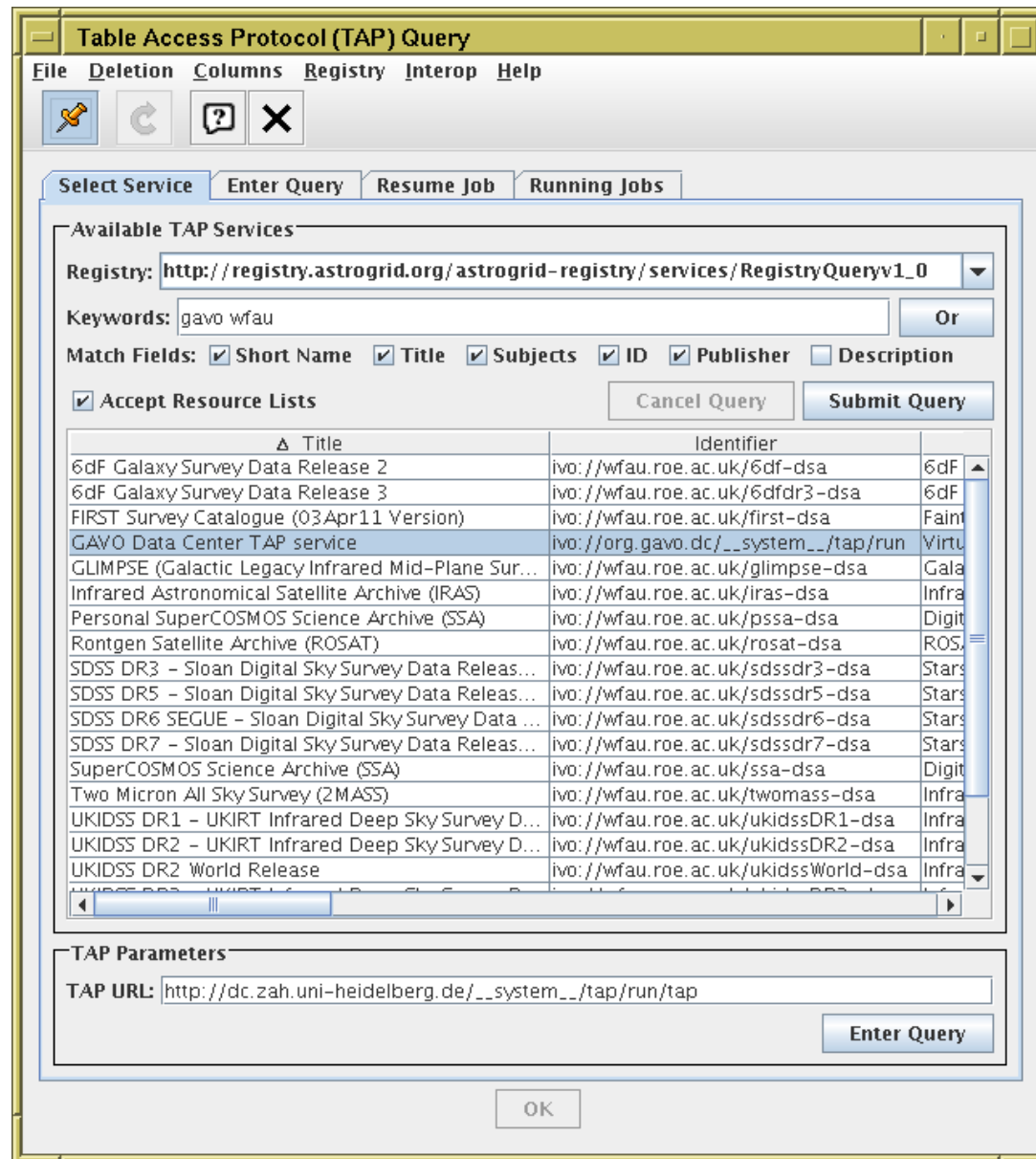
- TOPCAT/STILTS TAP client
- State of services
- TAP issues

TOPCAT TAP Load Dialogue

Normal sequence of actions to load TAP result into TOPCAT:

- TAP service selected by user:
 - ▷ *either* search registry for TAP services and pick one
 - ▷ *or* enter TAP service URL by hand
- `/tables` endpoint queried for table metadata (VODataService)
 - ▷ If present: displays table selector, column metadata, foreign key info
 - ▷ Currently ignores `<schema>/<table>` hierarchy
- `/capabilities` endpoint queried for service capabilities (TAPRegExt WD)
 - ▷ If present: displays QL selector, output limits (adjustable), upload support/limits
- ADQL entered by user
 - ▷ **Examples** menu, entries constructed from supplied table metadata
 - ▷ **Table uploads** supported (from tables already in TOPCAT)
 - ▷ Query Language name/version is selectable (from capabilities list)
 - ▷ Non-ADQL-like languages (PQL) not currently supported (any need?)
- Query submitted by user
 - ▷ Asynchronous: shows submitted job with status (QUEUED→EXECUTING→COMPLETED); loads when complete
 - ▷ Synchronous: loads table directly

TOPCAT TAP Dialogue: Select Service Tab



TOPCAT TAP Dialogue: Enter Query Tab

Table Access Protocol (TAP) Query

File Deletion Columns Registry Interop Help

Select Service **Enter Query** Resume Job Running Jobs

Table Metadata
Service: GAVO Data Center TAP service (59 tables)
Table: **fk6.data** The union of all published parts of FK6

Columns:

Name	DataType	Indexed	Unit	Description
alphaFloat	double	<input checked="" type="checkbox"/>	deg	Main value of right ascension
deltaFloat	double	<input checked="" type="checkbox"/>	deg	Main value of declination
c_x	float	<input type="checkbox"/>		x coordinate of intersection of radius vector and unit sph
c_y	float	<input type="checkbox"/>		y coordinate of intersection of radius vector and unit sph
c_z	float	<input type="checkbox"/>		z coordinate of intersection of radius vector and unit sph
localid	char	<input checked="" type="checkbox"/>		FK6 Number

Foreign Keys:

Target Table	Links	Description	Utype
--------------	-------	-------------	-------

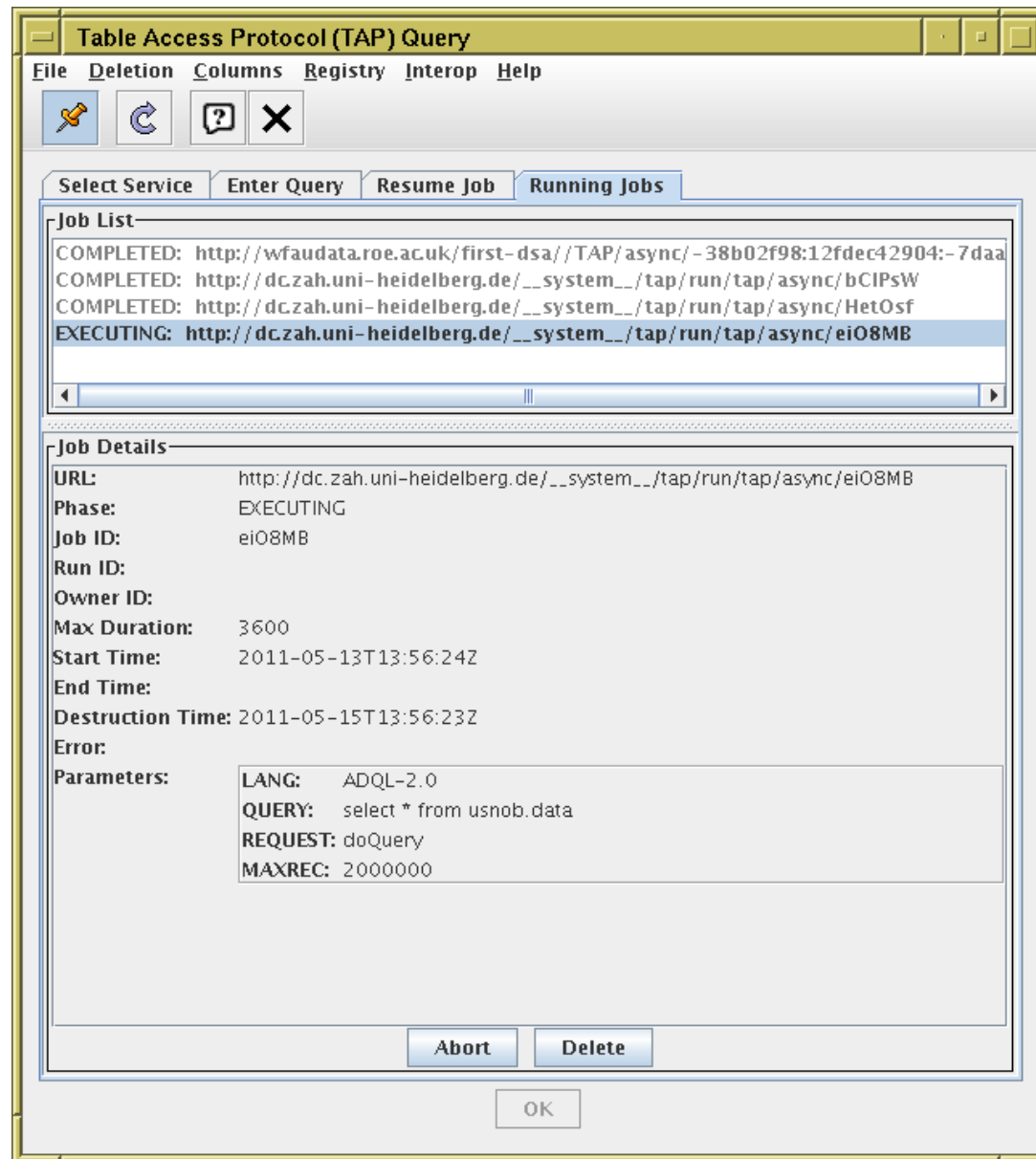
Service Capabilities
Query Language: **ADQL-2.0** Max Rows: **2000 (default)** Uploads: 50Mb

ADQL Text
 Synchronous Examples Clear

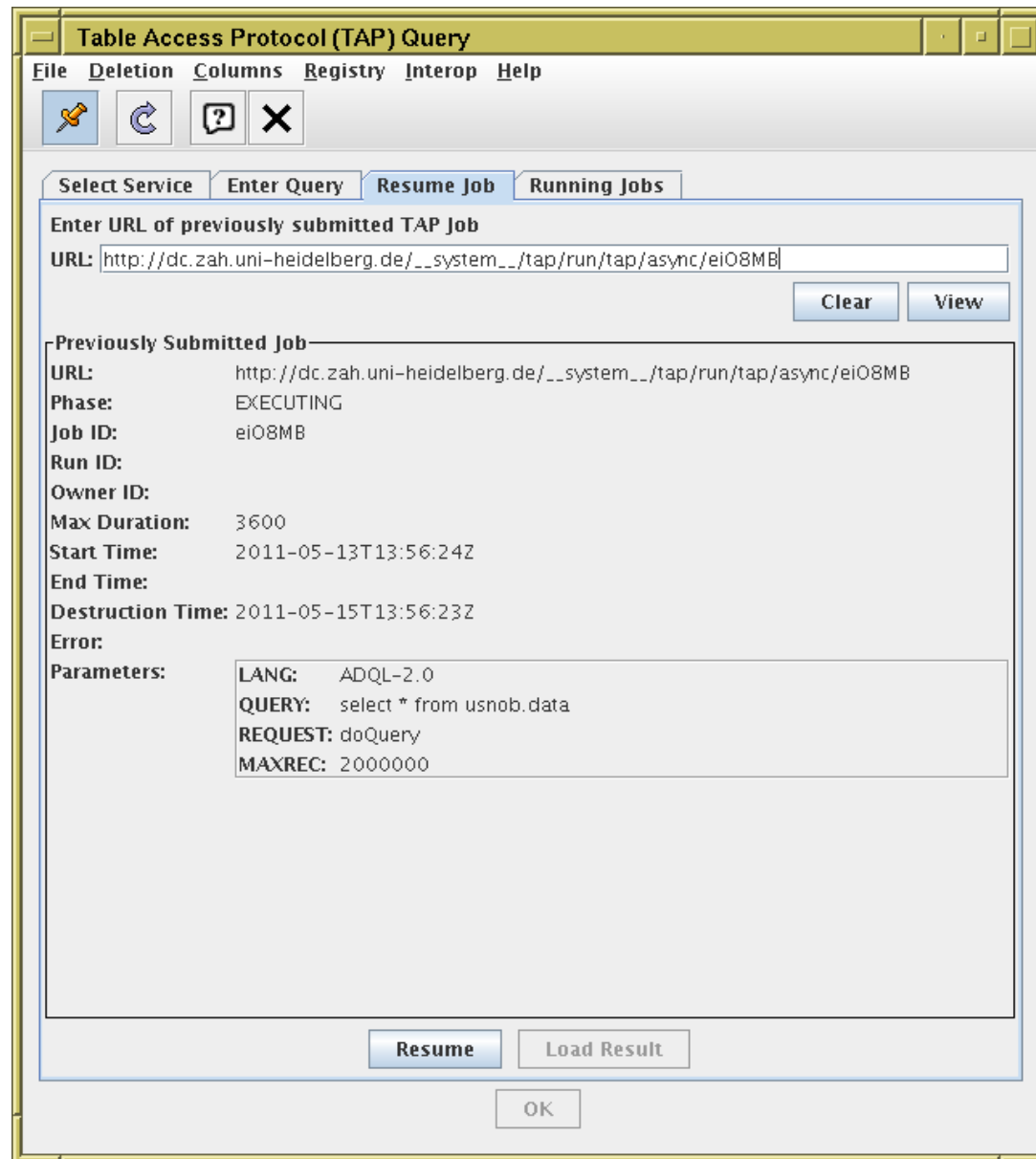
```
SELECT
TOP 1000
localid, alphaFloat, deltaFloat, c_x, c_y, c_z
FROM fk6.data
WHERE
alphaFloat BETWEEN 180 AND 190
AND
deltaFloat BETWEEN -10 AND 10
```

OK

TOPCAT TAP Dialogue: Running Jobs Tab

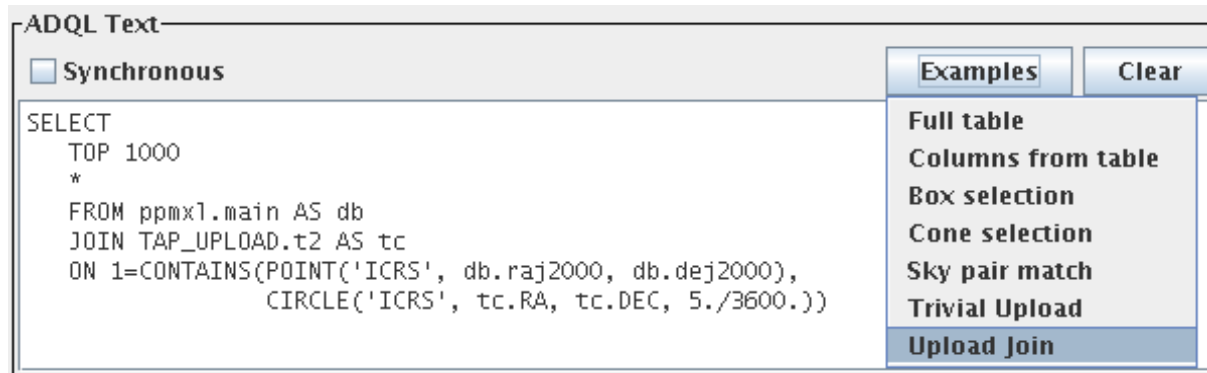


TOPCAT TAP Dialogue: Resume Job Tab



Help with ADQL

- Example queries available from a menu
 - Easier for users than reading the ADQL specification
 - ▷ Can pick an example a bit like what they want to do, and edit it
 - Example text is sensitive to:
 - ▷ Table metadata (available tables and columns)
 - ▷ Table metadata GUI state (which table currently selected by user)
 - ▷ Capabilities metadata (e.g. upload support, language variant)
 - Tries to generate a sensible query based on column UCDs etc
 - Scope for more/better examples (though too many could become confusing)



- Client-side parsing (presubmission ADQL validation etc)
 - Not done yet; will investigate

STILTS TAP Client

STILTS commands provide query submission capabilities:

- `tapquery` to submit a new query
- `tapresume` to resume a previous query
- Features:
 - ▷ Sync or Async submission (`sync=true|false`)
 - ▷ Configurable status polling (`poll=<millis>`)
 - ▷ Configurable UWS job deletion (`delete=finished|never|always`)
 - ▷ Table uploads (`nupload=<count>` `uploadN=<tableN>` `ucmdN=<cmds>`)
- Limitations:
 - ▷ Not interactive
 - ▷ Does not attempt to query service table/capabilities metadata
 - ▷ Offers no help in constructing queries
 - ▷ See Markus's [tapsh](#) for a more capable command-line client
- Example:

```
% stilts tapquery tapurl='http://dc.zah.uni-heidelberg.de/__system__/tap/run/tap'  
          adql='SELECT TOP 100000 * FROM ppmxl.main'  
          out=ppmxl.fits  
  
SUBMITTED ...  
http://dc.zah.uni-heidelberg.de/__system__/tap/run/tap/async/Wn3umi  
QUEUED ...  
EXECUTING ...  
COMPLETED  
DELETED
```

Release

First public release with TAP client 9 May 2011:



TOPCAT

v3.8

<http://www.starlink.ac.uk/topcat/>



STILTS

v2.3

<http://www.starlink.ac.uk/stilts/>

(TAP/UWS client Java classes somewhat reusable)

TAP Services

Existing TAP services:

- GAVO, Heidelberg
 - ▷ 1 service, 59 tables
 - ▷ Full implementation including uploads
- CADC
 - ▷ 1 service, 14 data tables
 - ▷ Not registered
- AstroGrid/DSA (CASU, WFAU)
 - ▷ CASU: 5 services, 19 data tables; WFAU: 20 services, ?? data tables
 - ▷ Some pre-TAP-1.0 features (ADQL-1.0, non-standard table metadata, ...)
 - ▷ No TAP_SCHEMA
 - ▷ No ADQL geometry functions
 - ▷ Upgrades in progress
- VO Paris Data Center
 - ▷ 3 services, 3 data tables
 - ▷ Synchronous only (no async)
 - ▷ MySQL not ADQL
- Also ESA-VO (moribund), HEASARC (experimental), CDS, JHU (under development)
- . . . any others?

TAP Service Registration

Attempt to locate TAP services through searchable registries

- Query the three public searchable registries using **Registry Interface 1.0** (the only standard interface to VO Registries):
 - 😊 AstroGrid Registry
 - ▷ Main registry correct and reliable
 - ▷ Backup registry also has TAP services since early April 2011
 - 😞 Euro-VO Registry
 - ▷ No way to specify TAP capability in searches
 - ▷ Should be fixed when TAPRegExt is published/finalized?
 - 😞 NVO Registry
 - ▷ RI 1.0 interface has been completely broken since mid-2010
- CADC service not registered anywhere (problems with NCSA publishing reg?)
- TAP registry records can be large → RI Search query slow

Issues/Thoughts

- Issues:

- Non-compliant services
 - ▷ e.g. ObsPM has no async, DSA has no geometry functions
 - ▷ Client has no (respectable) way of telling, so queries fail
 - ▷ Should a responsible client use private knowledge?
- Registry issues
 - ▷ Unregistered services, difficulty searching for TAP
 - ▷ Should I give up and use GloTS?
- Wide range of sizes for TAP services
 - ▷ ObsPM have 1 table per service, HEASARC 600. One day Vizier??
 - ▷ Does the same GUI make sense for big and small?
 - ▷ Can you make a sensible registry search that covers big and small?
- Complicated standard, many options
 - ▷ TOPCAT TAP GUI quite big/complex for user to grasp
 - ▷ Adding ObsTAP, PQL, better schema display... would make it more so

Issues/Thoughts

● Design questions:

- What would an ObsTAP client look like?
 - ▷ Do I just need to add ObsTAP-friendly examples to the Examples menu?
 - ▷ Should it send the same query to *all* TAP services? (What if many more services arise?)
 - ▷ Should the result be a single table?
- What would a PQL client look like?
- Should sync or async query be the default?
- [When] should the client delete jobs by default?

● Service provider support:

- Lots of people want to deploy TAP services
 - ▷ Guidance (web/wiki page?) with guidance for data providers on what's available (*in fact this is a general DAL issue*)
- Validation
 - ▷ Informal validation: hitting services with experimental client
 - ▷ Takes quite a bit of effort for both client and service developers
 - ▷ But leads to service compliance improvements