TOPCAT TAP Client Improvements

Mark Taylor (Bristol)

IVOA Interop
Sesto

17 June 2015

$Id: toptap.tex,v 1.20 2015/06/17 08:45:14 mbt Exp $

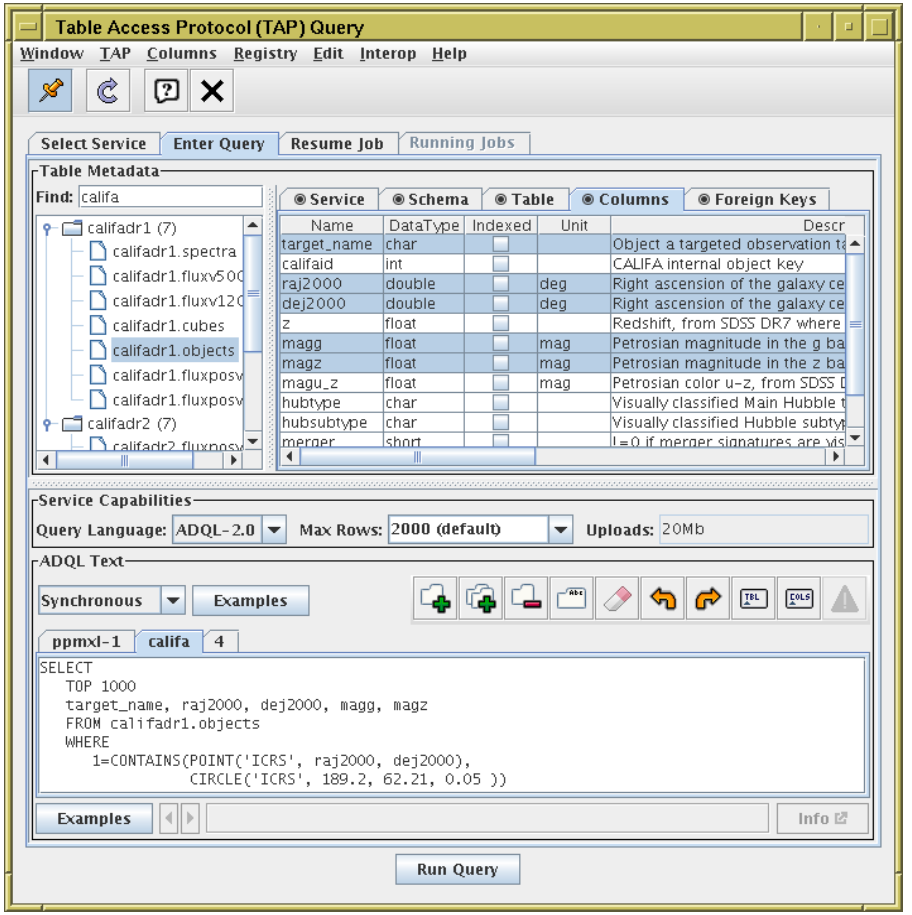
Outline

- Summary of improvements
- Demo
- Implementation feedback
- Status
- Outstanding issues

Improvements

New in this (pre-)release:

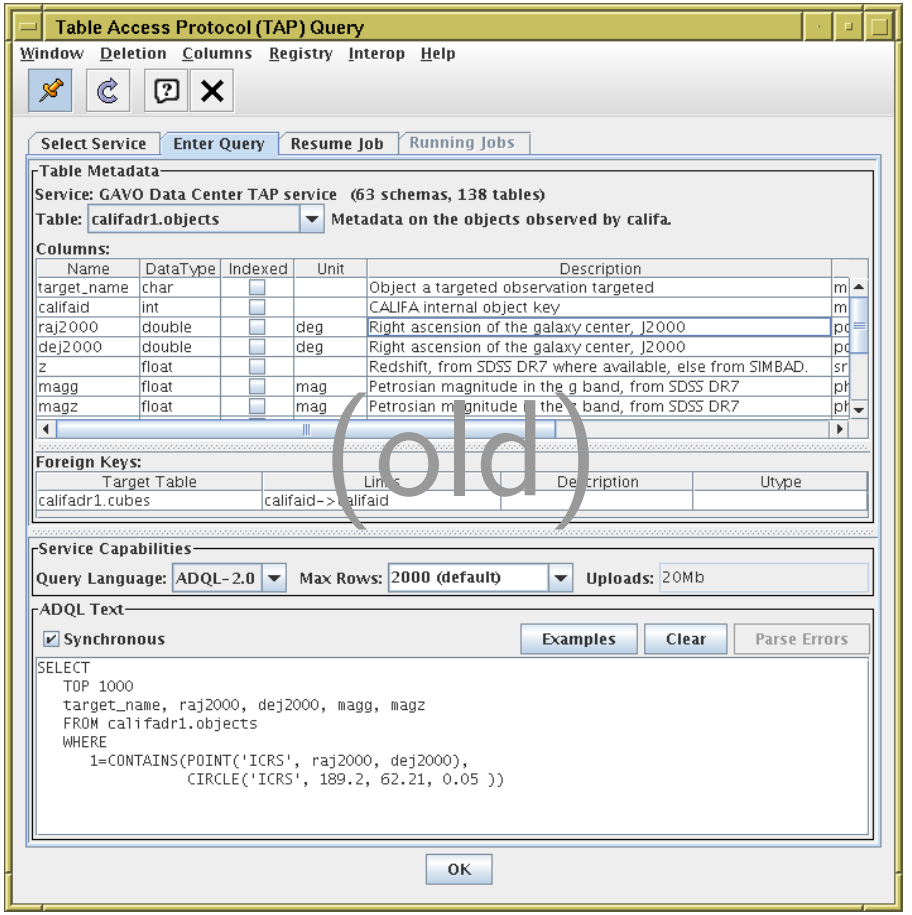
- Improved metadata acquisition: pluggable and scalable to many tables
- Improved metadata display GUI: scalable to many tables
- More per-service and per-table metadata displayed
- More examples, including service-provided ones
- Better ADQL editing panel
- Query result FORMAT negotiation
- Quick-look query mode



Improvements

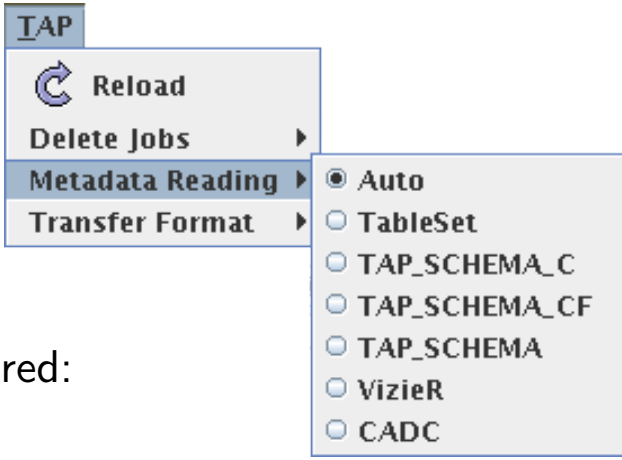
New in this (pre-)release:

- Improved metadata acquisition: pluggable and scalable to many tables
- Improved metadata display GUI: scalable to many tables
- More per-service and per-table metadata displayed
- More examples, including service-provided ones
- Better ADQL editing panel
- Query result FORMAT negotiation
- Quick-look query mode



Metadata acquisition

Pluggable metadata acquisition options:



- 1-stage: All metadata (schemas, tables, columns, keys) up front:
  - ▷ from standard `/tables` endpoint
  - ▷ from TAP_SCHEMA queries
- 2-stage: Tables and schemas up front, columns and keys as required:
  - ▷ from TAP_SCHEMA queries
  - ▷ from CADC-style multi-stage `/tables` endpoint (proposed VOSI 1.1)
  - ▷ from VizieR-style multi-stage `/tables` endpoint (non-standard)
- Auto
  - ▷ Chooses one of the above (e.g. by using total column count)

2-stage options are scalable

- Workable with VizieR-scale service (30k tables, 400k columns, $\lesssim$10 seconds initial load)
- Some care required when scheduling/threading column metadata requests

Additional tweaks made in some cases

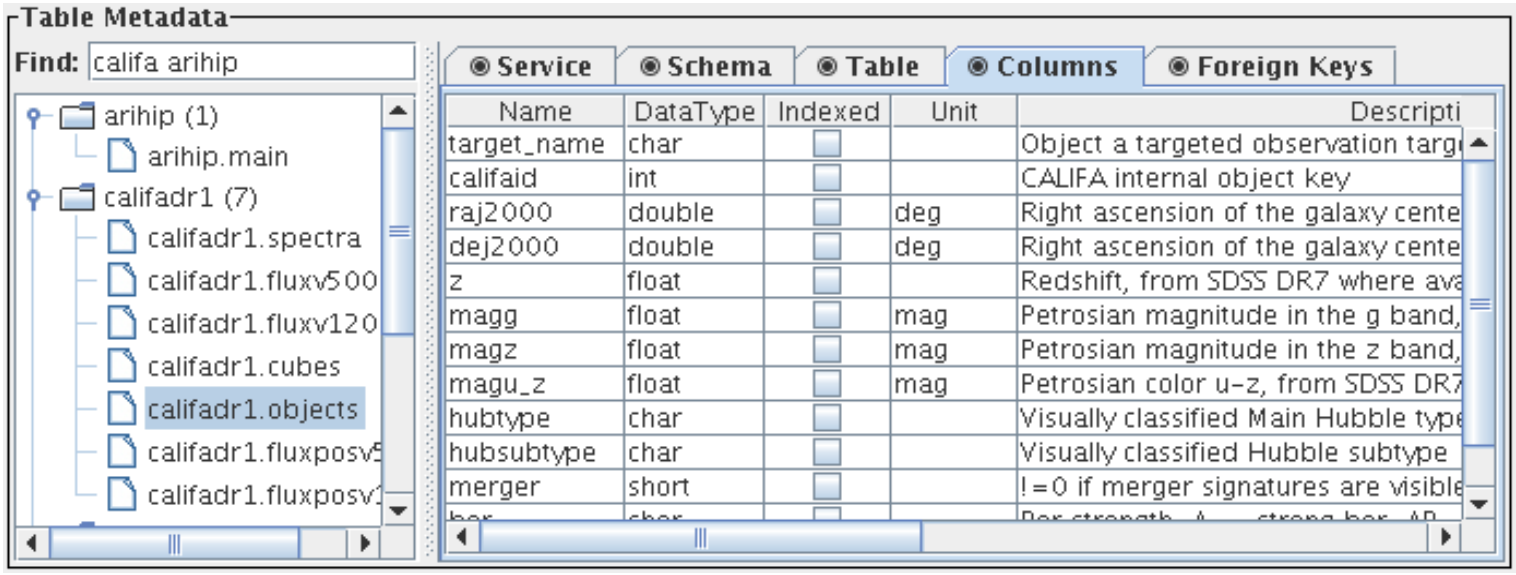
- Attempt to quote not-delimited-but-should-be column/table identifiers

Metadata display

- Better scalability
  - JTree not JComboBox
  - Tables displayed hierarchically by schema
  - Instant search by table name

    ⟶ *usable with TAPVizieR!*

- More information displayed
  - Service metadata from registry (description, reference URL, ...)
  - TAPRegExt language features (supported DMs, UDFs, geometry functions, ...)
  - Full table and schema metadata (descriptions)

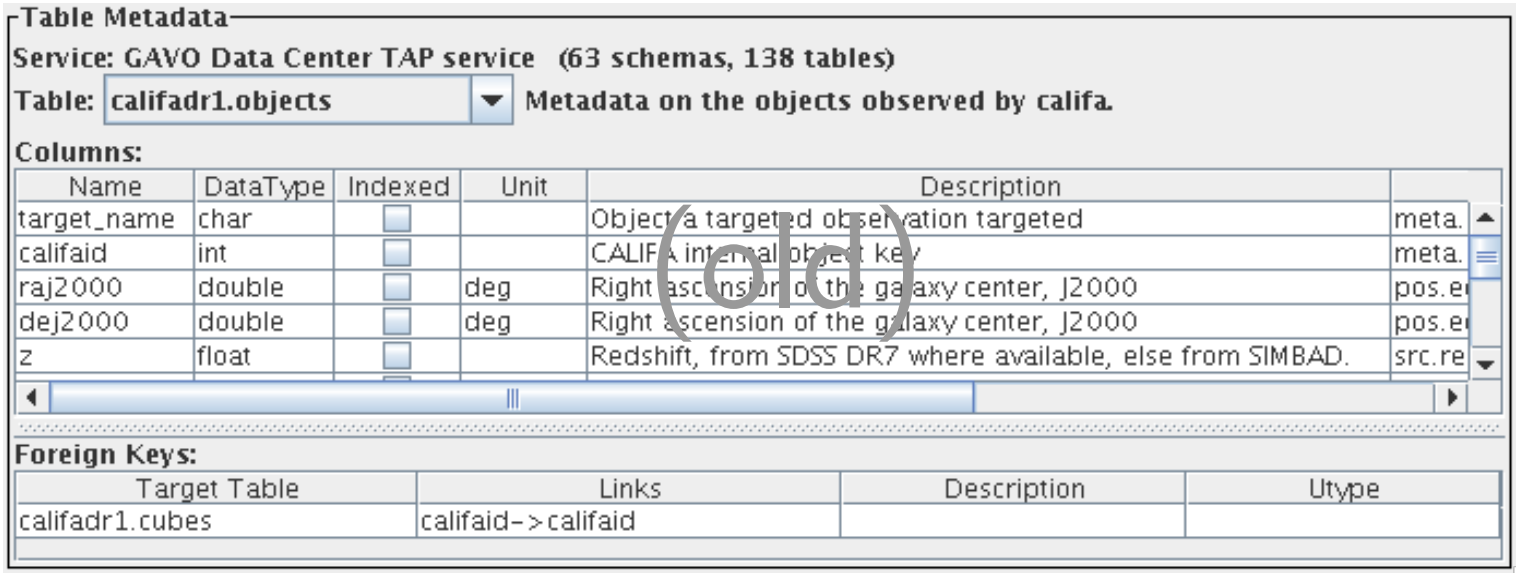


Metadata display

- Better scalability
  - JTree not JComboBox
  - Tables displayed hierarchically by schema
  - Instant search by table name

    ⟶ *usable with TAPVizieR!*

- More information displayed
  - Service metadata from registry (description, reference URL, ...)
  - TAPRegExt language features (supported DMs, UDFs, geometry functions, ...)
  - Full table and schema metadata (descriptions)



(old)

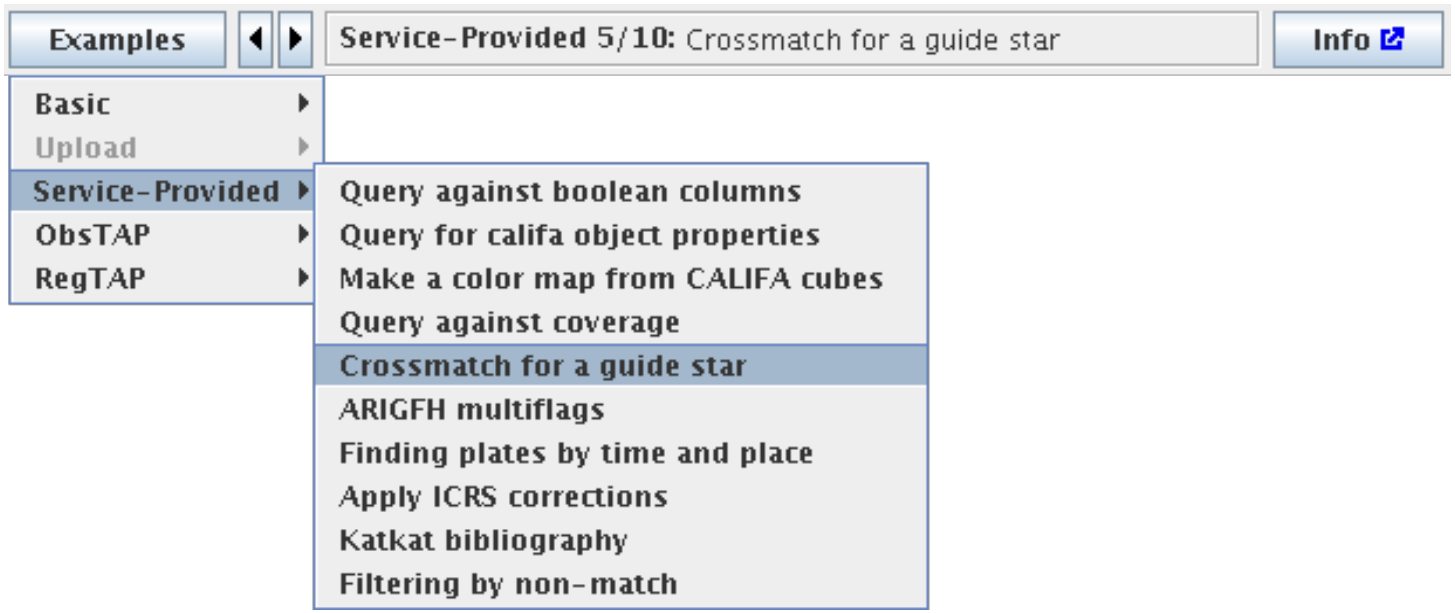
Examples

More ADQL example queries:

- Service-speciic examples from service DALI/TAP `/examples` endpoint (XHTML+RDFa)
- Canned RegTAP and ObsTAP examples if data models declared

Example browsing:

- Buttons to cycle through example lists
- Link to example documentation URLs (opened in browser)



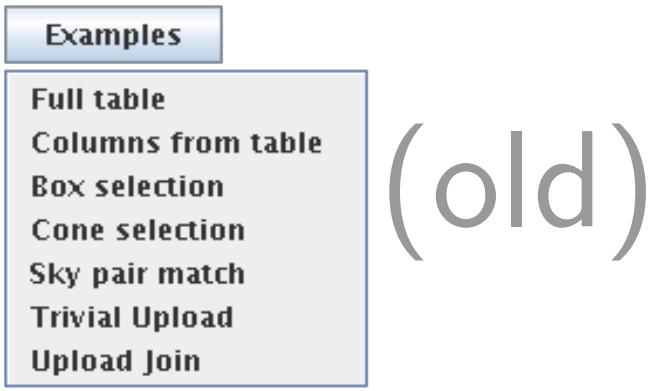
Examples

More ADQL example queries:

- Service-speciic examples from service DALI/TAP `/examples` endpoint (XHTML+RDFa)
- Canned RegTAP and ObsTAP examples if data models declared

Example browsing:

- Buttons to cycle through example lists
- Link to example documentation URLs (opened in browser)



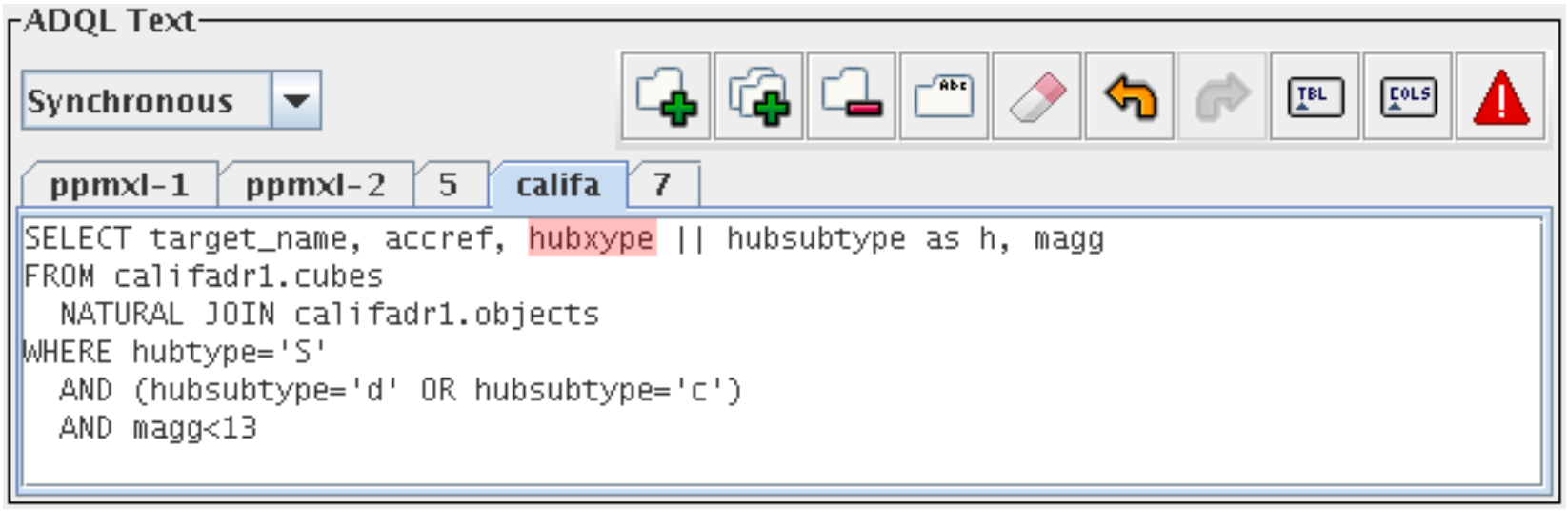
(old)

ADQL Editing

More functions added to ADQL editing window

- Multiple tabs allow multiple queries on the go
- Text edit Undo/Redo actions (buttons or Ctrl-[Shift]-Z)
- Convenience buttons for inserting selected table/column names
- Version 1.3 of CDS/ARI ADQL lib for metadata-aware ADQL parsing

  UDF-aware parsing coming soon?

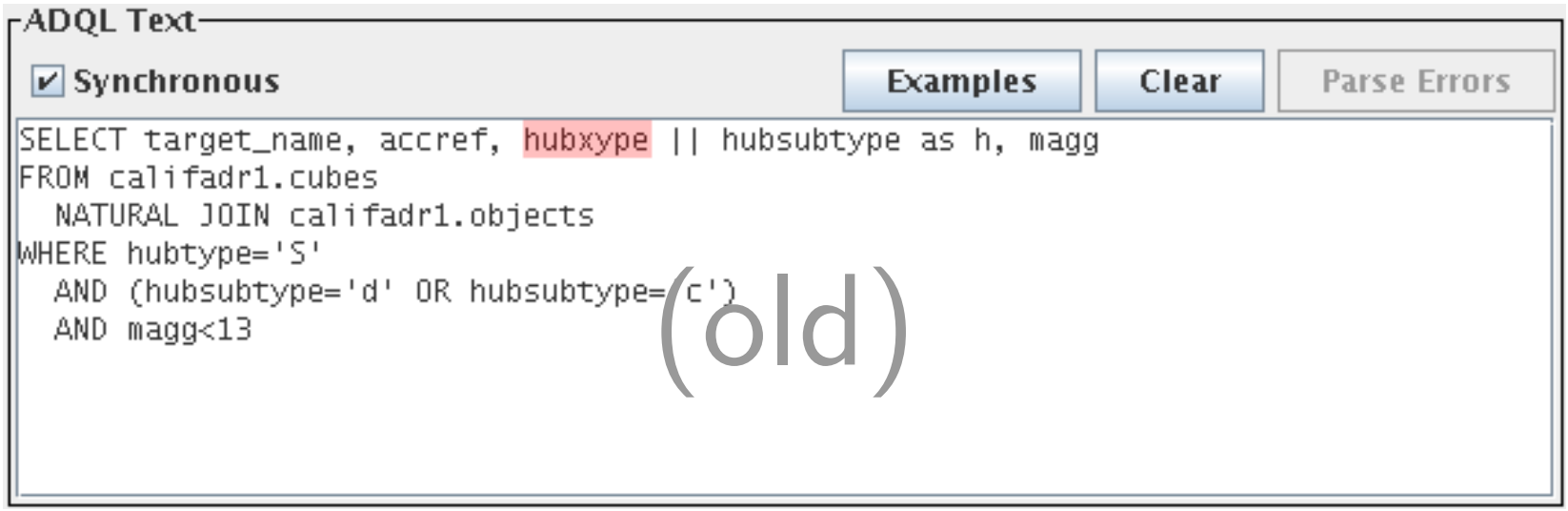


ADQL Editing

More functions added to ADQL editing window

- Multiple tabs allow multiple queries on the go
- Text edit Undo/Redo actions (buttons or Ctrl-[Shift]-Z)
- Convenience buttons for inserting selected table/column names
- Version 1.3 of CDS/ARI ADQL lib for metadata-aware ADQL parsing

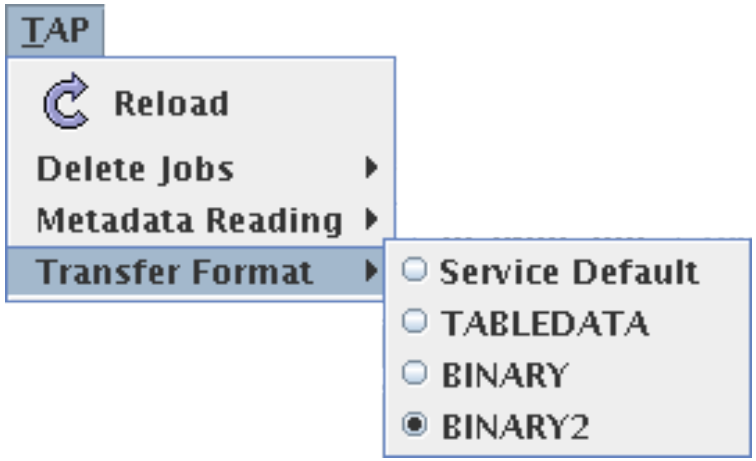
  UDF-aware parsing coming soon?



Result Format Negotiation

Uses FORMAT parameter for result transfer format selection

- Default to BINARY2 if service provides it
- User may choose preferred format by menu
- Performance may be affected, results not much



Use HTTP Content-Coding for HTTP compression

- Can deliver big reductions in bandwidth use
- See my talk in GWS I

Demo

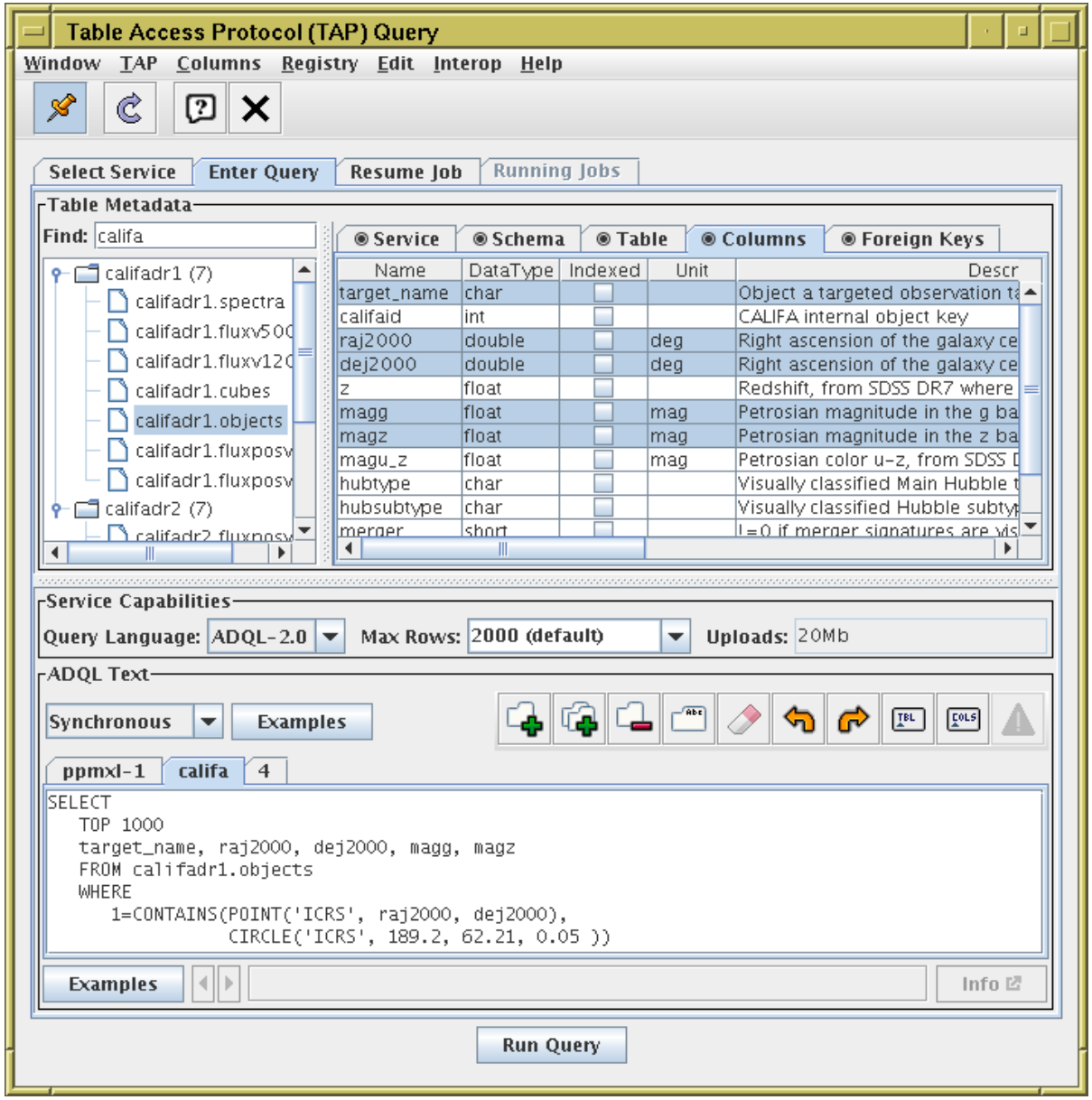


Table Access Protocol (TAP) Query
Window TAP Columns Registry Edit Interop Help
Select Service
Enter Query
Resume Job
Running Jobs
Table Metadata
Find: califa
califadr1 (7)
califadr1.spectra
califadr1.fluxv500
califadr1.fluxv120
califadr1.cubes
califadr1.objects
califadr1.fluxposv
califadr1.fluxposv
califadr2 (7)
Service
Schema
Table
Columns
Foreign Keys
Name DataType Indexed Unit Descr
target_name char Object a targeted observation ta
califaid int CALIFA internal object key
raj2000 double deg Right ascension of the galaxy ce
dej2000 double deg Right ascension of the galaxy ce
z float Redshift, from SDSS DR7 where
magg float mag Petrosian magnitude in the g ba
magz float mag Petrosian magnitude in the z ba
magu_z float mag Petrosian color u-z, from SDSS D
hubtype char Visually classified Main Hubble t
hubsubtype char Visually classified Hubble subtyp
merger short I=0 if merger signatures are vis
Service Capabilities
Query Language: ADQL-2.0
Max Rows: 2000 (default)
Uploads: 20Mb
ADQL Text
Synchronous
Examples
ppmxl-1
califa
4
SELECT
TOP 1000
target_name, raj2000, dej2000, magg, magz
FROM califadr1.objects
WHERE
1=CONTAINS(POINT('ICRS', raj2000, dej2000),
CIRCLE('ICRS', 189.2, 62.21, 0.05))
Examples
Info
Run Query

Implementation Feedback

General comments:

- TAP and related standards are complex, but mostly provide what's required
- RegTAP is a huge improvement on RI 1.0 (I can now rely on registry queries)
- TAP service quality is patchy (implementation errors, missing functionality)

Status

Available in pre-release

- Working
- User documentation TBD
- Available for standalone/third-party use (TOPCAT application classes not required)
- TAP service client fairly complete
- TAP service discovery still poor
- Your feedback is welcome, especially service providers
- Public release (TOPCAT v4.2-4?) in next few weeks?

`ftp://andromeda.star.bris.ac.uk/pub/star/topcat-full_tap.jar`

Outstanding Issues

TBD:

- **Better discovery UI for (useful) TAP services**
- Some remaining ADQL library issues
  - Periods in table names
  - UDF validation
- Standards under discussion
  - DALI/TAP `/examples` endpoint
  - Scalable `/tables` endpoint (VizieR vs. CADC)
- UWS 1.1 *(no services)*
- Use HTTP Content-Coding? (see my talk from GWS I)
- Fixup machinery for non-identifier names
- Multiple TAP window instances per TOPCAT?

Open Questions

- Authentication?
- Anything else I should be doing?
- Complaints/concerns from service providers?