

Client-side RegTAP in TOPCAT

Mark Taylor (Bristol)

Registry I

IVOA Interop

ESAC

21 May 2014

`$Id: tcregtap.tex,v 1.5 2014/05/21 06:49:11 mbt Exp $`

Overview

Context and requirements

Implementation

TOPCAT Requirements

- Locate VO services
 - Restricted by resource type:
 - ▷ Services Cone, SIA, SSA, TAP
 - Restricted by keyword search terms (e.g. “2mass”, “redshift”)
 - ▷ Match against IVO-ID, ShortName, Title, Publisher, Subjects, Description
 - Only a few items of information required:
 - ▷ Display ShortName, Title, Subjects, Publisher, Contact, Reference URL
 - ▷ Use Service URL
- (nice-to-have) Locate alternative searchable registries

Old GUI

Registry queries using RI1.0

Available Cone Services

Registry:

Keywords:

Match Fields: Short Name Title Subjects ID Publisher Description

Accept Resource Lists

Short Name	Identifier	Title	Subjects
6dF DR2	ivo://wfau.roe.ac.uk/6df-dsa	6dF Galaxy Survey Data Release 2	6dF survey galaxy galaxies re
6dF DR3	ivo://wfau.roe.ac.uk/6dfdr3-dsa	6dF Galaxy Survey Data Release 3	6dF survey galaxy galaxies re
VII/259	ivo://CDS.VizieR/VII/259	6dF galaxy survey final redshift rele...	Photometry, Galaxies, Redshif
VII/249	ivo://CDS.VizieR/VII/249	6dF-DR2 Galaxy Survey (Jones+, 20...	Galaxies, Redshifts
SIXDFGS	ivo://nasa.heasarc/sixdfgs	6dFGS Galaxy Survey Final Redshift ...	Galaxy
J/MNRAS/273/157	ivo://CDS.VizieR/J/MNRAS/273/157	A faint galaxy survey to B = 24. (Gla...	Photometry, Galaxies, Redshif
J/ApJ/200/8	ivo://CDS.VizieR/J/ApJ/200/8	AGES: the AGN and Galaxy Evolution...	AGN, Photometry, Galaxies, C
J/ApJ/691/705	ivo://CDS.VizieR/J/ApJ/691/705	AGN host galaxy morphologies in CO...	AGN, Galaxies, Redshifts
I/AI/120/2190	ivo://CDS.VizieR/I/AI/120/2190	Caltech Faint Galaxy Redshift Survey	Galaxies, Redshifts

AccessURL	Description	Version
http://wfaudata.roe.ac.uk/6dfdr3-dsa/...	twomass table: cone search for K, J and ...	
http://wfaudata.roe.ac.uk/6dfdr3-dsa/...	denisj table: cone search for sources fro...	
http://wfaudata.roe.ac.uk/6dfdr3-dsa/...	agn2mass table: cone search for candid...	

New GUI

Registry queries using RI1.0 or RegTAP

Available Cone Services

Registry: RegTAP

Keywords: And

Match Fields: Short Name Title Subjects ID Publisher Description

Accept Resource Lists

Short Name	Identifier	Title	Subjects
6dF DR2	ivo://wfau.roe.ac.uk/6df-dsa	6dF Galaxy Survey Data Release 2	6dF survey galaxy galaxies re
6dF DR3	ivo://wfau.roe.ac.uk/6dfdr3-dsa	6dF Galaxy Survey Data Release 3	6dF survey galaxy galaxies re
VII/259	ivo://CDS.VizieR/VII/259	6dF galaxy survey final redshift rele...	Photometry, Galaxies, Redshif
VII/249	ivo://CDS.VizieR/VII/249	6dF-DR2 Galaxy Survey (Jones+, 20...	Galaxies, Redshifts
SIXDFGS	ivo://nasa.heasarc/sixdfgs	6dFGS Galaxy Survey Final Redshift ...	Galaxy
J/MNRAS/273/157	ivo://CDS.VizieR/J/MNRAS/273/157	A faint galaxy survey to B = 24. (Gla...	Photometry, Galaxies, Redshif
J/ApJ/200/8	ivo://CDS.VizieR/J/ApJ/200/8	AGES: the AGN and Galaxy Evolution...	AGN, Photometry, Galaxies, C
J/ApJ/691/705	ivo://CDS.VizieR/J/ApJ/691/705	AGN host galaxy morphologies in CO...	AGN, Galaxies, Redshifts
I/AI/120/2190	ivo://CDS.VizieR/I/AI/120/2190	Caltech Faint Galaxy Redshift Survey	Galaxies, Redshifts

AccessURL	Description	Version
http://wfaudata.roe.ac.uk/6dfdr3-dsa/...	twomass table: cone search for K, J and ...	
http://wfaudata.roe.ac.uk/6dfdr3-dsa/...	denisj table: cone search for sources fro...	
http://wfaudata.roe.ac.uk/6dfdr3-dsa/...	agn2mass table: cone search for candid...	

RI1.0 Implementation

SOAP Search operation:

- ADQL/S:

```
Search WHERE ( capability/@standardID = 'ivo://ivoa.net/std/TAP' )
            AND ( ( title LIKE '%gavo%' ) OR ... )
```

- Issues:

- ▶ Retrieves full XML VOResource records (several MByte each in some cases)
 - No mandatory RI1.0 operations can *both* restrict by service type *and* select response elements
- ▶ Case sensitivity undefined for **LIKE** matching (OK in practice)
- ▶ Searching registry for searchable registries never worked

RegTAP Implementation

Single synchronous ADQL TAP query

- ADQL 2.0:

```
SELECT ivo_id, short_name, res_title, reference_url, base_role, role_name,
       email, intf_index, access_url, standard_id, cap_type,
       cap_description, std_version, res_subject
FROM rr.resource AS res
NATURAL JOIN rr.interface
NATURAL JOIN rr.capability
NATURAL LEFT OUTER JOIN rr.res_role
NATURAL LEFT OUTER JOIN rr.res_subject
WHERE (base_role='contact' OR base_role='publisher' OR base_role IS NULL)
      AND ( standard_id='ivo://ivoa.net/std/tap' AND
            ( intf_role='std' OR intf_type='vs:paramhttp' ) )
      AND ( 1=ivo_hasword(res_title, 'gavo') OR ... )
```

- Result somewhat denormalised ($N_{\text{row}}/N_{\text{resource}} \approx 4$)

- ▷ Multiple `base_role` (publisher/contacts) and `res_subject` values
- ▷ Aggregate functions `string_agg/array_agg` in ADQL/RegTAP would help, but not essential

Performance

RI1.0/RegTAP query response size comparison

- Results:

Query	N _{resource}	Response size in kbytes	
		RI1.0	RegTAP
All TAP	~50	25,000	150
All SSA	~80	580	160
Cone, title like "IUE"	~15	200	44
Cone, title/publisher like "ukidss dr9 wfau"	1	1,100	13

- Summary

- ▷ "Normal" resources: RegTAP $\sim 4\times$ smaller
- ▷ "Pathological" (fine-grained table metadata) resources: RegTAP $\sim 100\times$ smaller

Trials

- GAVO RegTAP (Heidelberg + AIP) seems fast and reliable, it all works
- No other RegTAP services to test against?

Conclusions

- Success! 😊
 - Implementation and service access working well
 - Relational Registry document PR in very good shape
 - Implementation took some iteration to get right, but no major problems
 - *Lots* of help from Markus, but mostly SQL help not clarifying document
 - Nice to have: `array_agg/string_agg`
- Release
 - Pre-release available: <ftp://andromeda.star.bris.ac.uk/pub/star/topcat/pre/>
 - Registry access default to RegTAP in next TOPCAT public release **if RegTAP is stable**
- Ready to point at other RegTAP implementations
 - I can hardwire persistent RegTAP endpoints into the code
 - TOPCAT can search registry dynamically for other RegTAP services
 - I'm happy to help testing new RegTAP services