TAP REGISTRATION EXTENSIONS FOR TAP FEDERATION FACTORY AND ASTRODABIS

Federating TAP Services

TAP Federation Factory Service

- Takes multiple TAP endpoints (and the table metadata endpoints), and creates a new temporary TAP service that federates the data from all specified endpoints
- Allows user to query federated datasets as though they are all in the same database
- Built using OGSA-DAI grid middleware and AstroGrid DSA software

How it works

- User submits HTTP POST message that contains a series of TAP endpoint URLS and their respective table metadata URLS
- Factory service harvests the table metadata and combines them into a single metadoc
- DSA instance is deployed using the generated metadoc, and is connected to OGSA-DAI through JDBC
  - The underlying connections to the TAP services are handled by OGSA-DAI
- The TAP endpoint for the federated TAP service is returned to the user
- Users can query the new service like any other TAP service
  - Query planning and execution are performed by OGSA-DAI behind the scenes

TAP Factory Service

- Detailed description:
  http://www.ogsadai.org.uk/demos/astro/

- HTML Interface:
  http://admire3.epcc.ed.ac.uk:8081/TAPFactory/

- See me anytime this week for a demo...

Future Plans

- Deploy and publish service to general users
  - Still deciding how to publish
- Deploy a small number of permanent TAP Federations for high use databases
  - Example: SDSS, UKIDSS, TWOMASS
  - Still determining other combinations that would be of use to a large number of users

General Issues Discovered

- Metadata access not yet standardized
  - Many current implementations don't support metadata queries
  - We utilize VOSI metadata currently, but this isn't ideal
- VOTable Binary output not always available, or it is not documented how to specify it (this appears to be addresses in the current TAPRegExt WD)
- Service restrictions not specified (again this appears to be addressed in the current WD)

AstroDAbis

- Annotation service for astronomical data
  - Allows users to annotate database rows (objects) with different tags and share these annotations with others
    - Users can upload a table or execute queries on TAP services to create their datasets
    - The AstroDAbis server can be queried to find objects with certain tags, or objects can be queried to see what tags have been applied to them
    - Tags can be applied to individual objects as well as crossmatch identifications

Extensions we'd like to see

- Histogram data for major database columns
  - All DBMS's provide this information, but it is not published
  - This data is extremely useful for federated query planning and execution (data reduction, JOIN algorithm ordering, and JOIN algorithm selection are all examples of areas where histogram data can be applied)
  - If histogram data seems too complicated registering the following would also be useful:
    - Column Metadata – minimum value, maximum value, number of unique values
    - Table Metadata – number of rows, number of columns
    - Indication of date for statistics (in case the DB changes)

Extensions continued...

- Primary Key flag
  - An indicator of which column(s) constitute the primary key for a table
    - Useful in query planning
    - Also useful for AstroDAbis project

For more details

- The TAP Factory Service will be presented at ADASS (poster) - http://www.eso.org/sci/php/meetings/adass2011/html/display.php?topic=Poster_Hume_1313833337.html
- AstroDAbis will likewise be presented at ADASS (poster) - http://www.eso.org/sci/php/meetings/adass2011/html/display.php?topic=Poster_Gray_1316092286.html