VIRTUAL ASTRONOMICAL OBSERVATORY

VAO SIAV2 Prototype - Advanced Capabilities

Doug Tody, NRAO, USVAO

SIAV2 Request Types

Request	Sync	Async	Description
QueryData	MAN	OPT	Query the service to discover available image datasets matching the given query constraints. The response is a table listing all available datasets. Full metadata is returned describing each image dataset. The described datasets may be either static ***archival*** datasets (pre-existing datasets in the remote archive), or ***virtual data*** datasets more closely matching the client request, to be computed and returned upon request by the client. An **access reference URL** is returned for each dataset that may be used to directly download the dataset.
AccessData	OPT	OPT	Directly access a single image dataset, returning the requested view of the data (usually a small subset, e.g., a filtered product, cutout or projection). AccessData by its nature normally computes and returns virtual data.
StageData	n/a	OPT	Request asynchronous computation and staging of data products (often virtual data) referenced in a prior QueryData operation.

QueryData

- Approach
  - Basic queryData is a conventional DBMS query
    - Query DBMS, format and return query response VOTable
  - Cutout mode "plans" a virtual image
    - Invokes (via *accessData*) an external task to do this
    - Updated metadata for virtual image returned in query response
      - Metadata = archival image metadata edited as necessary for cutout
    - Access reference URL provided to retrieve virtual dataset

Back-End Processing

- Usage
  - External task is used to "plan" virtual image
  - Simple example of back-end processing capability
- Motivation
  - Image computation can be complex, domain-specific
  - Custom processing often required
    - very similar to conventional pipeline processing
  - Scalability required for some use cases
    - large cubes mandate cluster computing on back-end

Back-End Processing

- *VOCutout* Task (in prototype)
  - Provides a generic cutout capability for FITS images
  - A single C program
    - built using Starlink AST WCS library, and cfitsio

- Functionality
  - "Plan" virtual image
    - Result written to *data staging area*
    - Metadata also used to build queryData response
  - Generate virtual image
    - In response to an access request
    - Generated images are cached in staging area

AccessData Request

- Functionality
  - Archive image retrieval
    - PubDID references a static archival image
    - Example: "ivo://nrao/vo#siav2model:3"
  - Retrieval of virtual data after a queryData
    - PubDID references the "plan" for a virtual image
    - Example: "ivo://nrao/vo#image-rKhul8"
    - Virtual image is generated and streamed back to client
  - Direct access to a dataset
    - Combines the "plan" and image generation into one operation
    - Client specifies image via PubDID, filter params for cutout

Virtual Image Plan and Metadata

```
MDFILE = image-rKhul8
image = /d1/testData/vaodata/jvla/orionall_hannclean_hotcore.fits
section = [72:24,24:72,1:24012,1:1]
Naxes = 3
Naxis = 47 49 24012
EstSize = 216014
DataLength = 55299636
CreationType = cutout
SpatialLocation = 83.8110425265584 -5.37548608932131
SpatialLoLimit = 83.8060204418066 -5.38048609103702
SpatialHiLimit = 83.8160646113101 -5.3704860876056
SpectralLocation = 0.011918431905180 6
SpectralStart = 0.012674608507831 5
SpectralStop = 0.011247403423014 6
END
```

Some Useful Standard Queries

- Initial image discovery
  - queryData <*constraints*>, or use TAP, or archive-specific query
- Retrieve metadata for an image
  - queryData PubDID=XX, possibly retrieve *dataref* to get WCS
- Plan an image cutout, e.g., to prepare a Datalink
  - queryData PubDID=XX <*filter-params*>
- Generate and Retrieve the Cutout Image
  - accessData PubDID=XX [FORMAT=XX]
- Direct Access to Image by a Client
  - accessData PubDID=XX <*filter-params*> [*WCS, pixel, function terms*]

Summary

- SIAV2
  - Provides essential capabilities for image data
    - Discovery, ImageDM metadata retrieval
    - Cutouts and other direct, precision image access
  - Based upon the Image Data Model
  - Image-specific access model and capabilities

- Mix and Match
  - Capabilities are available:
    - Integrated into a single, easy to use service interface
    - Separately available, e.g. for use with ObsTAP and Datalink