Data Access Meeting 2016-02-29

Date

29 Feb 2016

Attendees

- IPAC: Unknown User (xiuqin), Trey Roby
- SLAC: Brian Van Klaveren, Fritz Mueller, Unknown User (npease), John Gates, Vaikunth Thukral, Kian-Tat Lim
- AURA: Tim Jenness

Goals

- Initial survey of VO standards – which do we need/want to adopt?

Discussion items

VOTable

- SUI/T is concerned about efficiency of data transfer with VOTable for large results.
- VO BINARY2 may be a reasonably performant alternative. Need to check if reference implementations of ser/des are available for Java, C++, Python.
- Could we work with IVOA to get an additional encoding officially adopted, based on more contemporary and widely adopted ser/des frameworks (e.g. protobufs or capnproto)?
  - These frameworks have advantages (and disadvantages) of automatic stub generation for multiple languages.
  - Could be generic row-based result protocol; would not have to be touched/changed for every different result row type. This is done currently for Qserv, where protobufs are used for the wire ser/des for query results flowing from workers to czar.
  - It would be helpful to have a working example to demonstrate efficiency wins to IVOA in order to advocate adoption. Consider a test based on existing Qserv protocol description.
  - An advantage of getting IVOA adoption is that we could then perhaps dispense with poorly performing encodings which are otherwise currently mandated by IVOA.
- Tim notes that science pipeline is currently looking at migrating afwtable to astropy; once this is done it is much easier to go from astropy to VOTable.
- Firefly does have a VOTable parser.

TAP

- IPAC has some very preliminary experience accessing other catalogs via TAP. IRSA has just published via TAP.
- IPAC has work on their plate to adapt Firefly to TAP, but hasn't begun this work yet beyond a very simple prototype.
- IPAC considering/recommending working with Mark Taylor (TOPCAT) in consulting capacity.
- dbserv will be adapted to TAP; Brian will be undertaking this work.
- Brian has some preliminary concern about slightly different semantics of sync/async between TAP and Qserv as currently formulated. We shall see...
- Need to coordinate this work so we don't break/stall SUI/T pending TAP client-side work.

ADQL

- IPAC has a simple ADQL parser in C++. Qserv team should check this out and see if we could leverage this for upcoming parser work. Currently targets Oracle, but we will look anyway 😉

SIA

- Simple cone search, etc. may overlap with imageserv/metaserv. Should we adapt our services to SIA, include SIA as alternate access method in our services, or advocate the IVOA adopt our protocols here?
- Tim recommends we also take a look at VO data-link service (as an alternative to SIA hardcoded approach?) for drilling down from catalog results to images.

VOSpace

- SUI/T will be driving an investigation of this, with respect to L3 user workspaces, etc.
- There may be some involvement from dbserv though (for example, request that sends query results to L3 workspace rather than streaming results directly to client).

CAOM

- MAST and IRSA are considering moving to CAOM
- Adoption could affect some of our table designs
- Would need to coordinate this with data production and alert production for L1
- Maybe something new ("obsserv", "visserv"?)

VOSI

- Defines minimal core of VO which must/should be implemented. May be a good place for us to look first.
- Mandates some VO registry things.
    - MAST provides a registry; could we publish into theirs?
        - IPAC/IRSA experience with MAST registry was not positive.
        - Tim mentions that JWST is motivating cleanup/maintenance here, so may be improved/improving.
    - Operational def'n for sucess with registry: can TOPCAT find you?

Misc

- How do we pick VO standards that are well-adopted "winners", vs. new fads that will be dead in a couple years?
    - Could MAST tell us user statistics?
    - Brian might have some folks at Goddard to talk to.
    - XiuQin's experience is that TAP, SIA, SSA are de-facto expected.
- Bonus lengthy digression into remote-Butler design and issues, a la mode... 😉
- Let's have a follow-up meeting at a time when we can get both Gregory and Jacek involved.