Via Lactea as VO application

(a DAL implementation example)

Robert Butora, Marco Molinaro, INAF Trieste

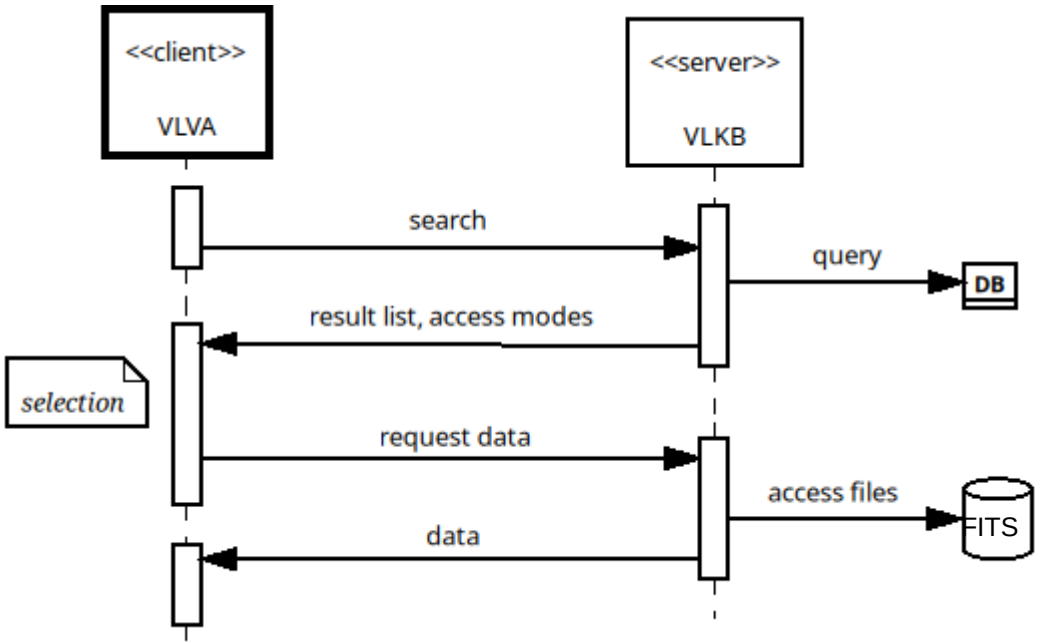
Via Lactea summary

- Study Galaxy in multiple wavelengths
- Based on ~25 surveys from last 10..20 years
  - 2D images
  - 3D velocity cubes
- Size total ~2TB
  - 1GB...300GB / survey
  - ~40000 FITS headers

WISE All-sky Survey (Wright et al. 2010, AJ 140, 1868)
CSO Bolocam Galactic Plane Survey (Ginsburg et al. 2013, ApJS 208, 14)
Hi-GAL - Herschel infrared Galactic Plane Survey (Molinari et al. 2016, A&A 591, 149)
…

Mopra Galactic Plane Survey (Burton et al. 2017)
JCMT CHIMPS Survey (Rigby et al. 2016, MNRAS 456, 2885)
Mopra CHaMP Survey (Barnes et al. 2011, ApJS 196, 12)
Nobeyama-45m Galactic Plane Survey (Umemoto et al. 2017, PASJ 69, 78)
VLA THOR Survey Inner Galactic Plane (Beuther et al. 2016, A&A 595, 32)
...

Interaction sequence

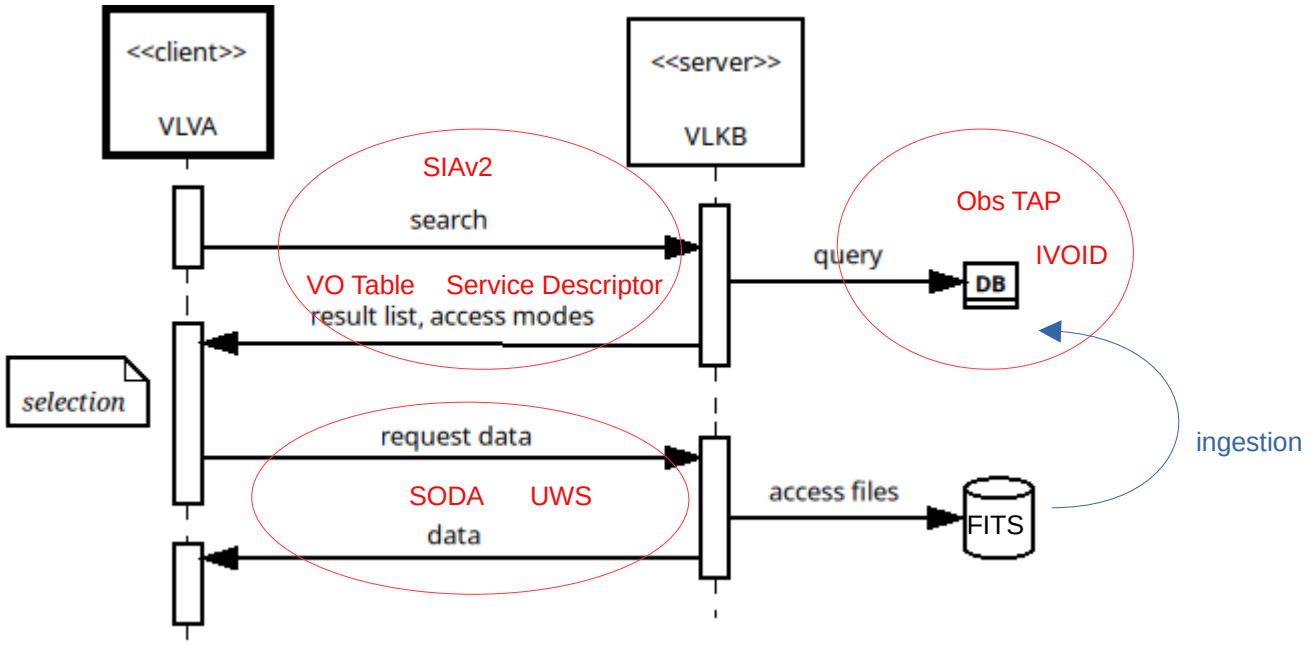


Functions:
- search
- cutout, mcutout*
- merge*

Used coords:
- GALACTIC
- VELO,LSRK,km/s

*mcutout = multi-cutout: multiple cutouts in one request, cuts returned in tar.gz
*merge = de-mosaicing of fits-data representing adjacent areas in the sky

VO protocol usage



mcutout, merge: no corresponding VO recommendation

search

ObsTAP extended:

- overlap code:
  - No overlap
  - Partial overlap
  - First inside second
  - Second inside first
- Sky region in GALACTIC
- Spectral range in VELO

xel	facility_name	instrument_name	overlap	overlapSky	overlapSpec	s_region_galactic	vel_min	vel_max
	JCMT	HARP	4	4		Polygon GALACTIC 40.350031 -0.105521 40.772750 -0.323795 40.991022 0.098920 40.568304 0.317195 unit deg	-50.0315308428	148.0008079588
	JCMT	HARP	3	3		Polygon GALACTIC 40.826413 -0.179712 41.179524 -0.179712 41.179524 0.173397 40.826413 0.173397 unit deg	-50.1684849247	121.3518675194
						Polygon GALACTIC		

Coord. system parametrization

Via Lactea data are velocity-cubes in galactic plane

- ‘Natural’ choice: GALACTIC, VELO/LSRK [km/s]
- VO standard: ICRS, WAVE/Barycentric [m]
- Added new params change interpretation of POS and BAND:
  - POSSYS: ICRS, GALACTIC
  - BANDSYS: WAVE_Barycentric, VELO_LSRK

POS=CIRCLE 40 0 0.1

POSSYS=GALACTIC

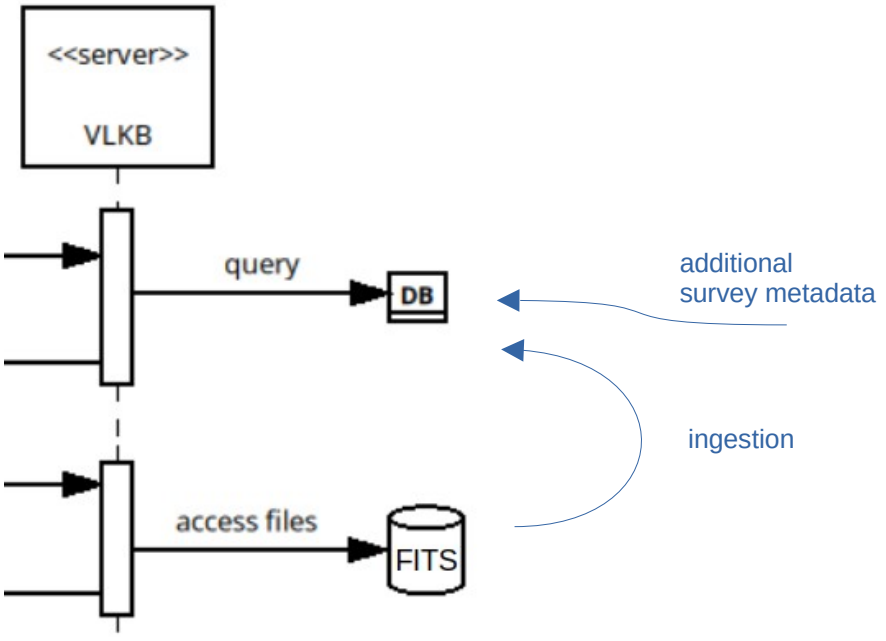
BAND=0 100

BANDSYS=VELO_LSRK

Coord. System conversion

Via Lactea data are velocity-cubes in galactic plane

- FITS header:
  - conversion needs more data then interpretation
    - like rest-frequency (not in header, FREQ0, in comments,...)
- Needs mechanism to complement FITS-header info:
  - Adding metadata to each survey
  - Performed by astronomers and added to data-store
  - Loaded to memory
  - Original FITS-headers not modified



Pixels coord. system

- For coord system not standard / unsupported
  - allow cutout by pixels
    - Params: PIXELS = [a:b c:d e:f]
    - Follows cfitsio extended filename syntax
  - In future separate by axis: PIXELSi = m:n:s
    - allows combine standard sky wcs-cut with pixels-cut on spec

POS=CIRCLE 84 -12 0.1 & PIXELS3=200:300:2

multiple cutouts

- Request n-cutouts in one request
- Response: n-cut FITS in one compressed file
- Solved by UWS JDL:
  - Collect standard SODA params in JSON array
    - Json arrays maintain order
  - Returns uws-results endpoints:
    - JSON array with result (ref to file or error)
    - Compressed cuts in tar.gz


```
{
  "id": "ivo://ia2.inaf.it/vlkb/datasets?CHIMPS/CHIMPS_C18O_40p00+0p00.fits",
  "pos": {
    "circle": {
      "lat": 0.045004,
      "lon": 39.9561,
      "radius": 0.1
    }
  },
  "possys": "GALACTIC"
},
{
  "id": "ivo://ia2.inaf.it/vlkb/datasets?JCMT-HARP/COHRS_40p00_0p00_CUBE_REBIN_R1.fit",
  "pos": {
    "circle": {
      "lat": 0.045004,
      "lon": 39.9561,
      "radius": 0.1
    }
  },
  "possys": "GALACTIC"
},
```



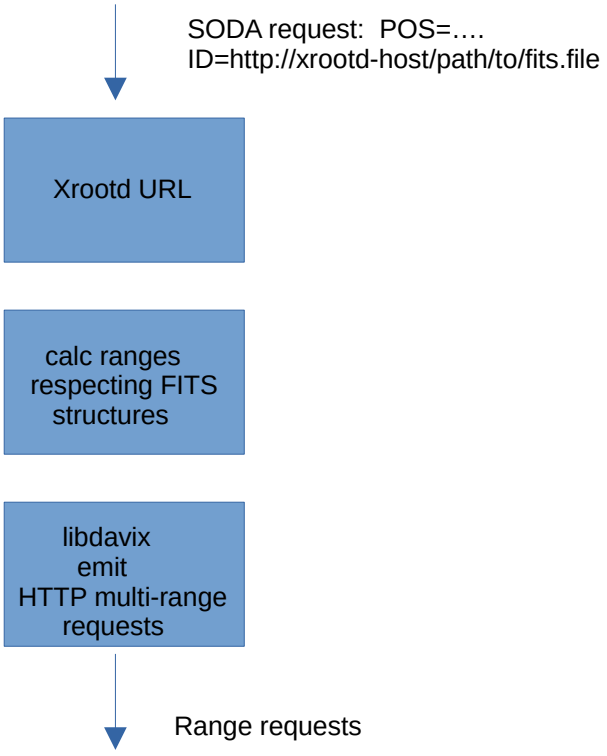
```
{
  "type": "FILENAME",
  "content": "/srv/cutouts/1739176796504A/22_CHIMPS-CHIMPS_13CO_40p00+0p00.fits"
},
{
  "type": "FILENAME",
  "content": "/srv/cutouts/1739176796504A/23_CHIMPS-CHIMPS_C18O_40p00+0p00.fits"
},
{
  "type": "FILENAME",
  "content": "/srv/cutouts/1739176796504A/24_JCMT-HARP-COHRS_40p00_0p00_CUBE_REBIN_R1.fit"
},
```

merge

- FITS headers represent adjacent areas in the sky
- Merge possible if files are from the same sub-survey:
    - The same survey (obs_collection) and the same species (obs frequency)
- Plan:
    - Extra column in search result VOTable: ‘merge_ids’
        - Set of mergeable data-id’s
        - Associate with merge service descriptor VOTable in search result
    - Implement as async/uws SODA call with multiplicity on ID
    - Request MIME-type:
        - RESPONSEFORMAT=application/fits;type=mosaic

SKA challenge

- Huge files
- Cutout is crucial
  - SODA
    - Posix local access
    - Davix remote cutout
- Davix: open source client for WebDav
  - C++, by CERN
- Storage access xrootd/http
  - If server does not support multi-range: libdavix simulates by single range-requests in parallel (default 10) connections



A&A

Metadata (search result) is public, data access (cutouts) is protected.

- Most of datasets are PUBLIC
- Non-authenticated access supported for PUBLIC data
- Connected to IA2 auth infrastructure, Rap-based, delivers user-groups
- OIDC
  - access_token validation
  - using Rap adapter (in-house)
- Authz:
  - One group for each private survey
  - groups based access

Technologies and packaging

- Interfaces (VO) → J2EE/Servlets
  - SIAv2: pgSphere/PostgreSQL
  - SODA: auth-adapters J2EE/Filters
- Computations (C++):
  - Overlap, coordinates: ASTlib
  - FITS access: cfitsio
  - Merge (demosaic): Montage
- Docker images
  - docker-compose/kubernetes point to:
    - git.ia2.inaf.it:5050/vialactea/vlkb-siav2/siav2:1.6.10
    - git.ia2.inaf.it:5050/vialactea/vlkb-soda/soda:1.7.14
  - Config DB-URI and FITS-files storage path, optionally auth