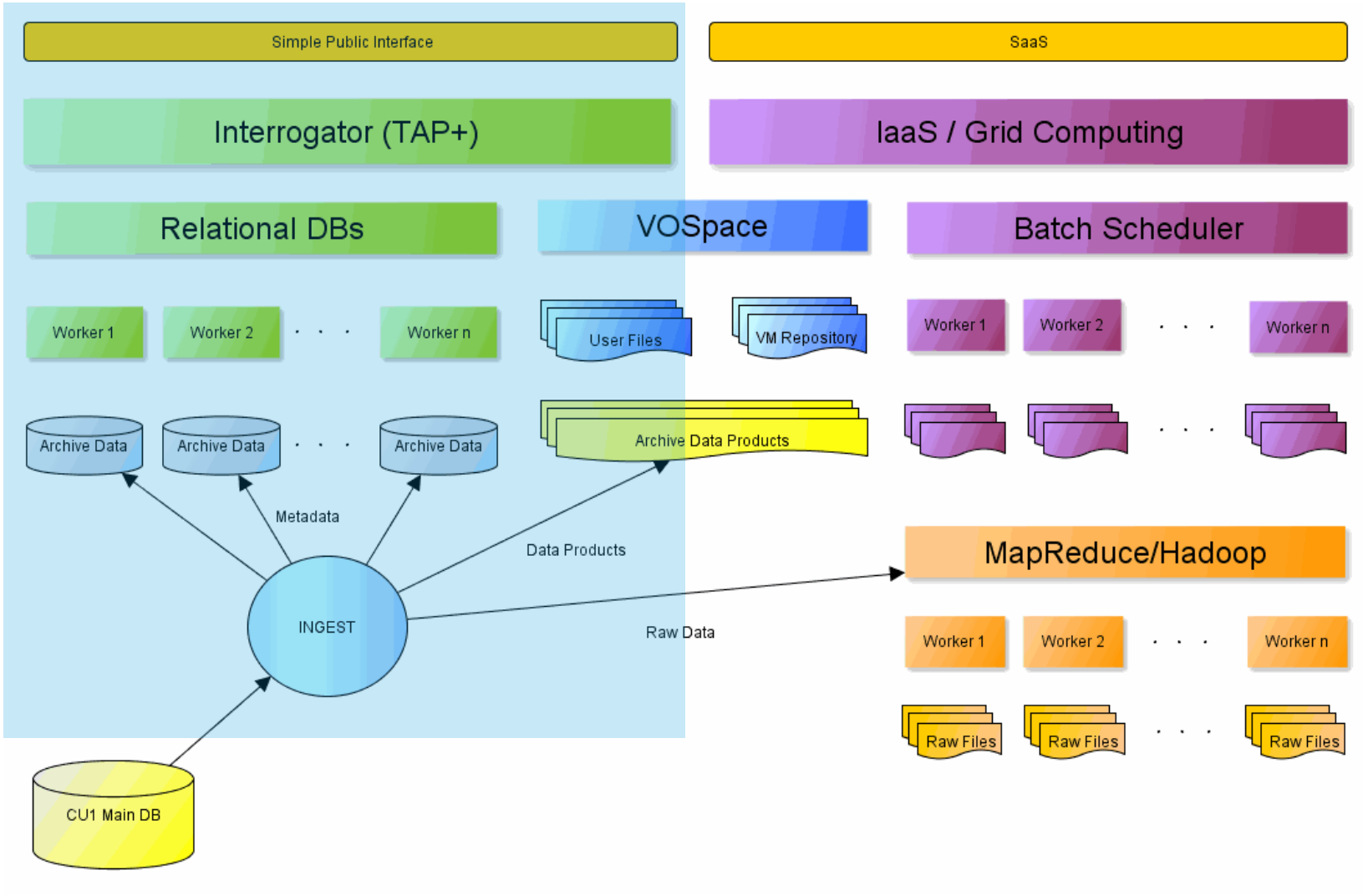
Gaia Archive Requirements: TAP+

Jesús Salgado Jesus.Salgado@sciops.esa.int
and the ESA Science Archives Team / Gaia Archive

Science Archives Team and VO Team/ESA
19/05/2014

Gaia Core Systems Architecture



Gaia Archive requirements for TAP+

Private data

- Users should have their own data/catalogues close to the archive

Gaia DM

- Gaia has its own DM that is propagated and used for all the software developed in the consortium

Crossmatch

- Crossmatch between Gaia catalogue and user catalogues to be executed as a user ADQL function

Big output results

- TAP server should be able to handle very big catalogues and very big output responses

TAP+ implementation: Private data

Jobs Visibility (TAP/UWS)

- ADQL queries are a description of ongoing scientific work so they have to be hidden
- Jobs should allow a “non anonymous” execution
- ADQL queries inside the jobs under this mode should be only accessible by the owner

Persistent upload (TAP)

- Data-centric architecture impose users data close to the server
- TAP Upload using workflow UPLOAD/QUERY/DELETE
- Tables uploaded by user should be maintained at the server in a persistent way
- Login/Authorization/server capabilities/sharing (VOSpace?)

TAP Schema (TAP)

- Queries should be allowed on TAP schema tables
- Some tables are user tables but they cannot be public
- Intermediate solution (content hidden, users tables names public). Is this acceptable?

TAP+ implementation: Gaia DM

DM Preserved in TAP_SCHEMA? (TAP?)

- Gaia DM should be preserved and propagated through TAP response
- Should TAP_SCHEMA reflect this DM? How?
- DM exploration? Cleaner/self-consistent solution?

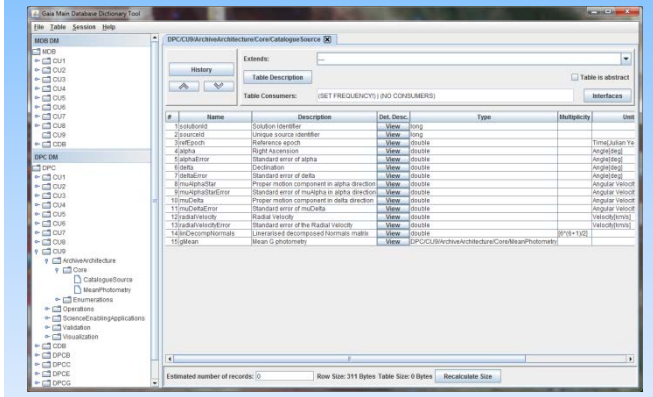
Output response OO complaint (TAP?)

- Compulsory output of TAP response is a single table VOTable
- Other output formats are allowed that could reflect better the Gaia DM (Gaia bin – a.k.a. gbin) so they should probably be implemented
- In the case of VOTable format, a VO-DML could/would be implemented

ADQL-VOTable or other languages/responses (ADQL?/TAP)

- Although TAP allows different query languages, we are trying to use only ADQL to query the Gaia catalogues
- Relational based language/tabular result so mapping techniques needed
- Could we need an OO query language and OO response to simplify this DM propagation?

Gaia DM propagation



Gaia DM

DB Schema Creator

gbin

Gaia Archive TAP server

- VO-DML Annotations in TAP_SCHEMA?
- Different structure?

TAP_SCHEMA

```
<TABLE>
<GROUP utype="vo-dml:Instance.root" >
  <PARAM      name="type"      utype="vo-dml:Instance.type"      value="src:source.Source"
datatype="char" arraysize="*"/>
  <FIELDref ref="_designation" utype=" vo-dml:ObjectType.ID"/>
  <FIELDref ref="_designation" utype="src:source.Source.name"/>
  <PARAM name="type" utype="src:source.Source.classification" value="galaxy">
  <VALUES><OPTION value="galaxy"/><OPTION value="star"/>...</VALUES></PARAM>
  <GROUP utype="src:source.Source.position">
    <PARAM    name="type"    utype="vo-dml:Instance.type"    value="src:source.SkyCoordinate"
datatype="char" arraysize="*"/>
    <FIELDref ref="_ra" utype="src:source.SkyCoordinate.longitude"/>
    <FIELDref ref="_dec" utype="src:source.SkyCoordinate.latitude"/>
    <GROUP ref="_icrs" utype="src:source.SkyCoordinate.frame"/>
  </GROUP>
</GROUP>
<FIELD name="designation" ID="_designation" .../>
<FIELD name="ra" ID="_ra" unit="deg" .../>
<FIELD name="dec" ID="_dec"  unit="deg" .../>
<TR><TD>08120809-0206132</TD><TD>123.033734</TD><TD>-2.103671</TD></TR>
...
</TABLE>
```

VOTable response with VO-DML

TAP+ implementation: Crossmatch

No crossmatch support in ADQL (ADQL?)

- Long debate in the past about this
- Crossmatch strategies
- UDF are allowed to be exposed
- Requirement for Gaia
- OK for complex crossmatches but... close-neighbours?
- distance_match(point, point, double)
- “CREATE AS”?

Current possible approach

DISTANCE(POINT('ICRS',a.ra, a.dec), POINT('ICRS',b.ra, b.dec)) < 1.0/3600

(difficult to handle at server side)

Option a

CREATE TABLE user_jsalgado.my_xmatch_table AS

SELECT * FROM public.g10_mw as a, public.igsl_source as b

WHERE

DISTANCE_MATCH(POINT('ICRS',a.ra,a.dec),POINT('ICRS',b.ra,b.dec), 1.0/3600)

Option b

CREATE TABLE user_jsalgado.my_xmatch_table AS

SELECT * FROM public.g10_mw as a, public.igsl_source as b

WHERE

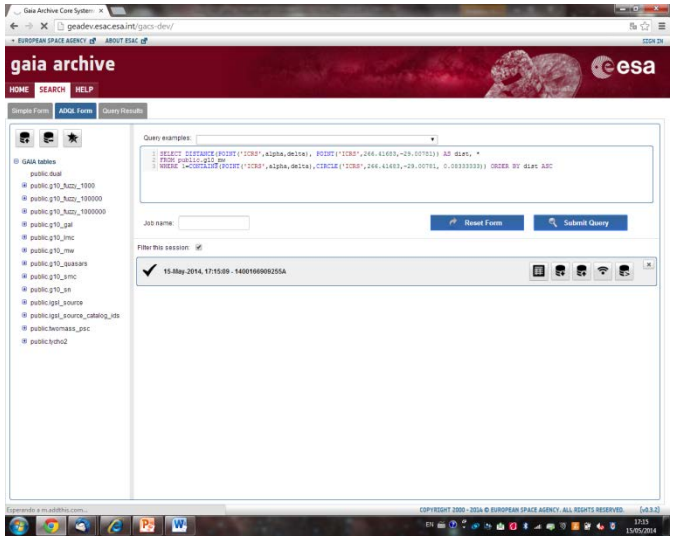
DISTANCE_MATCH(a.point,b.point, 1.0/3600)

TAP+ implementation: Big output results

Pagination (TAP?/ADQL?)

- Not clear how to handle big results
- Hard limit and “soft” limit already implemented in TAP
- User requirements to “navigate” results (not only TAP)
- Two approaches:
  - Table Pagination
  - Infinite scrolling
- Both approaches could imply specific pagination keywords
- If agreed,... where to add this support? TAP input? ADQL sentence?

TAP+ implementation: Short Demo



THANK YOU

Jesus Salgado

Jesus.Salgado@sciops.esa.int