

The Gaia Archive at ESAC: a VO-inside Archive

J. Gonzalez-Núñez – jgonzale@sciops.esa.int

*J. Salgado, R. Gutiérrez-Sánchez, J.C. Segovia, J. Duran, B. Merín,
S. Nieto, J. Hernández, C. Arviset, W. O'Mullane*

25/11/2015

Issue/Revision: 1.0

Reference: Gaia Archive – VO inside

Status: Issued

ESA UNCLASSIFIED - Releasable to the Public



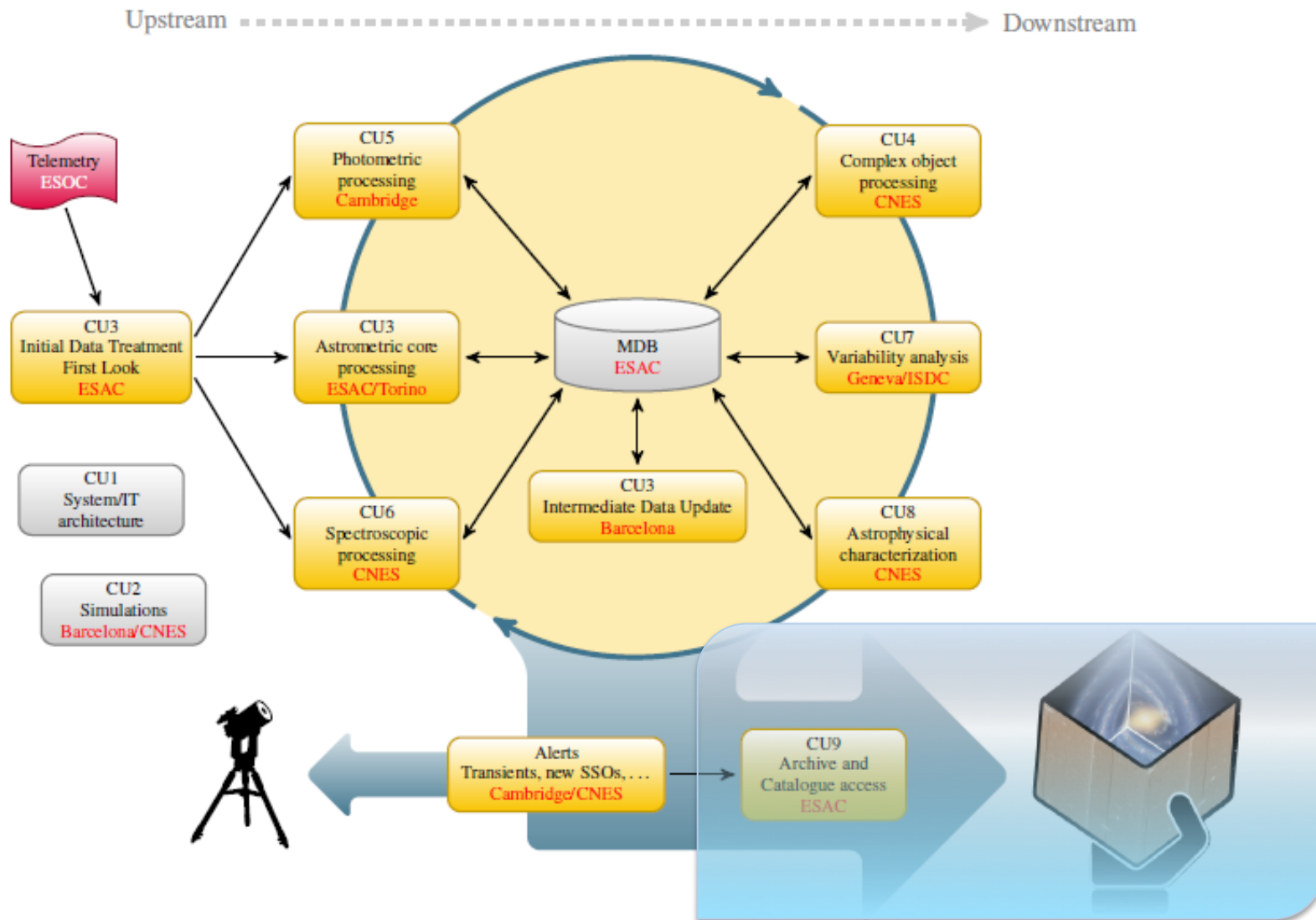
gaia

European Space Agency

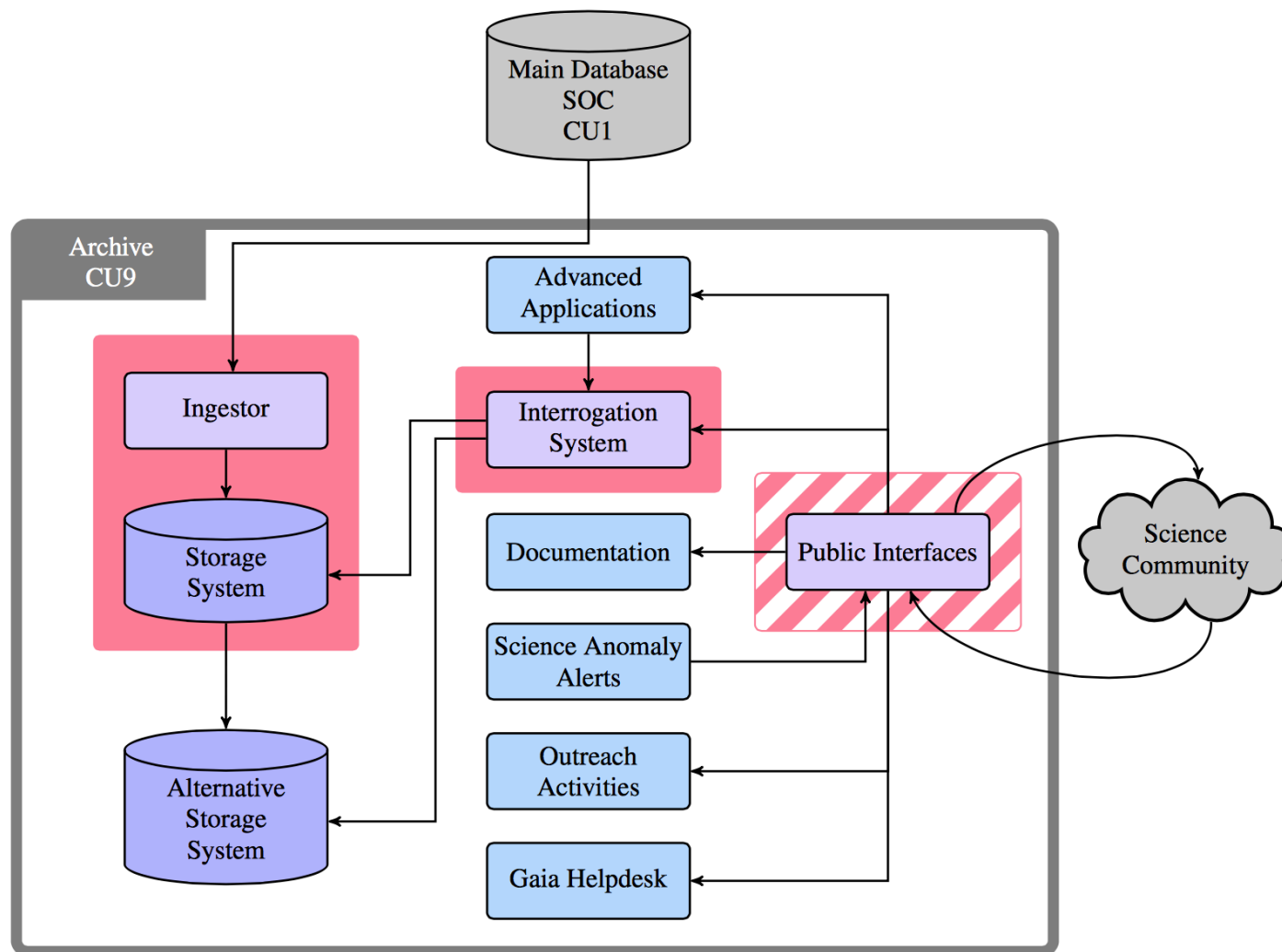
- Homogeneity through common Data Models
- Interoperability through Open Protocols
- Transparent access to Archives and Data providers worldwide



The DPAC consortium



The DPAC consortium



Querying Gaia catalogues:

TAP and UWS

[EUROPEAN SPACE AGENCY](#)
[ABOUT ESAC](#)

Juan Gonzalez (jgonza01)

gaia archive

[HOME](#)
[SEARCH](#)
[STATISTICS](#)
[HELP](#)
[DOCUMENTATION](#)
[SHARE](#)
[ADMIN](#)

Simple Form

ADQL Form

Query Results

public.twomass_psc

public.tycho2

public.ucac4_original_valid

User tables

user_jgonza01.table1

user_jgonza01.test_mio

user_jgonza01.trialUpXmatch

user_jgonza01.trialUpXmatch3

user_jgonza01.trialUpXmatch4

user_jgonza01.trialUpXmatch

user_jgonza01.xmatch_test

user_jgonza01.xmatch_twomass_psc_table1

Shared to me (from satgaia)

user_satgaia.igsl_tmass_15s_bestneighbour

user_satgaia.igsl_tmass_15s_neighbourhood

user_satgaia.igsl_tmass_1s_bestneighbour

user_satgaia.igsl_tmass_1s_neighbourhood

user_satgaia.igsl_tmass_5s_bestneighbour

Query examples:

```

1 SELECT DISTANCE(POINT('ICRS',alpha,delta),
2 POINT('ICRS',266.41683,-29.00781)) AS dist, *
3 FROM public.gog_cataloguesource
WHERE
1=CONTAINS(POINT('ICRS',alpha,delta),CIRCLE('ICRS',266.41683,-29.00781,
0.08333333)) ORDER BY dist ASC

```

Job name:

Reset Form

Submit Query

Filter by job name

Refresh

Select all jobs

Delete selected jobs

✓

02-Feb-2015, 11:43:48 - 1422873828311D

Rows: 100, Size: 39 KB

✓

09-Jun-2015, 18:59:51 - 1433869190582D

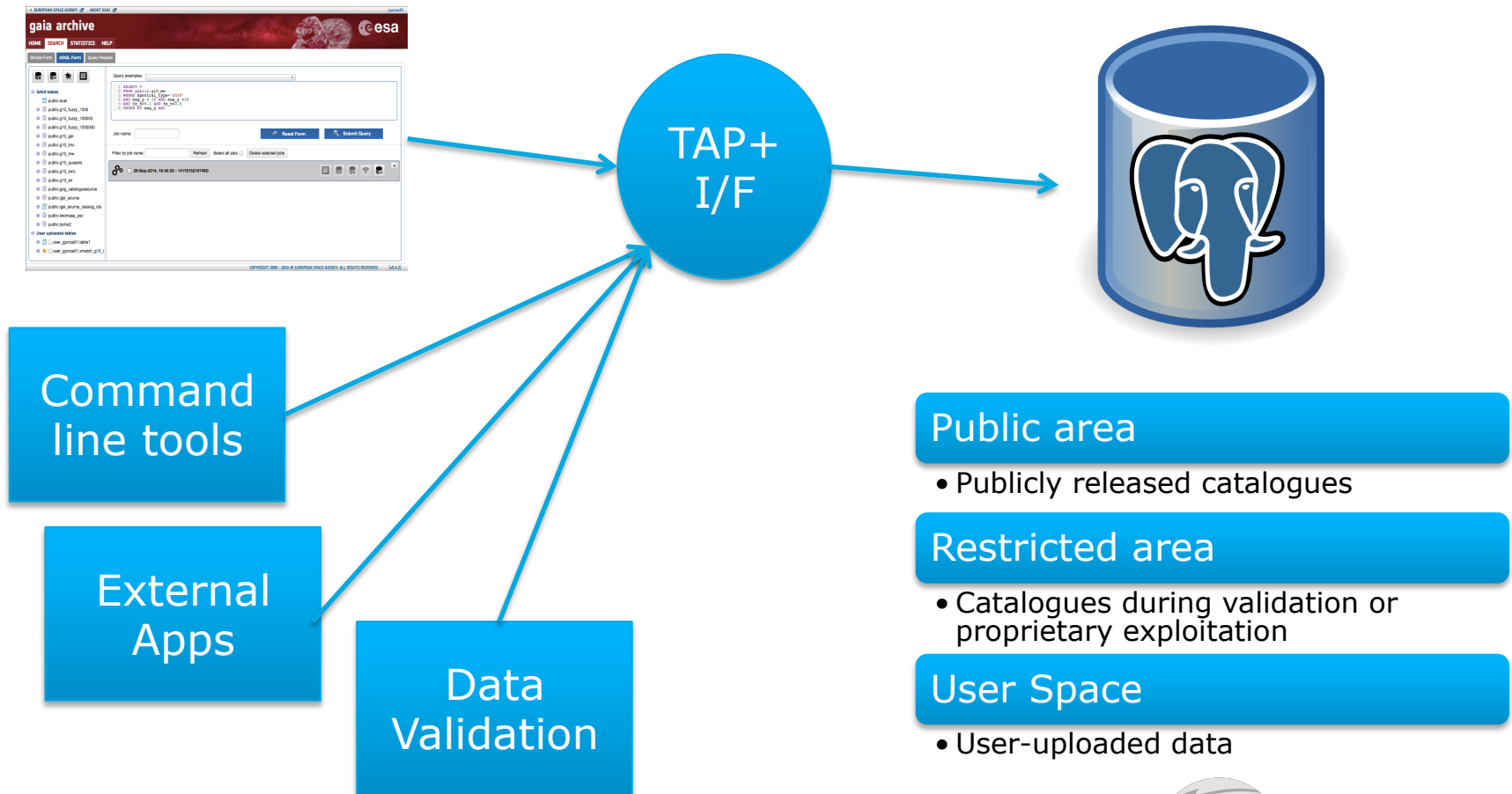
Rows: 4266, Size: 1 MB

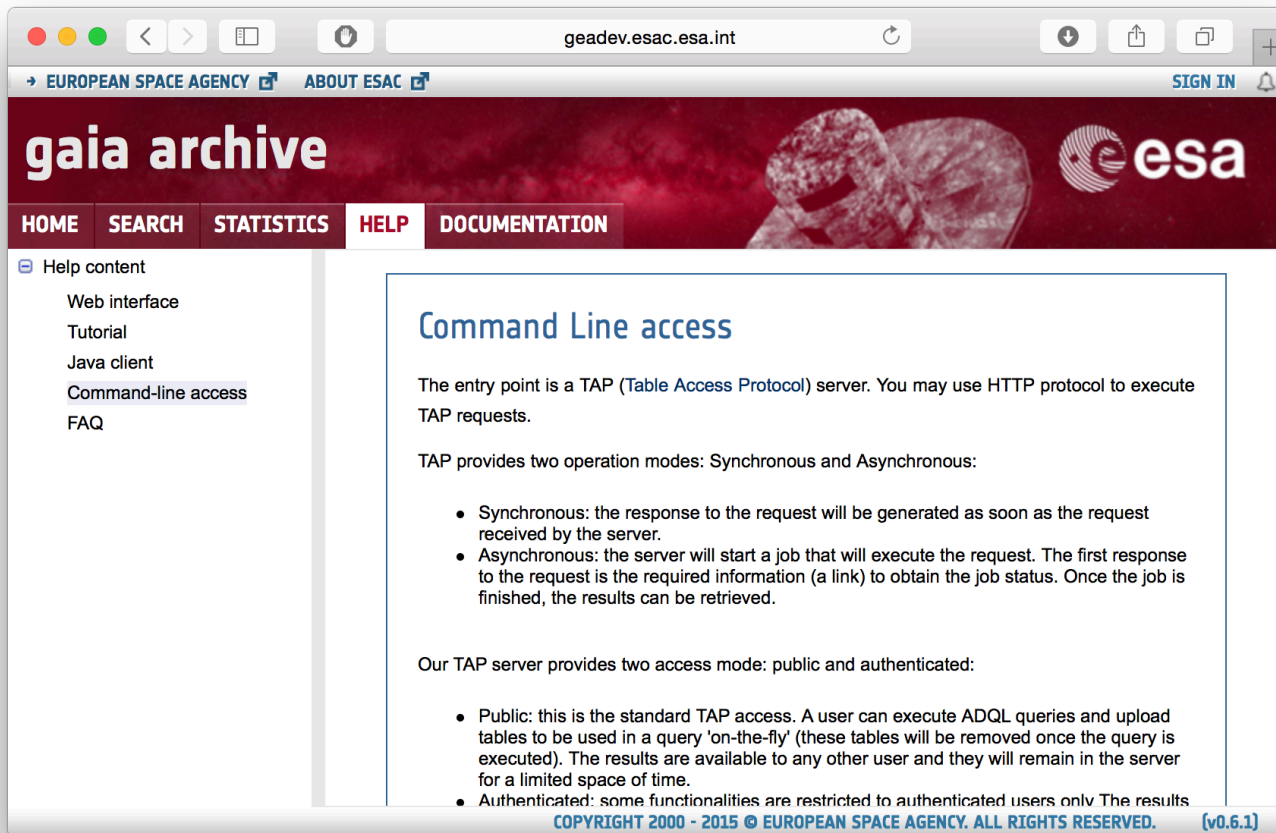
COPYRIGHT 2000 - 2015 © EUROPEAN SPACE AGENCY. ALL RIGHTS RESERVED.

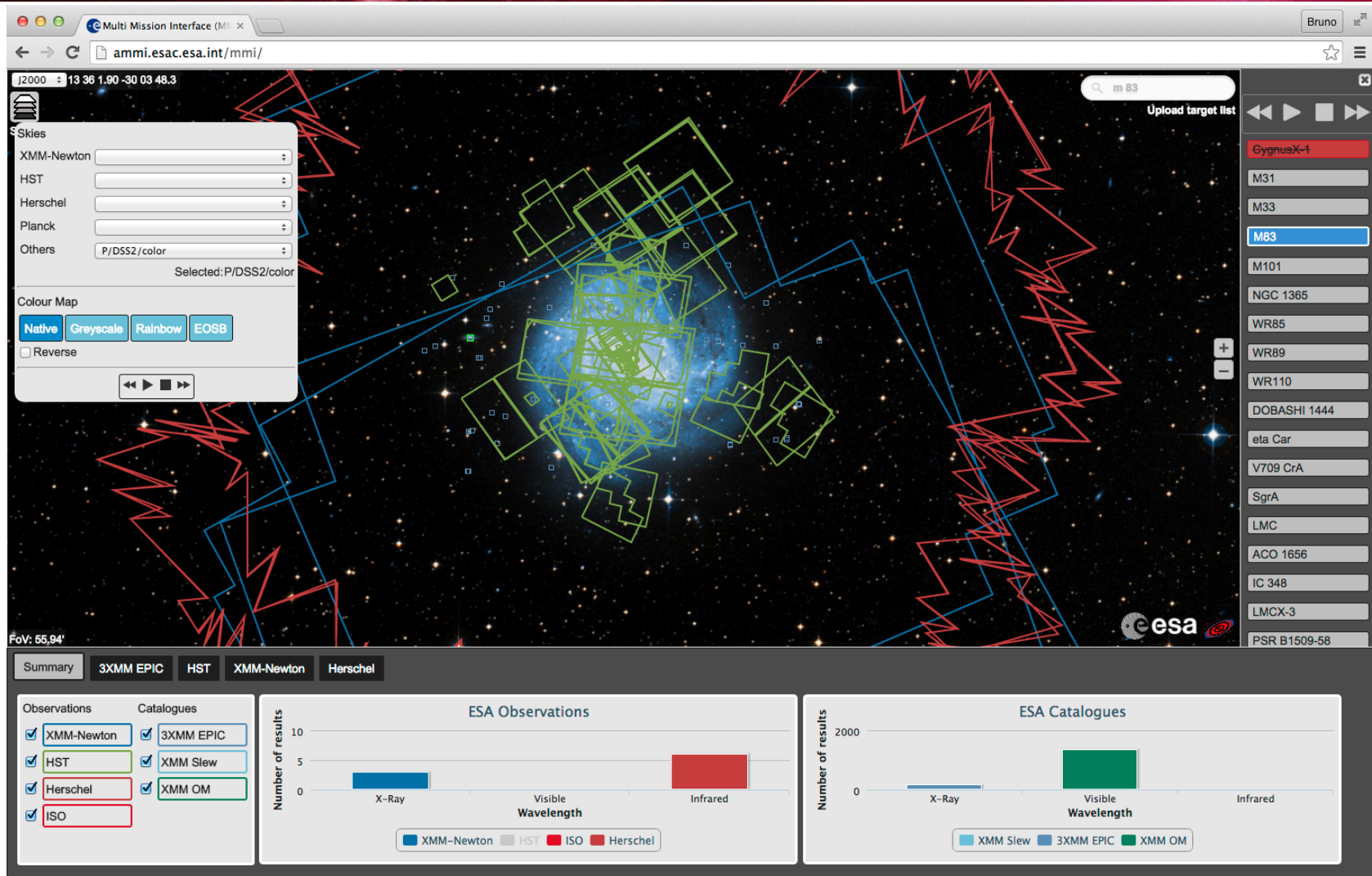
[v0.5.3]

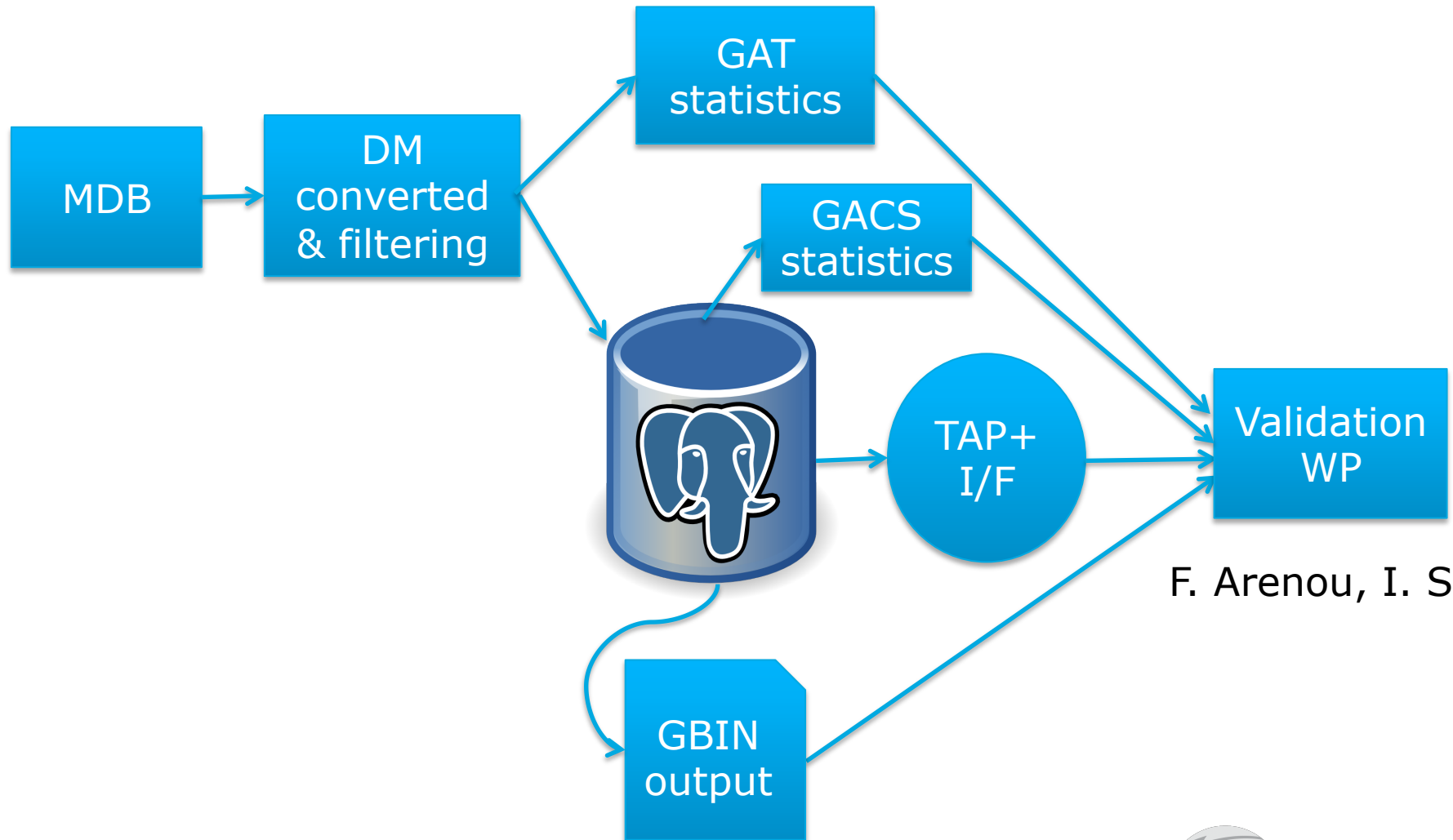


Open APIs



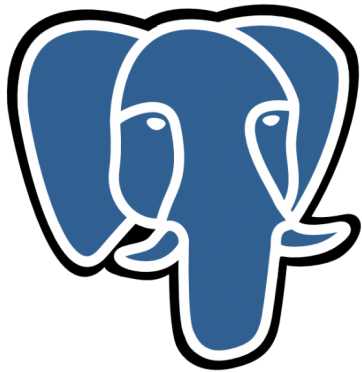






F. Arenou, I. Shih

PostgreSQL + PgSphere + Q3C



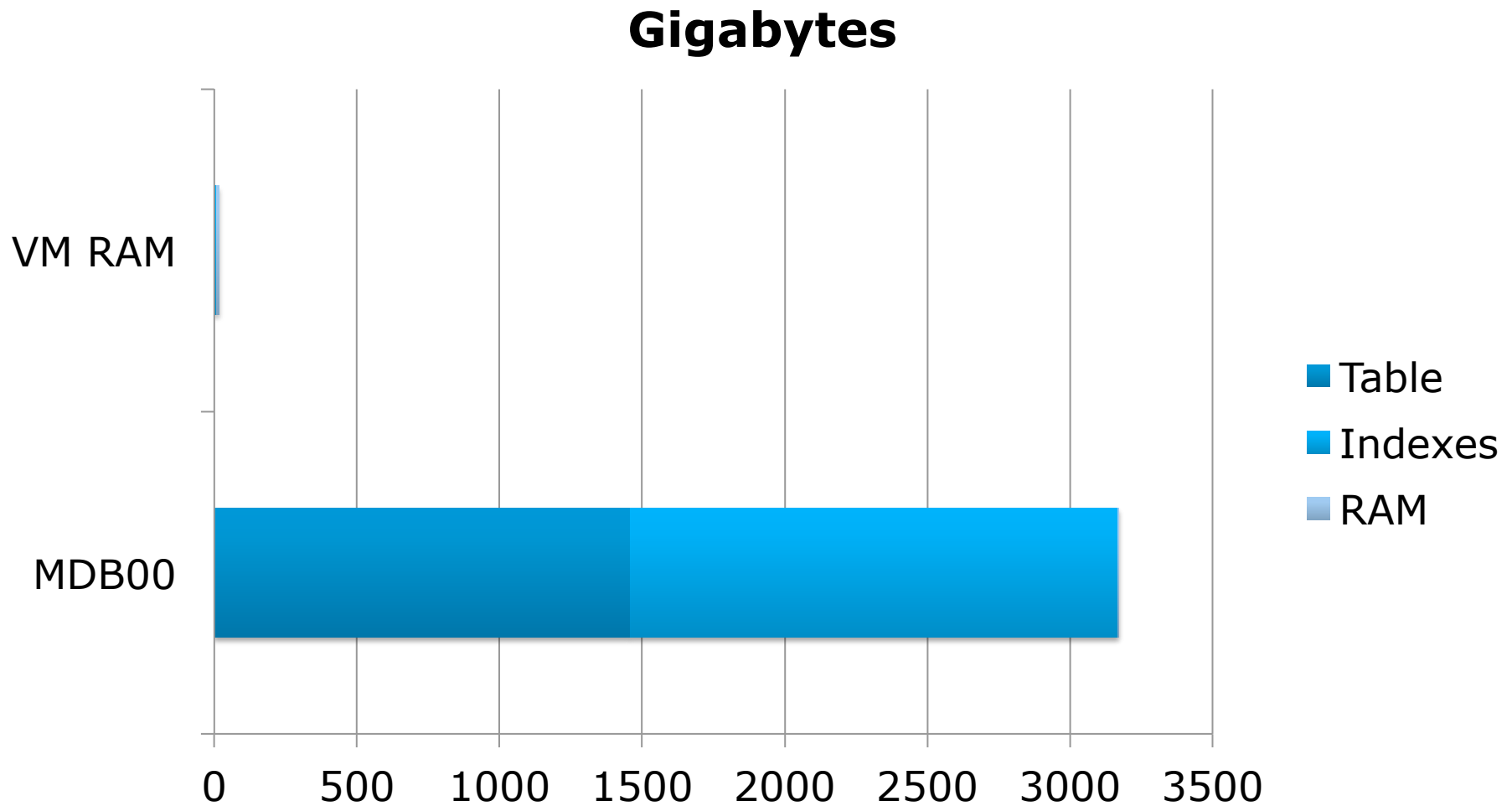
PostgreSQL

“The world’s most advanced open source database”

PostgreSQL Global Dev Group



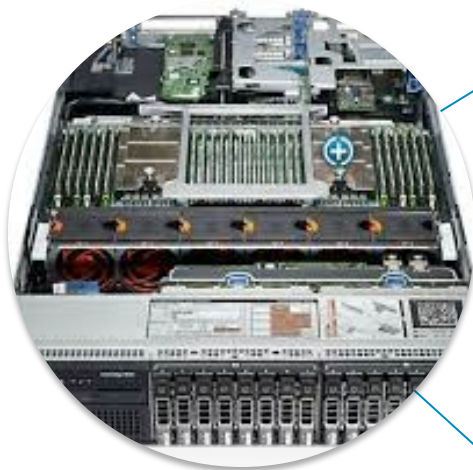
Multibillion catalogues @ GACS



Dell R820

1TB RAM

4 CPUs
32 Cores



PCIe
SSDs

- 2 x 1.21 TB
- Fast Random IO

SAS
HDDs

- 14x900GB
- Fast Sequential IO

NetAPP
NAS

- ~Unlimited
- 10 Gbps
- HA, DR, etc.

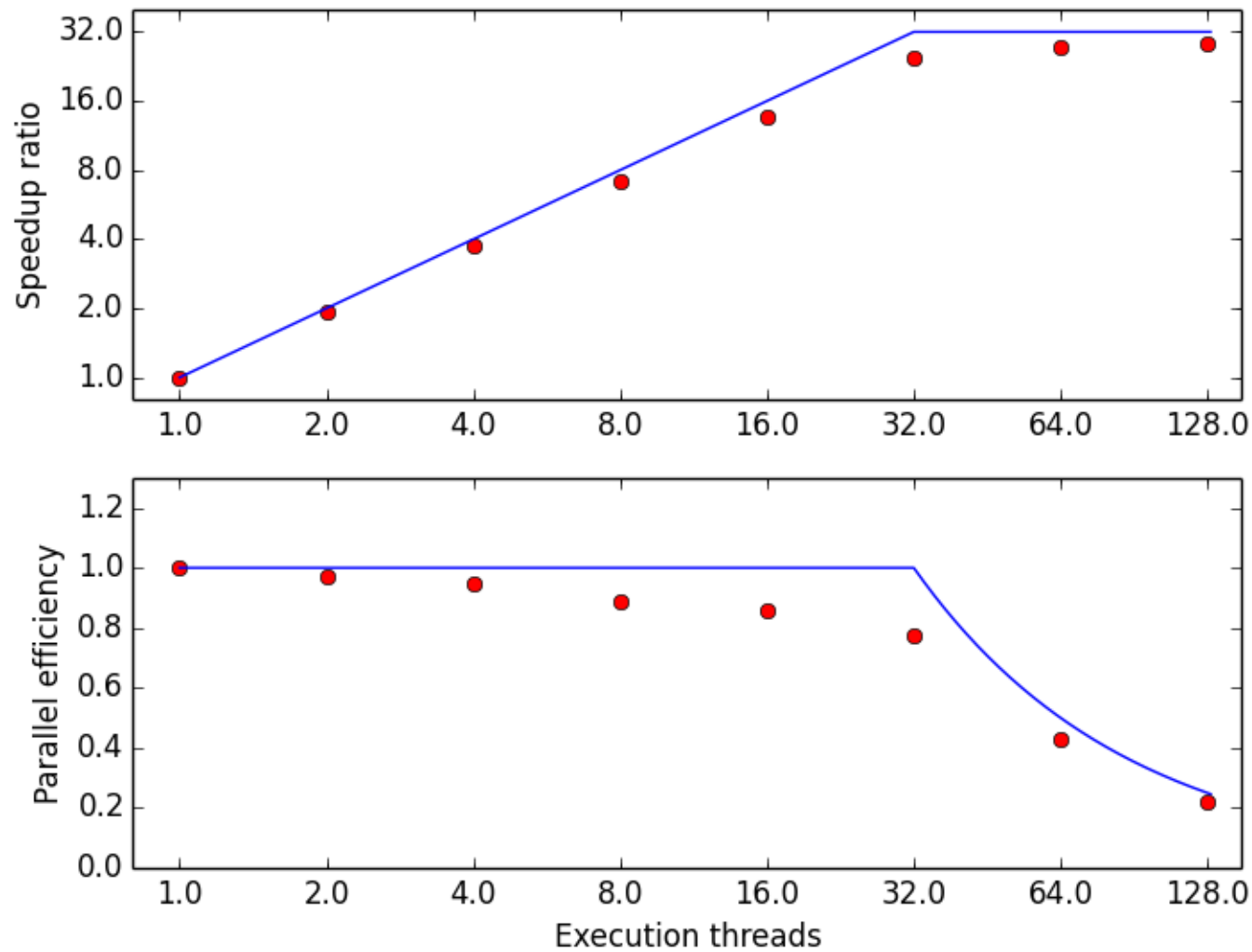
Data selection: Cone Search to GUMS Milky Way table (2.2E9 sources)

```
SELECT pos <-> spoint'(266.41683d, -29.00781d)' AS
dist, *
FROM gl0_ss
WHERE pos @ scircle'<(266.41683d, -29.00781d), Rd>'
ORDER BY dist ASC
```

PgBench Throughput: random positions

Search Radius	Transaction s/s	Avg service time	Output
1'	9430	2.5 ms	169 rows
5'	451	54 ms	4733 rows
10 deg	5.8	3.9 s	720K rows

TAP & UWS

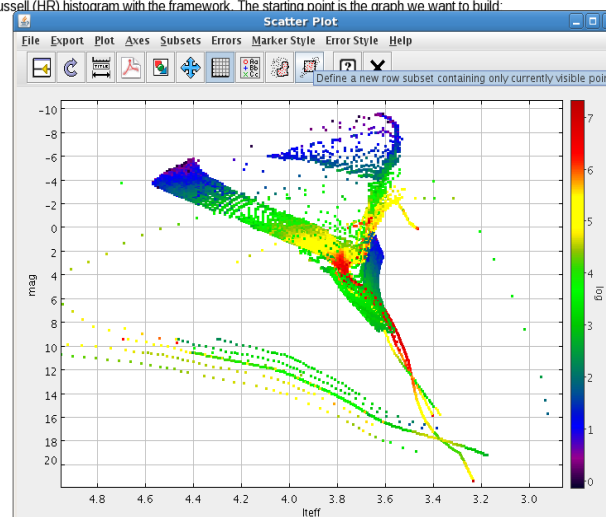
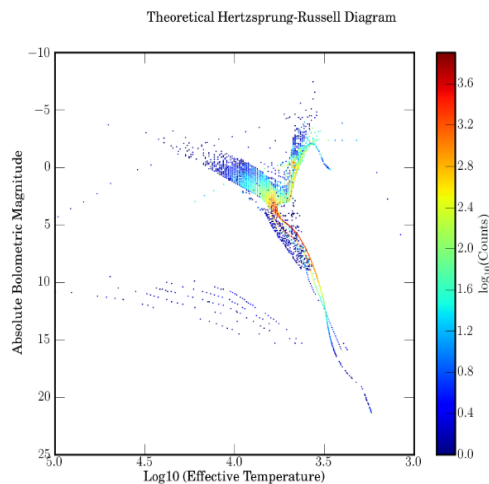


H-R Diagram (GUMS, 2.1 Billion sources)

Threads	Map	Reduce
10	14 min	<10 s
20	7 min	<10 s
30	5 min	<10 s

Theoretical Hertzsprung-Russell histogram example

In this section, we will go through the steps needed for computing a theoretical Hertzsprung-Russell (HR) histogram with the framework. The starting point is the graph we want to build:



TAP+?

- Started dev on TAP 1.0 library (G. Mantelet)
- Validation
 - Part of the Gaia catalogues validation will **run through the Archive TAP service**
 - Catalogues shall **only be available to validation teams** during this period

➤ **Authentication mechanisms**

- Log in/out
- Session handling

➤ **Persistent Upload**

- /upload user catalogues
- Schema per user

➤ **Table sharing**

- Creation of user groups
- Permissions handling

➤ **Quotas**

- DB/results storage limit

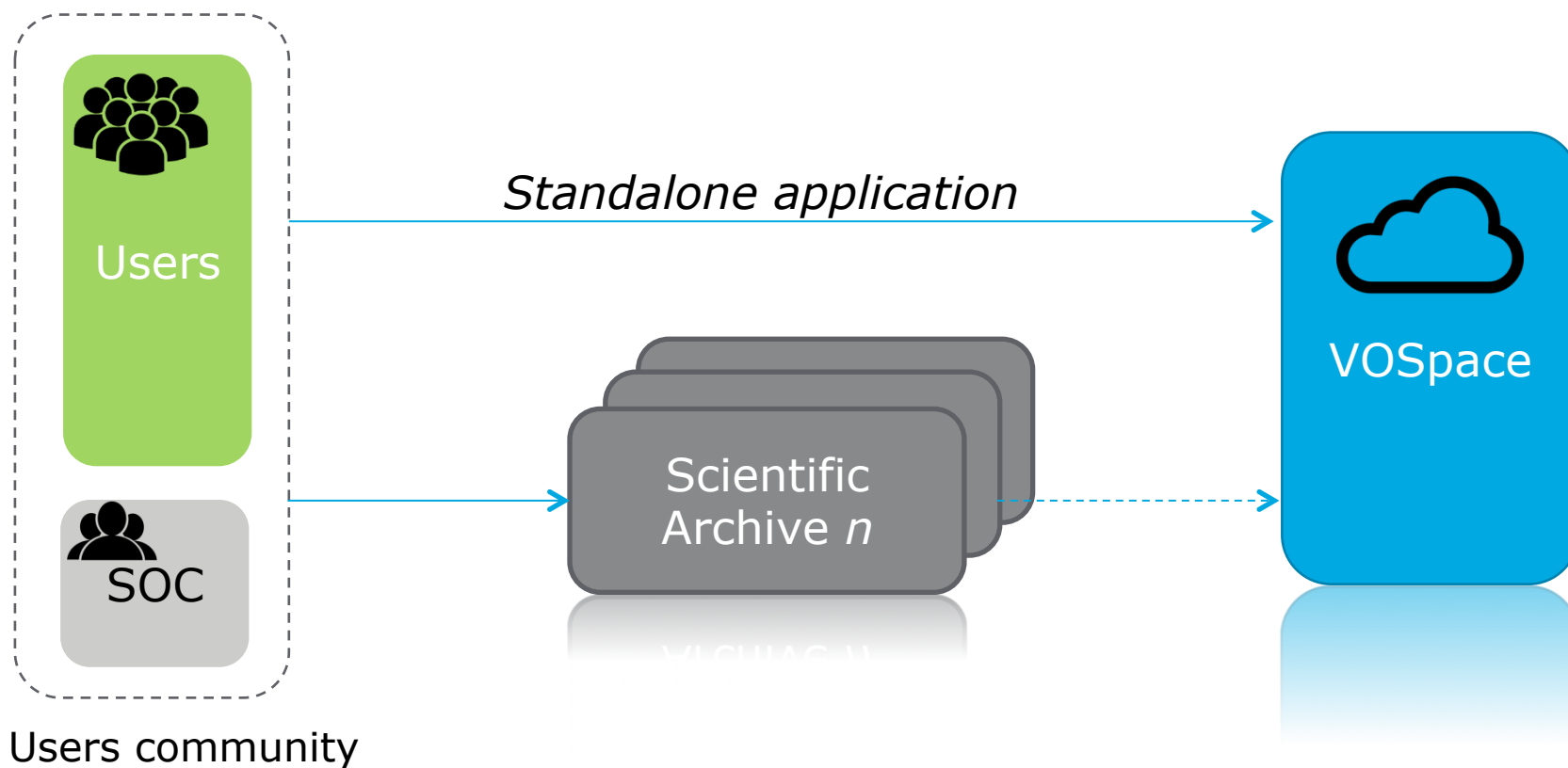
➤ **/tables filtering**

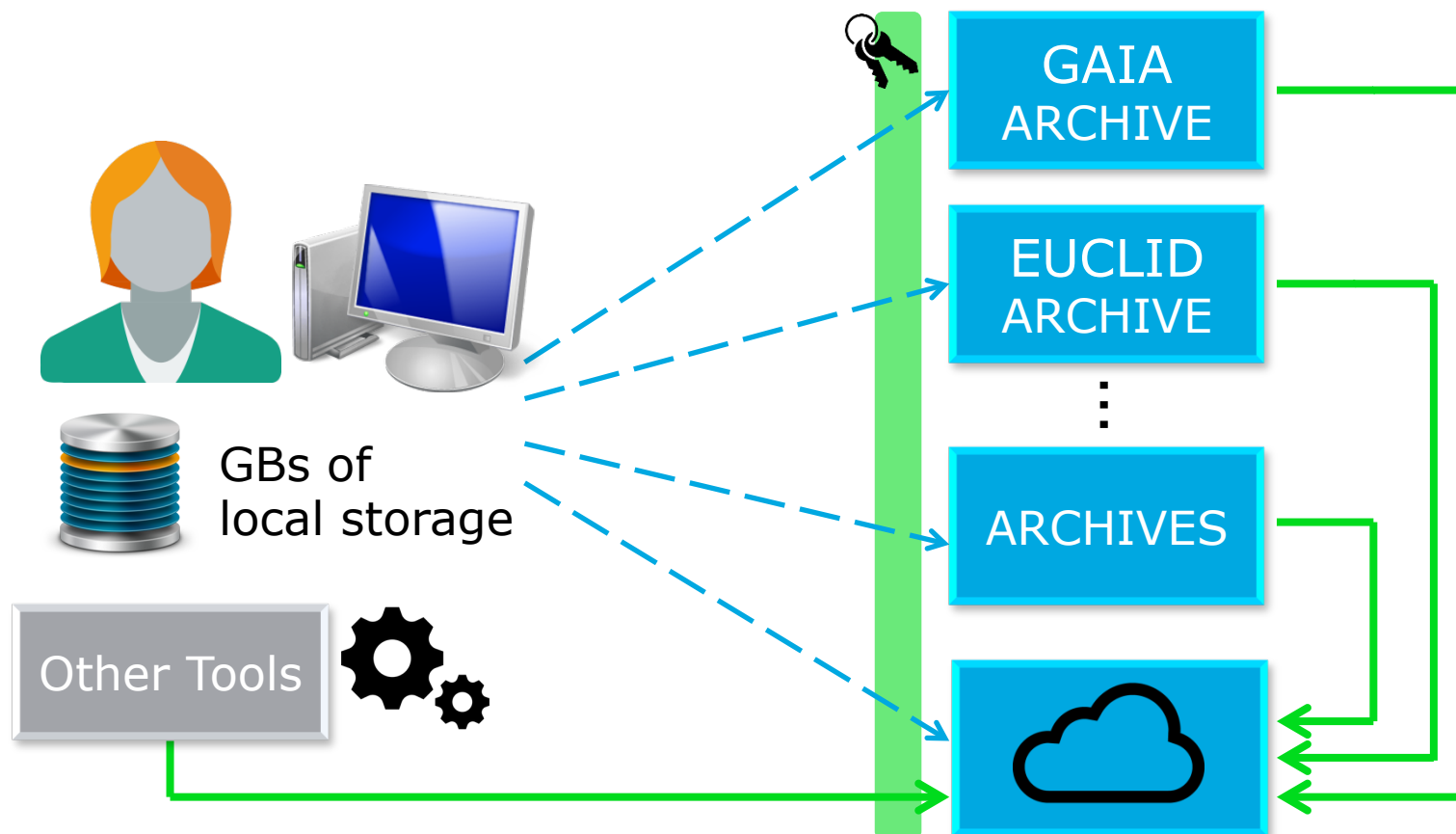
- Progressive browsing

➤ **User defined functions**

➤ **Notifications**

VOSpace





gaia archive

[HOME](#)
[SEARCH](#)
[STATISTICS](#)
[HELP](#)
[DOCUMENTATION](#)
[VOSPACE](#)
[SHARE](#)

[Simple Form](#)
[ADQL Form](#)
[Query Results](#)

Job name:
















[Query examples](#)

```

1 SELECT *
2 FROM public.igsl_source
3 WHERE spectral_type='GSIV'
4 AND mag_g > 15 AND mag_g < 16
5 AND fe_h>0.1 AND fe_h<0.5
6 ORDER BY mag_g ASC
    
```

[Reset Form](#)

[Submit Query](#)

Status	Job	Creation date	Num. rows	Size	
✓	<input type="checkbox"/> 1443861206344I	03-Oct-2015, 10:33:26	0	4 KB	    
✓	<input type="checkbox"/> 1443860851244I	03-Oct-2015, 10:27:31	0	0 KB	    
✓	<input type="checkbox"/> 1443860722274I	03-Oct-2015, 10:25:22	0	4 KB	    

GAIA send to VOSpace

Job 1443860722274I

Destination folder /carviset/GACS/

File name

☐ Overwrite file

[Cancel](#)











































[Send](#)

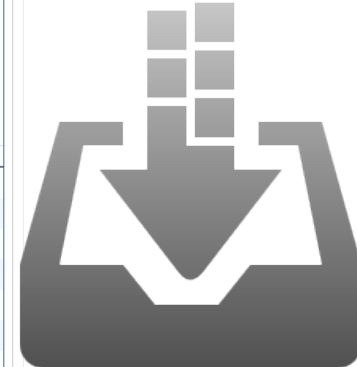
1-3 of 3

[Select all jobs](#)
[Delete selected jobs](#)

1


Reset Form [Submit Query](#)

Status	Job	Creation date	Num. rows	Size	
✓	 1444834077993JC	14-Oct-2015, 16:47:57	5	6 KB	     
✓	 1434544316098JC	17-Jun-2015, 14:46:16	3	0 KB	     
✓	 1434544147844JC	17-Jun-2015, 14:29:08	3	0 KB	     
✓	 1434447471777JC	16-Jun-2015, 11:37:51	3	0 KB	     
✓	 1434447400050JC	16-Jun-2015, 11:36:40	3	0 KB	     
✓	 1434447281271JC	16-Jun-2015, 11:34:42	3	0 KB	     



VOSpace Browser

snieto [Help](#) [Sign Out](#)

1.9 GB (3%) of 50GB



▼ My VOSpace

Euclid

for Pilar

Migration

Migration2

new-node

Security


Test

▼ Shared with me

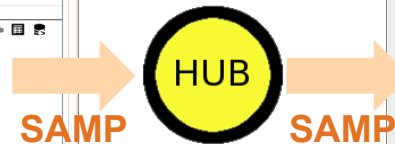
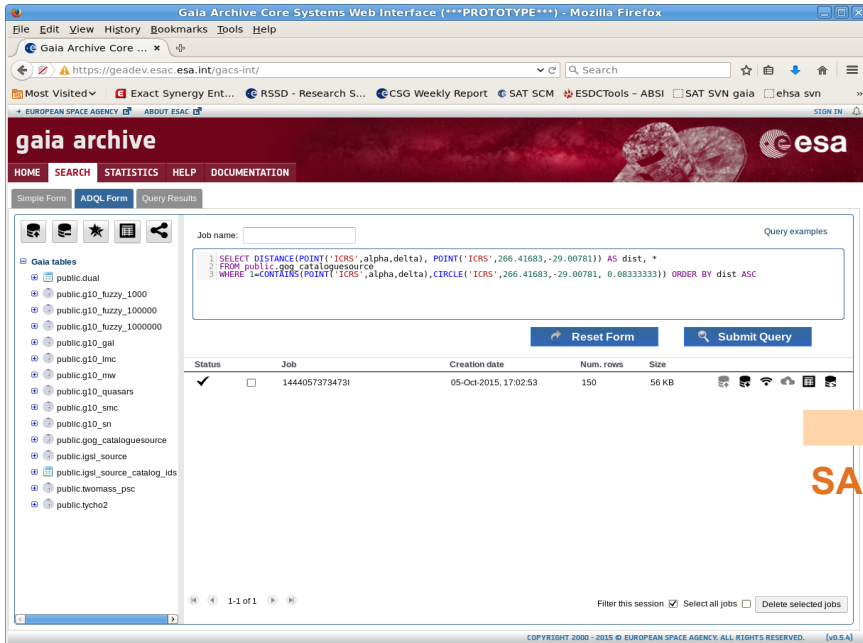
hola

test

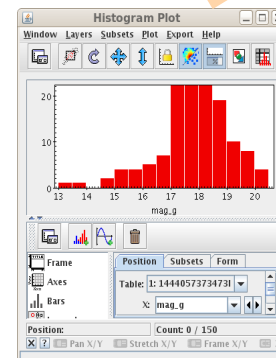
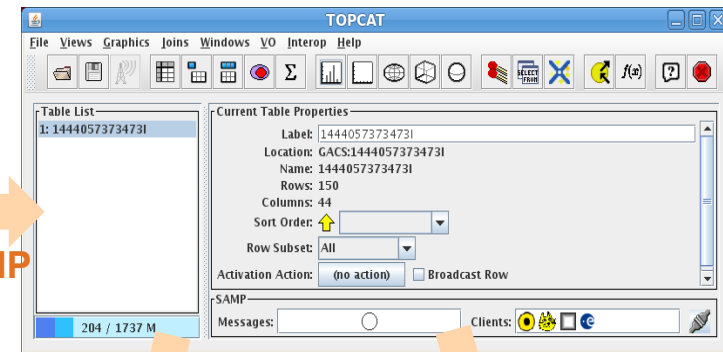
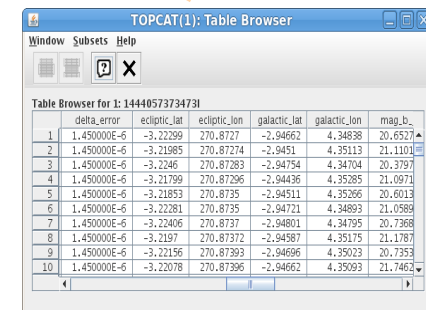
► [My VOSpace](#)

	TITLE	OWNER	SIZE	LAST MODIFIED
<input type="checkbox"/>	Test	Sara Nieto	--	06/13/2015 12:26:36
<input type="checkbox"/>	Security	Sara Nieto	--	06/13/2015 13:36:26
<input type="checkbox"/>	new-node	Sara Nieto	--	06/12/2015 17:14:50
<input type="checkbox"/>	Migration2	Sara Nieto	--	06/15/2015 21:09:15
<input type="checkbox"/>	Migration	Sara Nieto	--	06/15/2015 21:08:00
<input type="checkbox"/>	for Pilar	Sara Nieto	--	06/12/2015 14:17:52
<input checked="" type="checkbox"/>	Euclid	Sara Nieto	--	06/12/2015 14:21:42
<input type="checkbox"/>	 EAS-PL-ESDC-SNR-001.pdf	Sara Nieto	401.6 KB	07/14/2015 10:39:11
<input type="checkbox"/>	 1505.02291.pdf	Sara Nieto	1.5 MB	07/14/2015 10:39:36
<input type="checkbox"/>	 1433328333_coud.pdf	Sara Nieto	6.7 KB	06/14/2015 11:02:56

SAMP



VO Application : TOPCAT

	delta_error	ecliptic_lat	ecliptic_lon	galactic_lat	galactic_lon	mag_b
1	1.450000E-6	-3.22299	270.8727	-2.94662	4.34838	20.6527
2	1.450000E-6	-3.21985	270.87274	-2.9451	4.35113	21.1101
3	1.450000E-6	-3.2246	270.87283	-2.94754	4.34704	20.3797
4	1.450000E-6	-3.21799	270.87296	-2.94436	4.35285	21.0971
5	1.450000E-6	-3.21853	270.8735	-2.94511	4.35266	20.6013
6	1.450000E-6	-3.22281	270.8735	-2.94721	4.34893	21.0989
7	1.450000E-6	-3.22406	270.8737	-2.94801	4.34795	20.7368
8	1.450000E-6	-3.2197	270.87372	-2.94587	4.35175	21.1787
9	1.450000E-6	-3.22156	270.87393	-2.94696	4.35023	20.7353
10	1.450000E-6	-3.22078	270.87396	-2.94662	4.35093	21.7462

Source Data Model

1. Objective:

- a. Create a simple (but useful) SourceDM definition that could help in the publishing of catalogues through VO protocols
- b. Support on the definition for catalogues metadata for near-future missions (e.g. Gaia, Euclid,...) and allow the mapping for old ones
- c. Support interoperability operations between protocols (SED creation, crossmatch operations, etc)

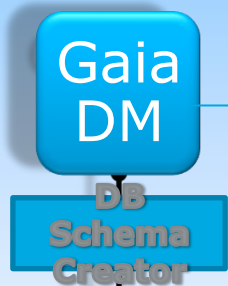
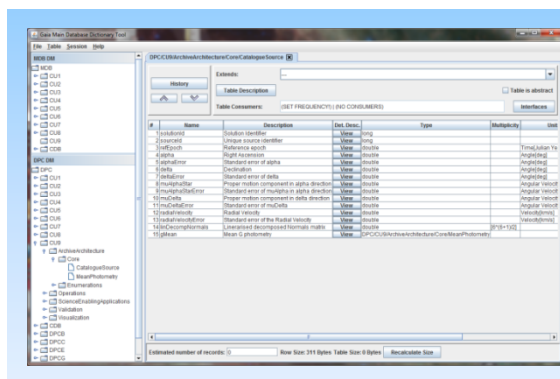
2. Initial Team:

- a. Jesús Salgado, Gerard Lemson, Sebastien Derrier, Laurent Michel, Bruno Merin, Tom Donaldson and Arnold Rots
- b. Open for contributors

3. Status:

- a. DM activities still quite packed
- b. Discussions at Gaia DPAC Consortium level

Gaia: DM in the core



```
<TABLE>
<GROUP type="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 type="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>
```

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

TAP_SCHEMA

VOTable response with VO-DML

Gaia Data Release Plans



Currently : Gaia Archive Core System (GACS) - internal access only to Gaia Consortium

- Available since early 2015
- Gaia simulations catalogues and some external catalogues

December 2015 : TGAS release - internal access only to Gaia Consortium

- Testing and verification phase until public release
- Most of GACS functionalities available

First Public release: summer 2016

- Positions (α , δ) and G-magnitudes for all stars with acceptable formal standard errors of positions. 90% of the sky can be covered
- Proper motion for Hipparcos stars

Further releases, full details at: <http://www.cosmos.esa.int/web/gaia/release>

- Second release : early 2017 (TBC), Third release: 2017/2018 (TBC), Fourth release: 2018/2019 (TBC)

Final release: 2022 (TBC)

- Full astrometric, photometric, and radial-velocity catalogues.
- All available variable-star and non-single-star solutions.
- Source classifications (probabilities) plus multiple astrophysical parameters (derived from BP/RP, RVS, and astrometry) for stars, unresolved binaries, galaxies, and quasars. Some parameters may not be available for faint(er) stars.
- An exo-planet list.
- All epoch and transit data for all sources.
- All ground-based observations made for data-processing purposes.



Thank you!

