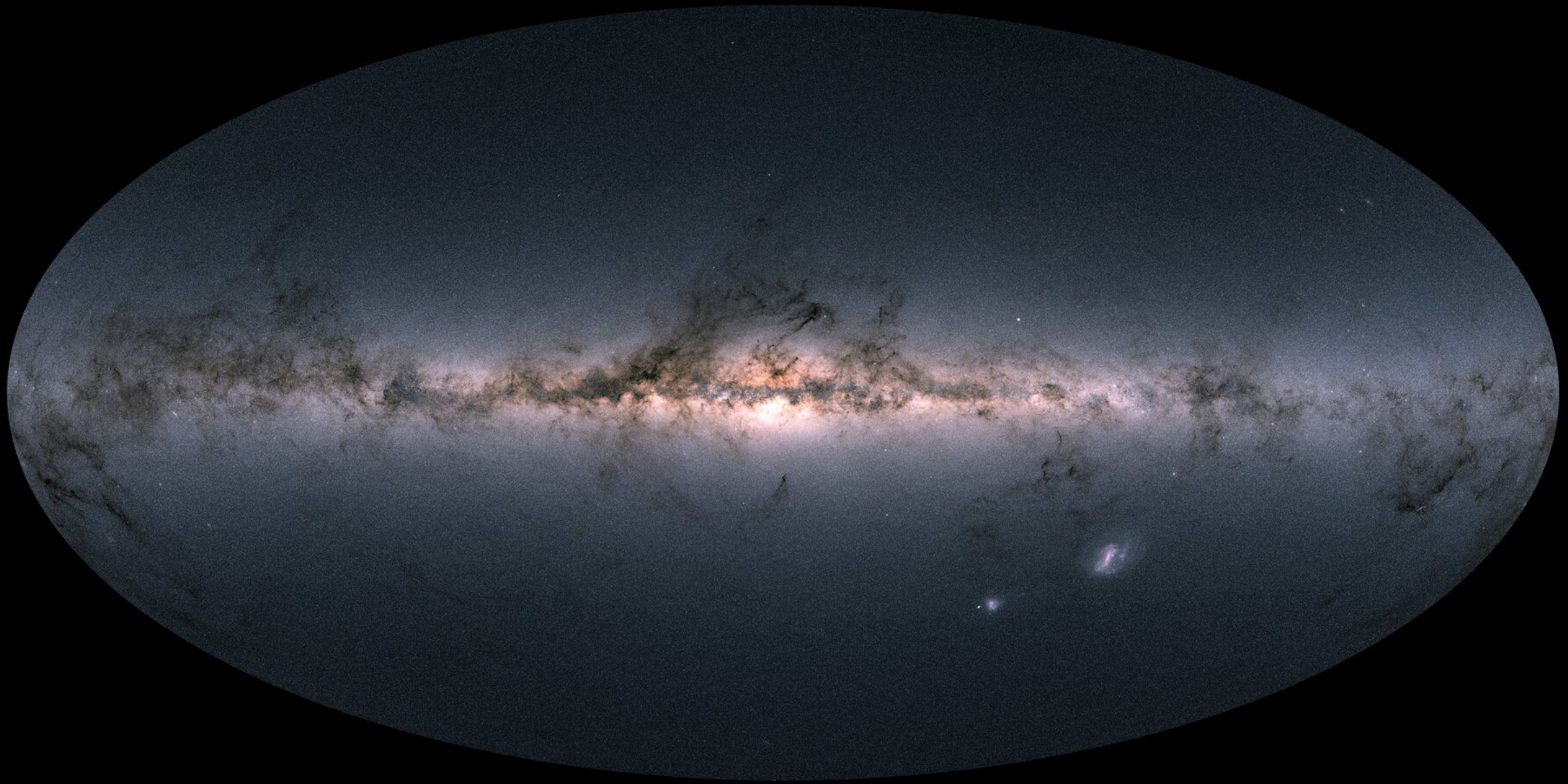
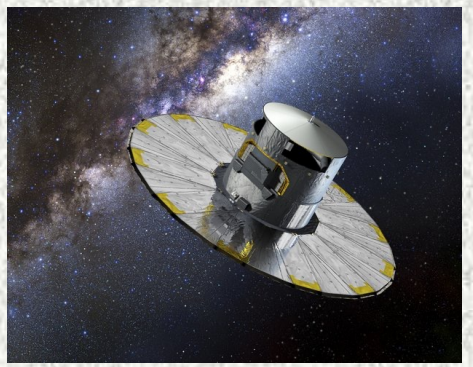
GREENPLUM IS NO LONGER OSS: CHANGE OF OPERATIONS IN MID-FLIGHT

Joaquim Oliveira , Pilar de Teodoro

PgConf'24 October 2024



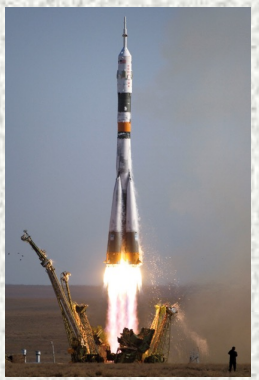
Data Flow



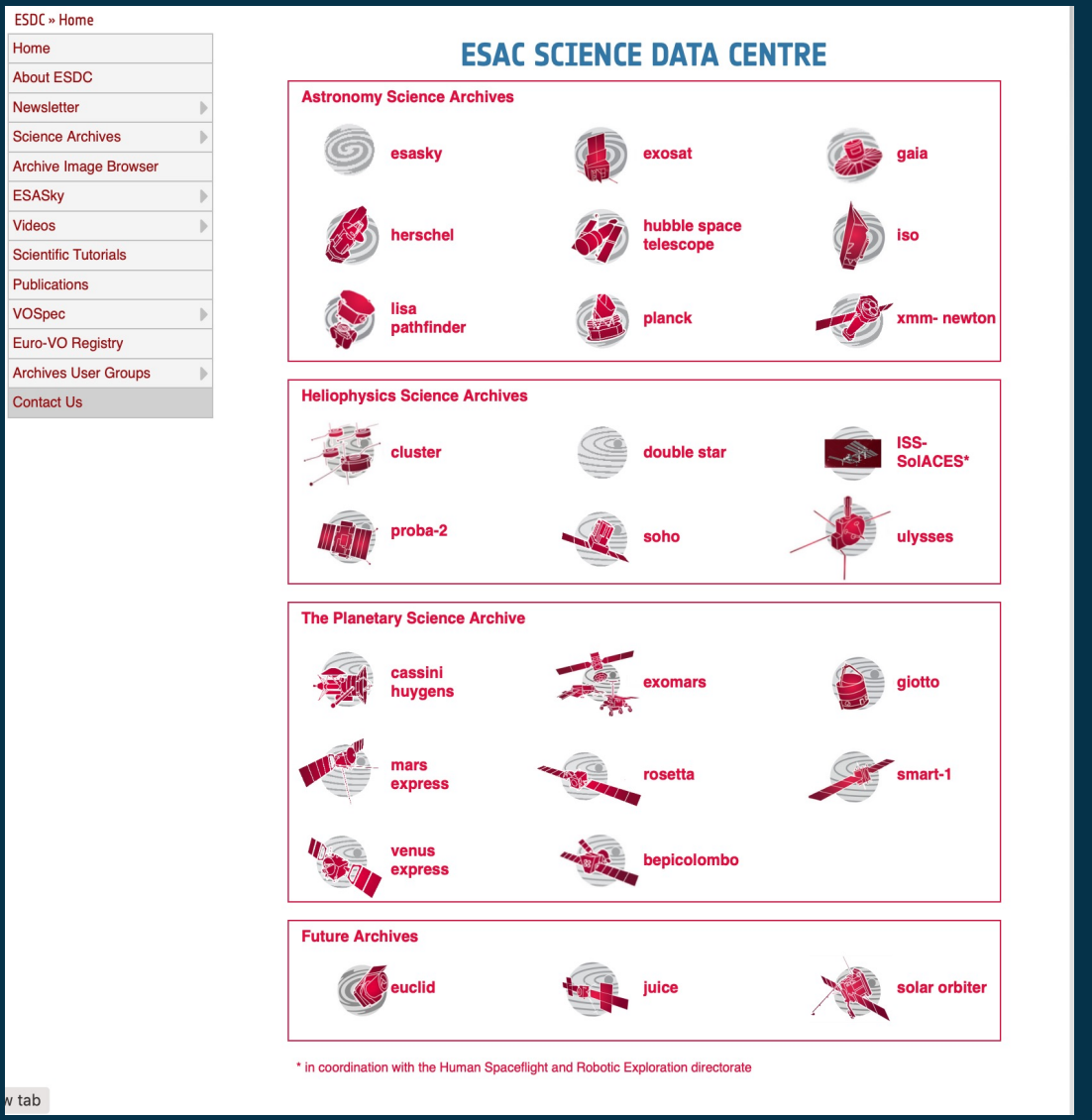
Mission Operations Center
ESOC: Germany

Science Operations Center
ESAC: Spain

ESAC Science Data Center
ESAC: Spain



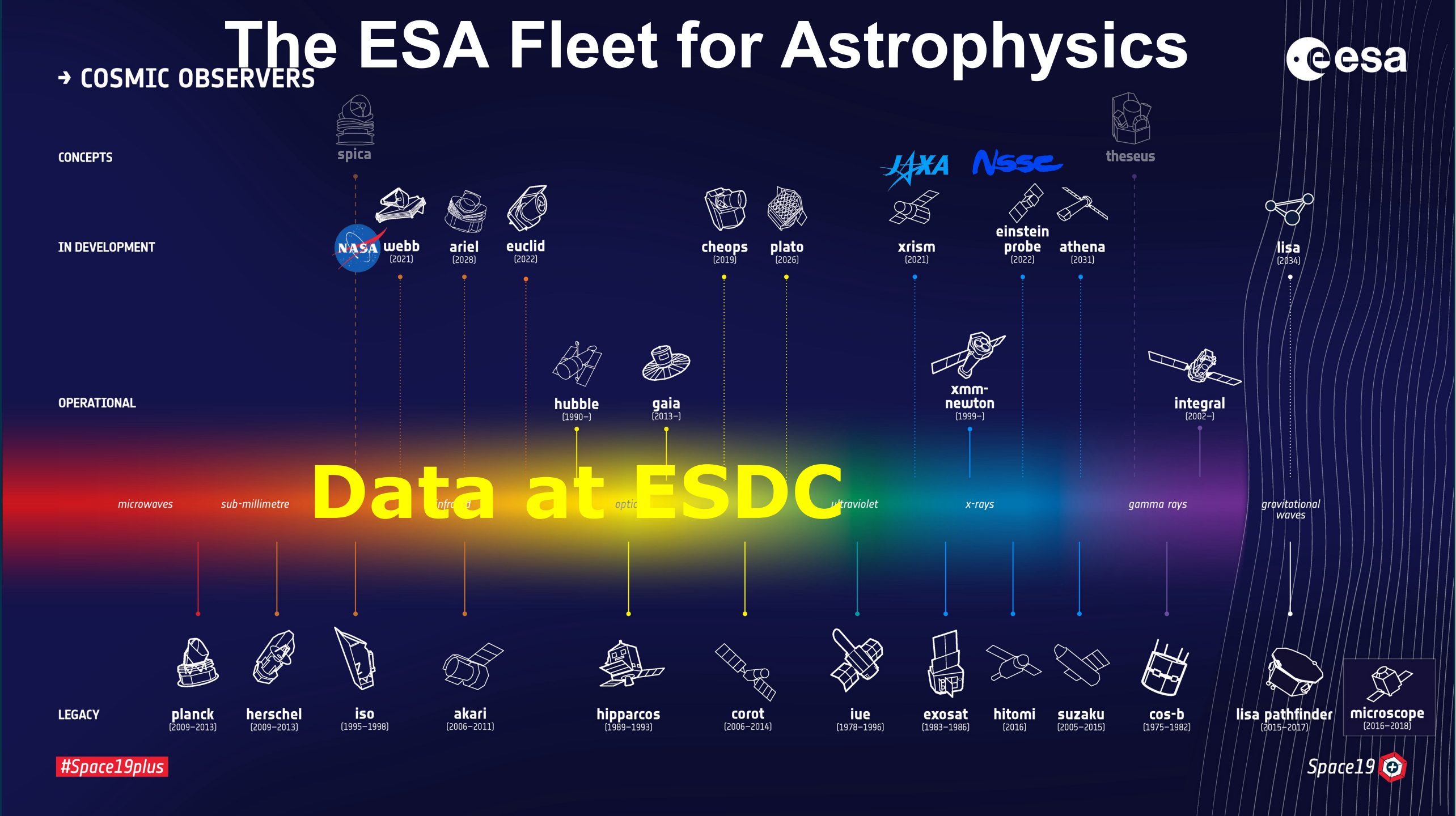
http://archives.esac.esa.int http://sky.esa.int



The ESA Fleet for Astrophysics

esa

→ COSMIC OBSERVERS



#Space19plus

Space19

Database Systems at ESDC

Most archives: **PostgreSQL**: 10.2 to 16.1

- Open Source
- Big community
- Geometrical queries : extensions: (pg_sphere, q3c, healpix, postgis)
- No professional support

Euclid DPS: **Oracle**

ESASky, PSA (some functionalities): **ElasticSearch**

Euclid Prototype: Spark

Import from Files, MySQL…

Databases HW

Chosen to scale up (3):

Largest (GACS):

- Dell PowerEdge R920
- 1.5TB RAM
- 96 cores
- 40 TB SSD

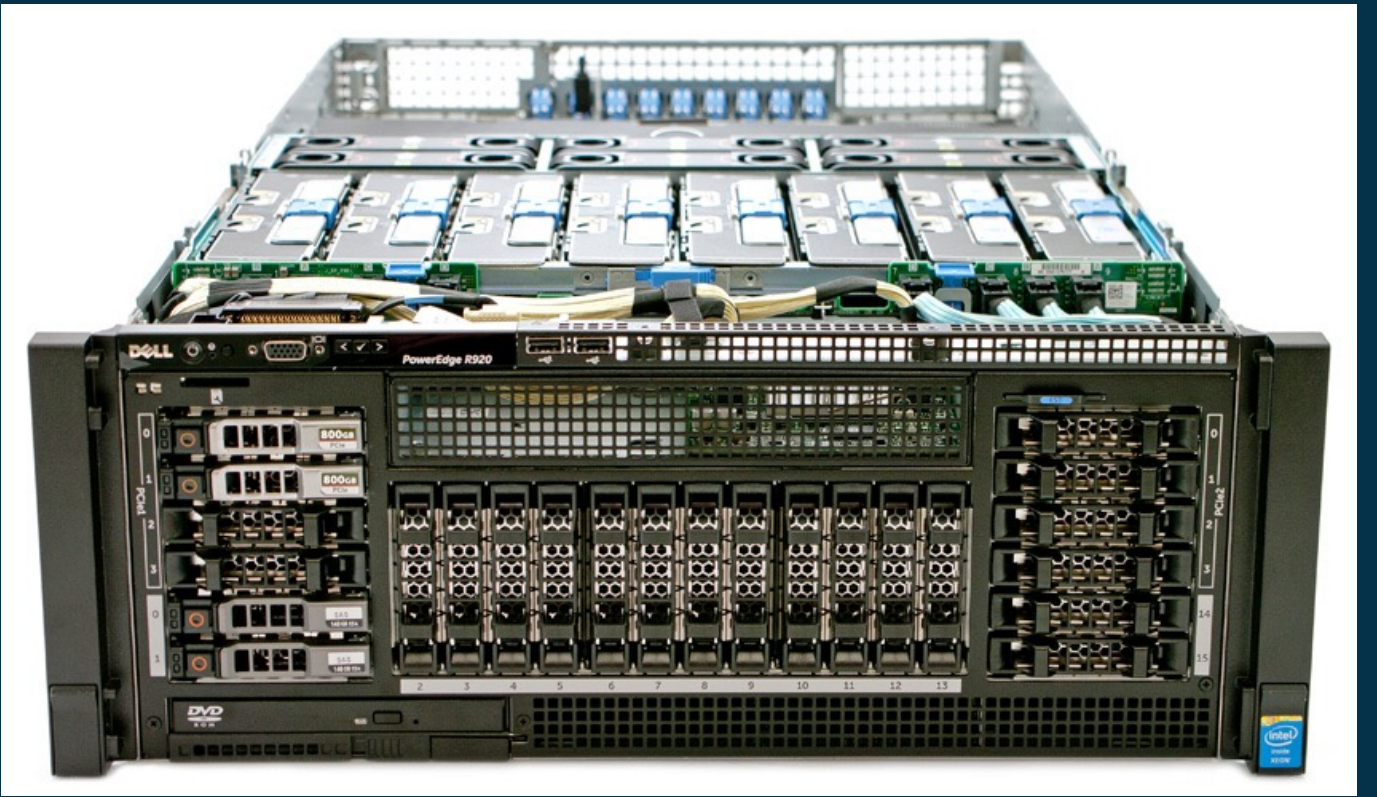
Chosen to scale out (3 GP bare metal clusters):

MASTER(2):

- 256GB RAM
- 3,5TB SSD
- 40 cores

WORKER(4):

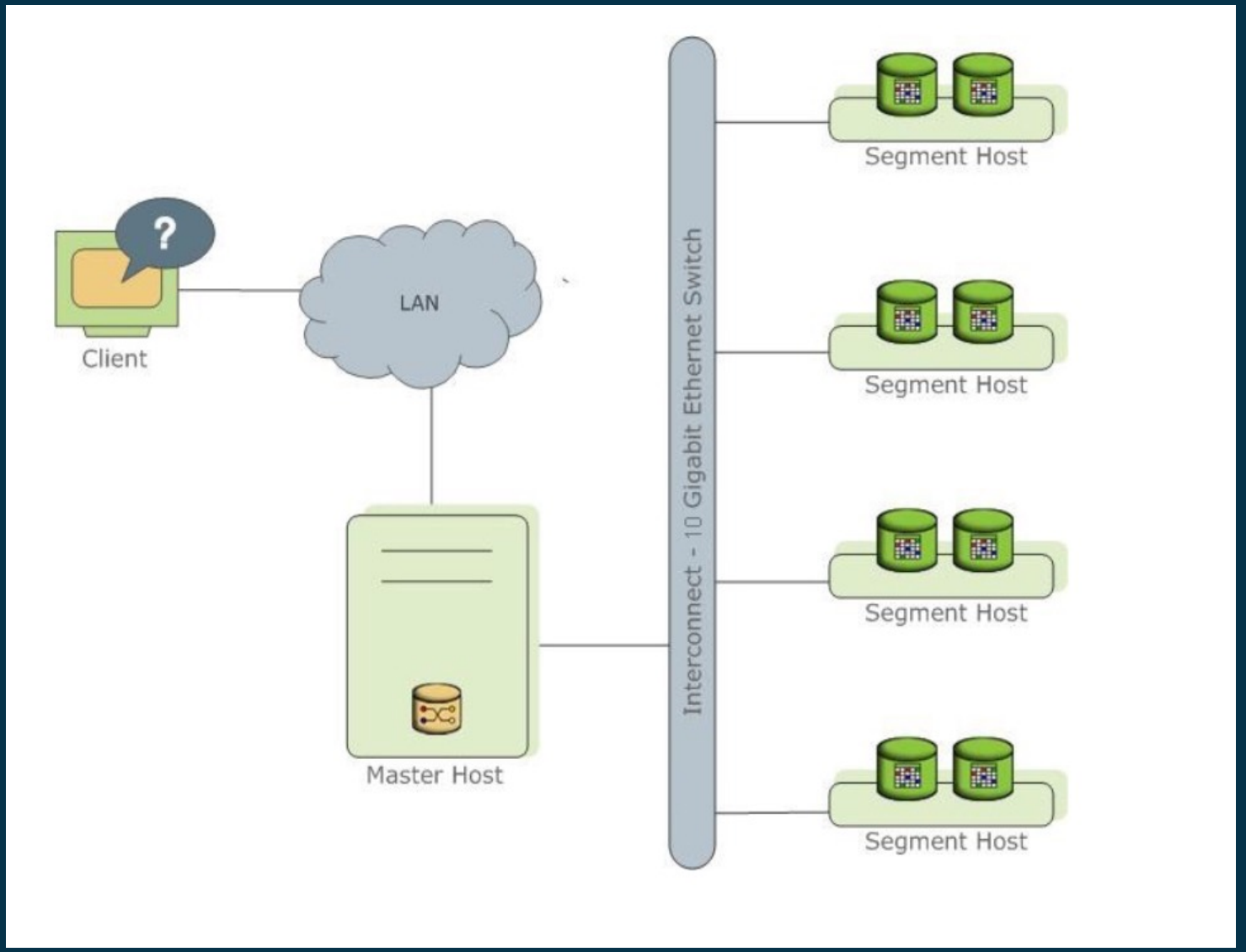
- 512GB RAM
- 40 cores
- 70TB SSD for GACS (35 TBfor Euclid)



Distributed open source flavours tested

- Postrgres-XL(10)
- Citus(7)
- Greenplum (5.9->7.1)

Greenplum Architecture



Greenplum OpenSource

- Used at ESDC since 2018, Greenplum Support tried to negotiate to move to Commercial option twice: 2019 and 2023.
- Greenplum Support companies: Pivotal->Dell->VMWare->Tanzu->Broadcom
- In May 2024, Broadcom decided to discontinue Open Source support and only continue with closed software. Open source last version 7.1
- ESDC last version used 6.27.1 Euclid archive and 7.1 for Gaia archive.

Testing

- New hardware
- Test with real data, no synthetic data
- Test real use cases
- Learn from working with the data
  - Table format options
  - Query plan optimizers
  - Indexes
  - Data distribution
  - …

Reference tables

Table name	# columns	Heap size (GB)	Columnar size with compression (GB)	Compression %
gaiadr3.gaia_source_lite	51	342	180	48%
gaiadr3.gaia_source	152	877	509	42%
user_dr3int6.gaia_source	243	1230	751	39%

```
Distributed by: (source_id)
Options: appendonly=true, orientation=column, compresstype=zlib,
compresslevel=7
```

#rows: 1.811.709.771

GP Testing: count(*)

Query	# columns	Postgres Heap format	Greenplum Heap format	Greenplum Heap format GPORCA	Greenplum Columnar format	Greenplum Columnar format GPORCA
Q1	51	77,9	22,3	21,7	15,9	15,6
Q1	152	63,6	40,9	42,3	11,9	11,9
Q1	243	64,4	55,2	58,5	11,9	12
Q2	51	75,8	45,9	26,9	7,1	11,4
Q2	152	68,3	127,6	66,7	9,5	11,6
Q2	243	82,9	93,9	84,8	9,8	14,1
Q3	51	87,4	39,1	23,2	9,9	15,2
Q3	152	73,3	94,6	53,7	11,4	15,5
Q3	243	192,7	62,7	67	13	15,3

Execution times in seconds

```
Q1=# SELECT COUNT(*) FROM <table>;
     count
---------------
 1.811.709.771
(1 row)
```

```
Q2=# SELECT COUNT(*)
Q2=# FROM <table>
Q2=# WHERE phot_g_mean_mag <= 18.25
Q2=# AND has_mcmc_msc = 'true';
    count
-------------
 348.630.727
(1 row)
```

```
Q3=# SELECT COUNT(*)
Q3=# FROM <table>
Q3=# WHERE parallax IS NOT NULL
Q3=# AND phot_g_mean_mag <= 18.25;
    count
-------------
 353.992.031
(1 row)
```

GP Testing: selected columns

Query	# columns	selected columns	Postgres Heap format	Greenplum Heap format	Greenplum Heap format GPORCA	Greenplum Columnar format	Greenplum Columnar format GPORCA
Q1	51	*	347,5	500,2	507,3	466,2	468,7
Q1	152	*	1.141,20	881,9	883,3	800,7	806,5
Q1	243	*	918,2	1153,1	1.138,50	1247,1	1.244,50
Q2	51	*	44,4	11,5	8,5	85,1	85,5
Q2	152	*	238,9	17,9	22	244,8	244
Q2	243	*	365,8	93,9	76,4	370,1	373
Q2	51	7	12,6	7,8	1,2	25,8	25,8
Q2	152	7	82,4	12,4	2,1	26,9	26
Q2	243	7	13,2	80,8	63,9	25,9	25,9
Q3	51	*	656,6	110,6	105,2	150,4	150,7
Q3	152	*	3.398,90	208,4	194,1	435,4	462,7
Q3	243	*	8.896,20	246,4	248,3	704,8	753,5
Q3	51	7	669,3	71	64,6	69,7	69,9
Q3	152	7	1.441,40	101,6	70,4	67	73,2
Q3	243	7	5.536,40	69	70,4	66,4	70,7

Execution times in seconds

```
Q1=# SELECT * FROM <table>;
...
(1.811.709.771 rows)
```

```
Q2=# SELECT <select>
Q2=# FROM <table>
Q2=# WHERE phot_g_mean_mag <= 18.25
Q2=# AND has_mcmc_msc = 'true';
...
(348.630.727 rows)
```

```
Q3=# SELECT <select>
Q3=# FROM <table>
Q3=# WHERE parallax IS NOT NULL
Q3=# AND phot_g_mean_mag <= 18.25;
...
(353.992.031 rows)
```

GP Testing: Q3C[1] indexing

# columns	angle	Postgres Heap format	Greenplum Heap format	Greenplum Columnar format
51	0,5	308,8	0,2	0,2
51	1	306,6	1,9	1,9
51	2,5	470,7	5,8	4,8
152	0,5	12,6	7,5	0,2
152	1	498,1	8,6	9,8
152	2,5	648,2	14	13,2
243	0,5	12,1	90,3	0,2
243	1	529,7	85,7	9,8
243	2,5	701	86,3	13,2

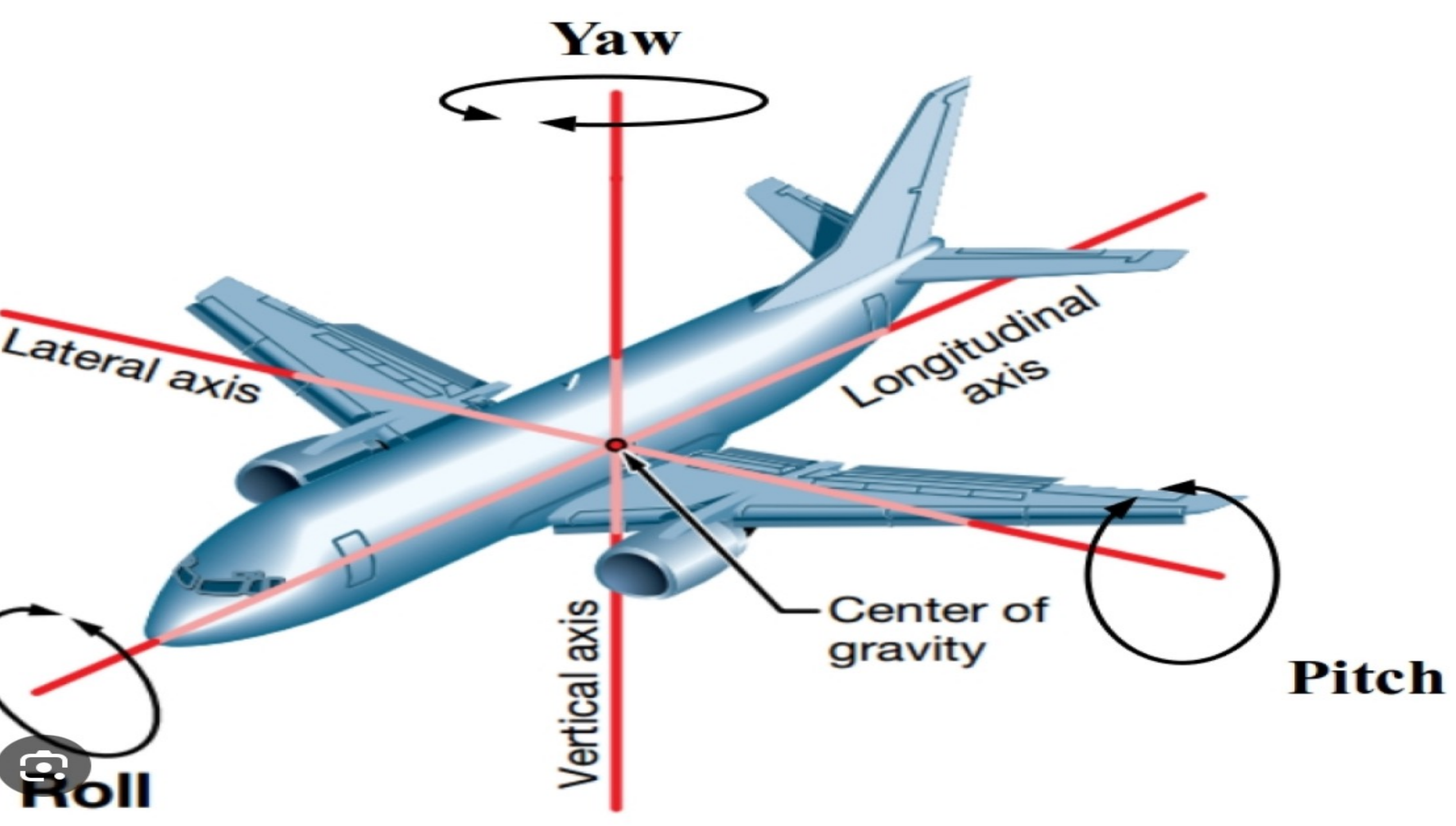
Query optimizer will be chosen by Greenplum
Execution times in seconds

```
Q=# SELECT ...
Q=# FROM <table>
Q=# WHERE '1' = (q3c_join(270.89,-30.03,ra,dec,0.5))
Q=# AND ruwe < 1.4;
...
(358.976 rows)
```

```
Q=# SELECT ...
Q=# FROM <table>
Q=# WHERE '1' = (q3c_join(270.89,-30.03,ra,dec,1))
Q=# AND ruwe < 1.4;
...
(1.514.361 rows)
```

```
Q=# SELECT ...
Q=# FROM <table>
Q=# WHERE '1' = (q3c_join(270.89,-30.03,ra,dec,2.5))
Q=# AND ruwe < 1.4;
...
(10.125.050 rows)
```

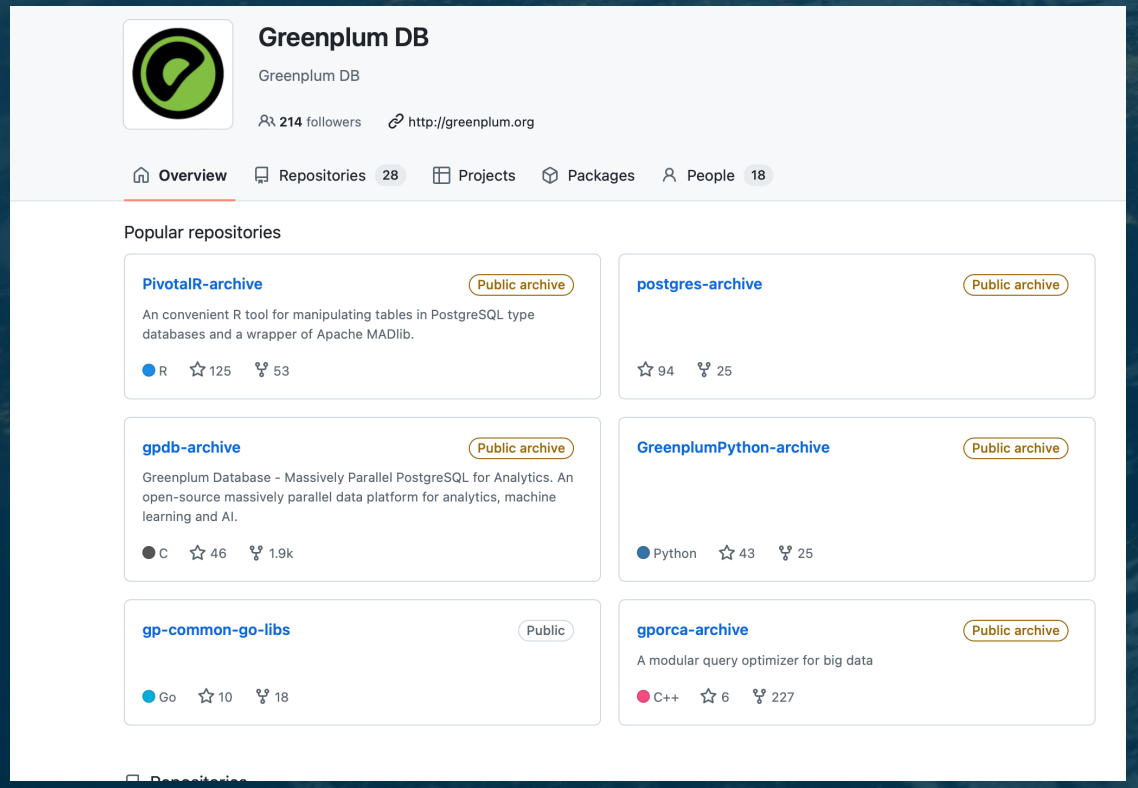
[1] Q3C - PostgreSQL extension for spatial indexing on a sphere.


Yaw
Lateral axis
Longitudinal axis
Center of gravity
Vertical axis
Roll
Pitch

Broadcom cutting ropes

May 29, 2024

Google group: We're writing to let you know that the group you tried to contact (gpdb-users) may not exist,



Alternatives to Greenplum Open Source

1- Greenplum Commercial

2- CloudberryDB

3- CitusData

4- Other MPP databases

MPP databases considered

Engine	License	Postgres based	Q3c/PgSphere compatible
Greenplum	~~Open Source /~~ Closed source	Yes	Yes
CloudberryDB	Open Source	Yes	Yes
Citusdata	Open Source	Yes	Yes
Cockrach Labs	~~Open Source / BSL /~~ Closed source	Yes	No
YugabyteDB	Open Source	Yes	No

Functionality matrix (I)

Functionality	Greenplum	CloudberryDB	Citusdata
Overall cost of the use including cost of purchase, support, maintenance, update and indirect costs (specific hardware, training for personal etc.)	Commercial license: 200K/year (6 nodes) negotiable/ Open Source up to 7.1 Other costs: same for all	Open Source Other costs: same for all	Open Source Other costs: same for all
Migration cost for the existing software (in manpower)	same for all	same for all	same for all
ESDC use cases can be implemented by the selected system	YES	YES	YES
ESDC requirements are satisfied by the selected system	YES	YES	YES
Availability of the support of the product for the duration of the mission and beyond, including a review of support by current customers	YES, by Broadcom/ Stackoverflow or similar	Community	Community
Postgres version	12	14	16
Dependency on third-party products	PgSphere q3c FDW	PgSphere q3c FDW	PgSphere q3c FDW
DBMS maturity and review by current customers	Many years in production. Many customers, more DWH: JP Morgan, Conversant...	Recent fork from Greenplum (2023)	Many years in production. Many customers, more DWH: Microsoft Azure, Upenn..

Functionality matrix (II)

Functionality	Greenplum	CloudberryDB	Citusdata
Support of SQL standards (SQL3) with minimal exceptions	YES	YES	YES
Support for complex queries in the case of exceptions to SQL standards	YES	YES	YES
Possibility to implement SSO based on ESA CAS SSO	N/A same as postgres	N/A same as postgres	N/A same as postgres
Support of replication	YES (mirror segments)	YES (mirror segments)	YES with Patroni
Support for HA	?	?	YES with Patroni
Support of materialized views	YES	YES	YES
Support of external procedures (routines, written in external - non-SQL language)	YES (with segment-level limitations)	YES (with segment-level limitations)	YES
Availability of Python interface	YES	YES	YES
Availability of online backups	YES	YES	YES
Availability of native monitoring tools	YES	YES	YES

Functionality matrix (III)

Functionality	Greenplum	CloudberryDB	Citusdata
Table partitioning	YES, distributed.	YES, distributed.	YES, distributed.
Multiple users and roles	YES	YES	YES
IEEE floating point support (is included in recent SQL standards, but there were earlier problems so keep as separate requirements	YES	YES	YES
ACID compliant	YES	YES	YES
Unique primary key	YES, must be part of distribution key	YES, must be part of distribution key	YES, must be part of distribution key
Foreign key support	NO	NO	YES
Table normalisation	De-normalise tables is advised	De-normalise tables is advised	Support for co-located tables (shard-based on distribution keys)
Support Company	Broadcom	HashData	Citusdata (part of Microsoft)

First tests – CloudberryDB and Citusdata (I)

First tests using local deployment (Mac laptop using docker)

- CloudberryDB supported natively on Mac (aarch64)
- Citusdata supported via emulators (rosetta)
- Clusters with 3 nodes (1 master + 2 worker nodes)

Query	Citusdata vs CloudberryDB
Q1	-5%
Q2	24%
Q3	15%
Q4	-77%
Q5	-58%
Q6	-82%
Q7	-47%

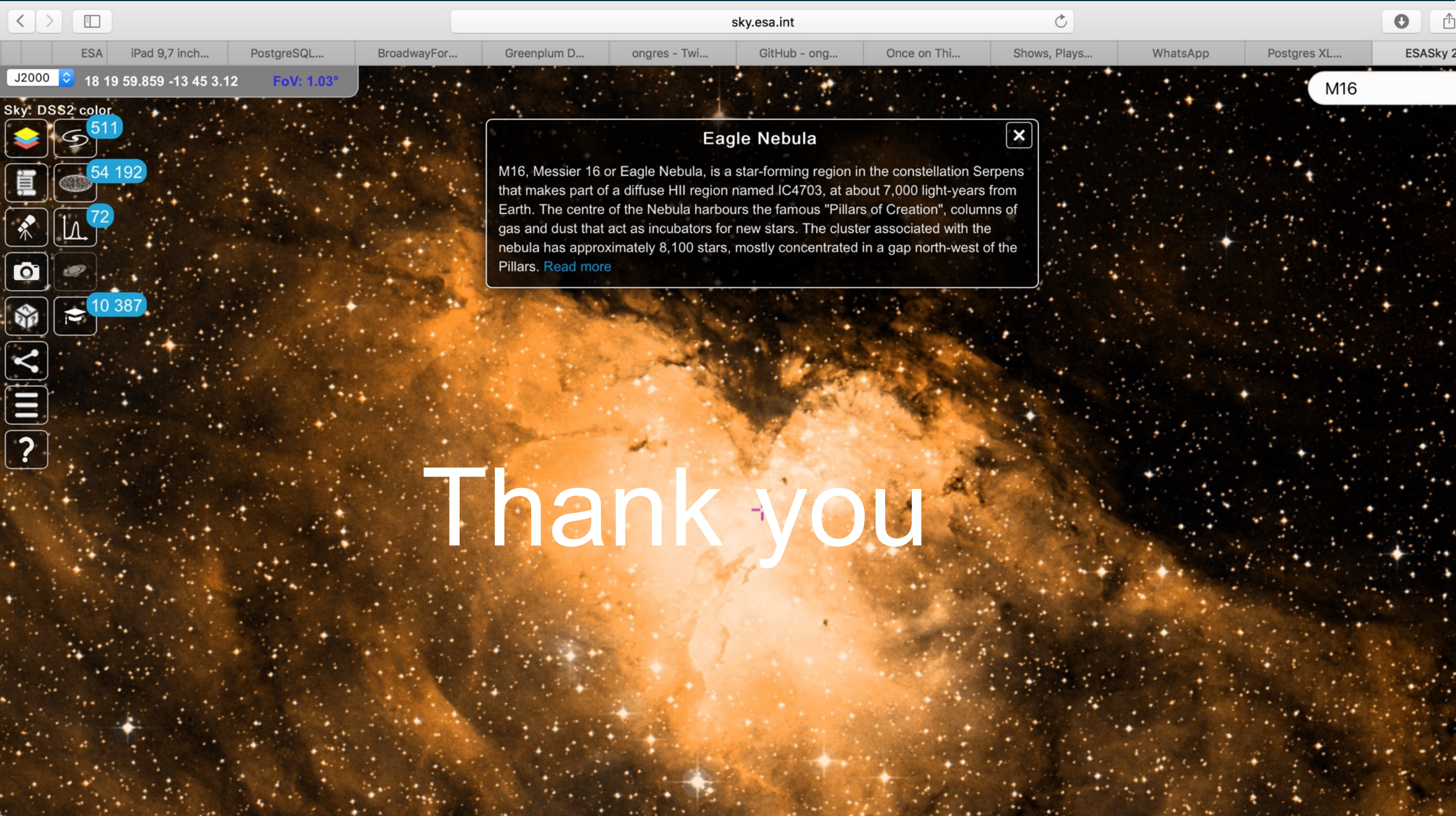
Conclusions

- Both solutions support our test queries (usage of extensions)
- Results are promising
- Running on bare metal will show the real performance (not docker, not virtual machines and not emulating)

Summary

- GP OSS no longer an option because it is now closed sofware and we cannot afford it
- Alternatives
- Risks of using OSS software. Trade-offs

We continue flying……
not landing yet


sky.esa.int
ESA
iPad 9,7 inch...
PostgreSQL...
BroadwayFor...
Greenplum D...
ongres - Twi...
GitHub - ong...
Once on Thi...
Shows, Plays...
WhatsApp
Postgres XL...
ESASky 2
J2000
18 19 59.859 -13 45 3.12
FoV: 1.03°
Sky: DSS2 color
511
54 192
72
10 387
M16
Eagle Nebula
M16, Messier 16 or Eagle Nebula, is a star-forming region in the constellation Serpens that makes part of a diffuse HII region named IC4703, at about 7,000 light-years from Earth. The centre of the Nebula harbours the famous "Pillars of Creation", columns of gas and dust that act as incubators for new stars. The cluster associated with the nebula has approximately 8,100 stars, mostly concentrated in a gap north-west of the Pillars. Read more
Thank you