

The Euclid NL- Science Data Center

Edwin A. Valentijn

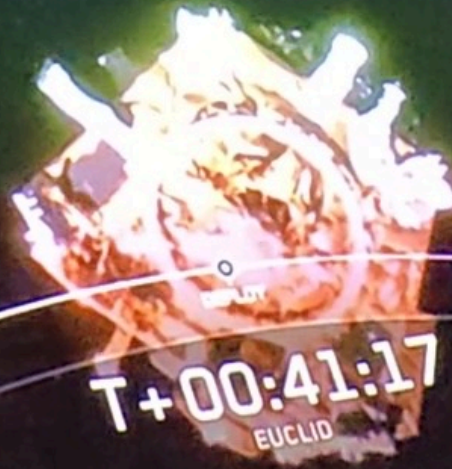
Rijks Universiteit Groningen - Kapteyn Astronomisch Instituut

NL – Euclid SGS Project manager

Big Data processing in Astronomy and Earth Observations

23 jan 2024 – SURF/NSO- Utrecht

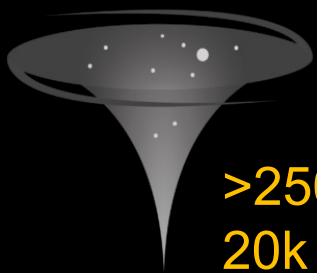




SPEED
27286
KM/H

ALTITUDE
5711
KM

STAGE 2 TELEMETRY



>250 days Hubble Legacy Field
20k x 20k
265.000 galaxies



30'

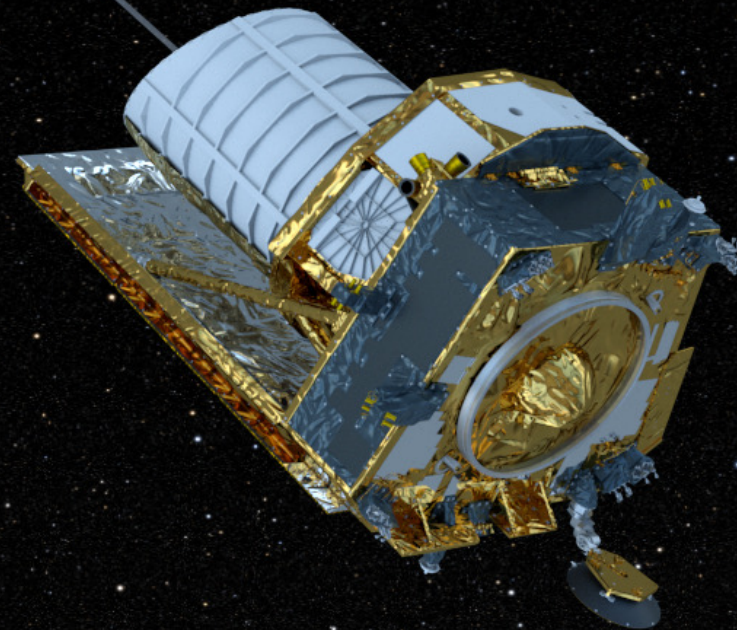
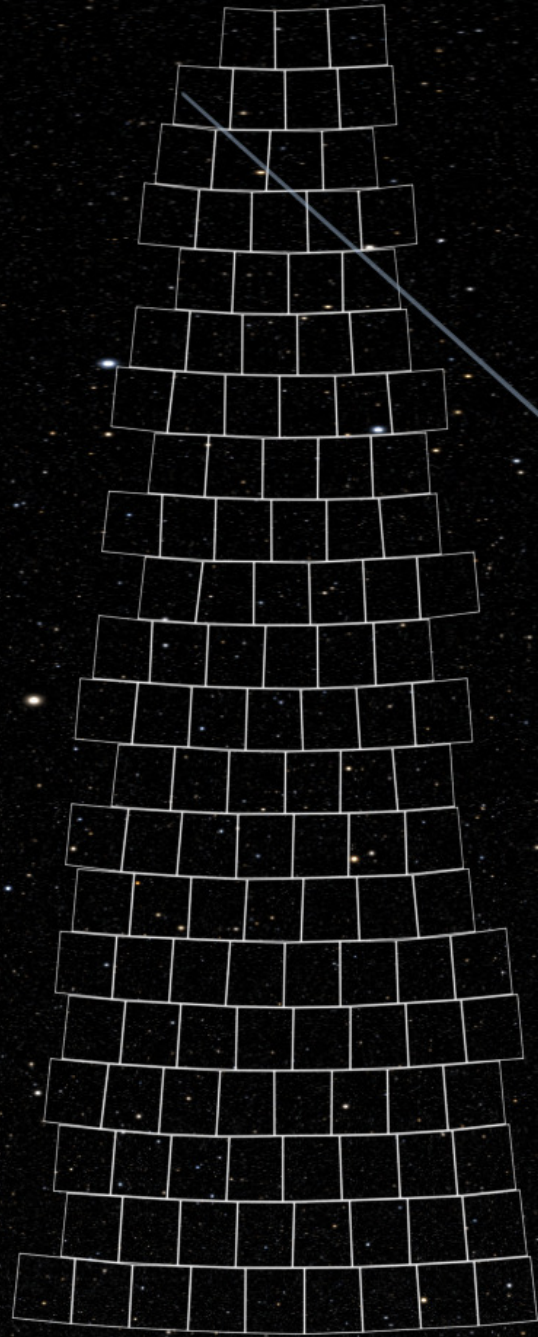


4200 sec nominal 45' x 45' 0.57 Sq deg
VIS: 4x 565 sec + 2x100s 0.1"/pix 24k x 24 k 609 Mpix
NISP: 565 sec spectra + Y, J, H 112 sec
300.0000 galaxies
50.000 distances



30'





- 15.000 Sq deg Wide
 - 50 Sq deg - 3 deep fields
 - calibrations
- 6 year



Cosmology :

Evolution $z < 2$ last 10 Billion year - 3 probes:

- Weak lensing tomography $\sim 10-12$ slices
- BAOs
- Large scale structure

very high precision :

milli-arcsec milli-mag

PSF Phase diversity cal.

Legacy Science:

- Local Universe -LSB
- Strong lensing
- Theory
- Galaxy, AGN evolution
- Milky Way
- Primeval Universe
- Solar system

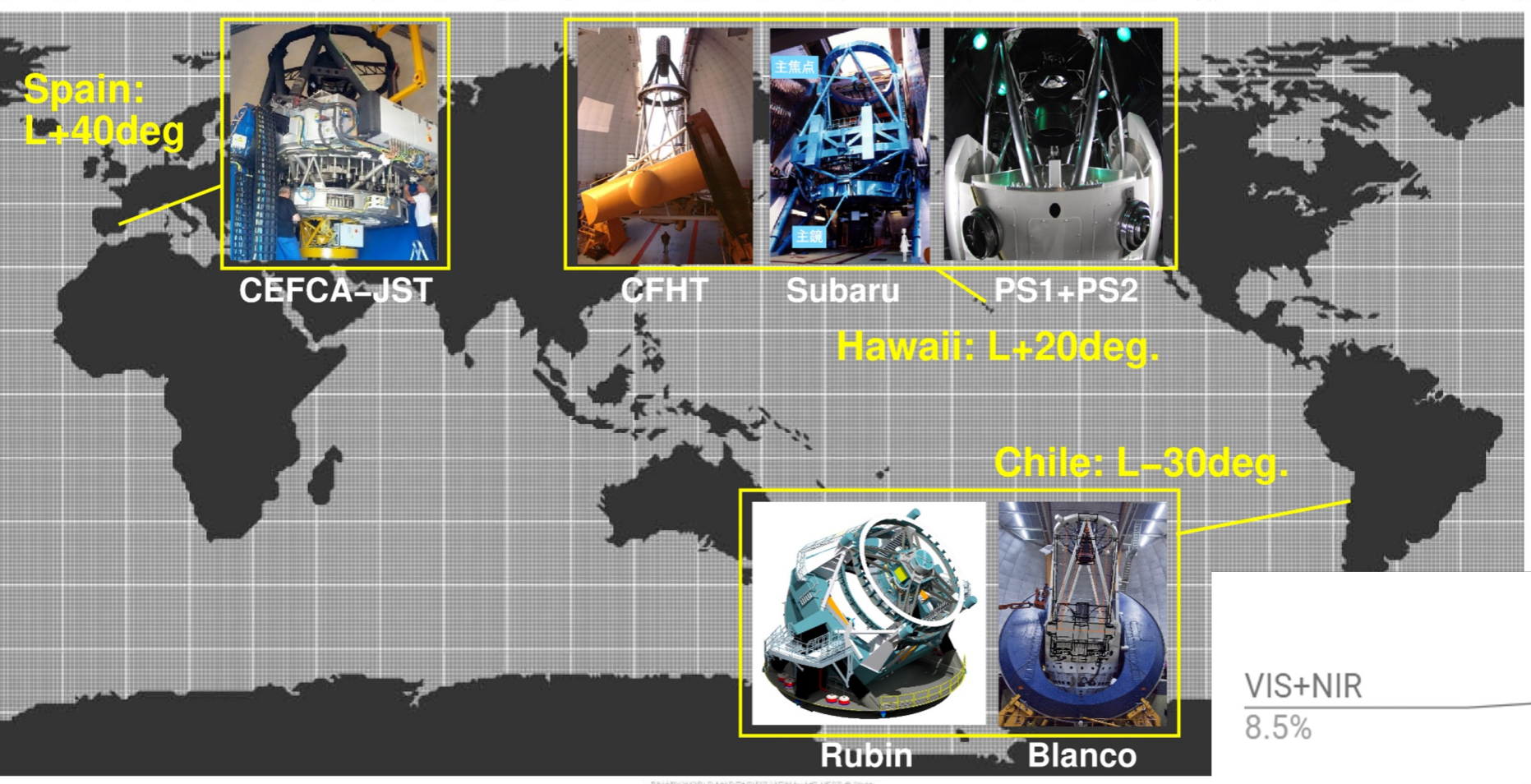


“ Euclid is Big Data from Space ”

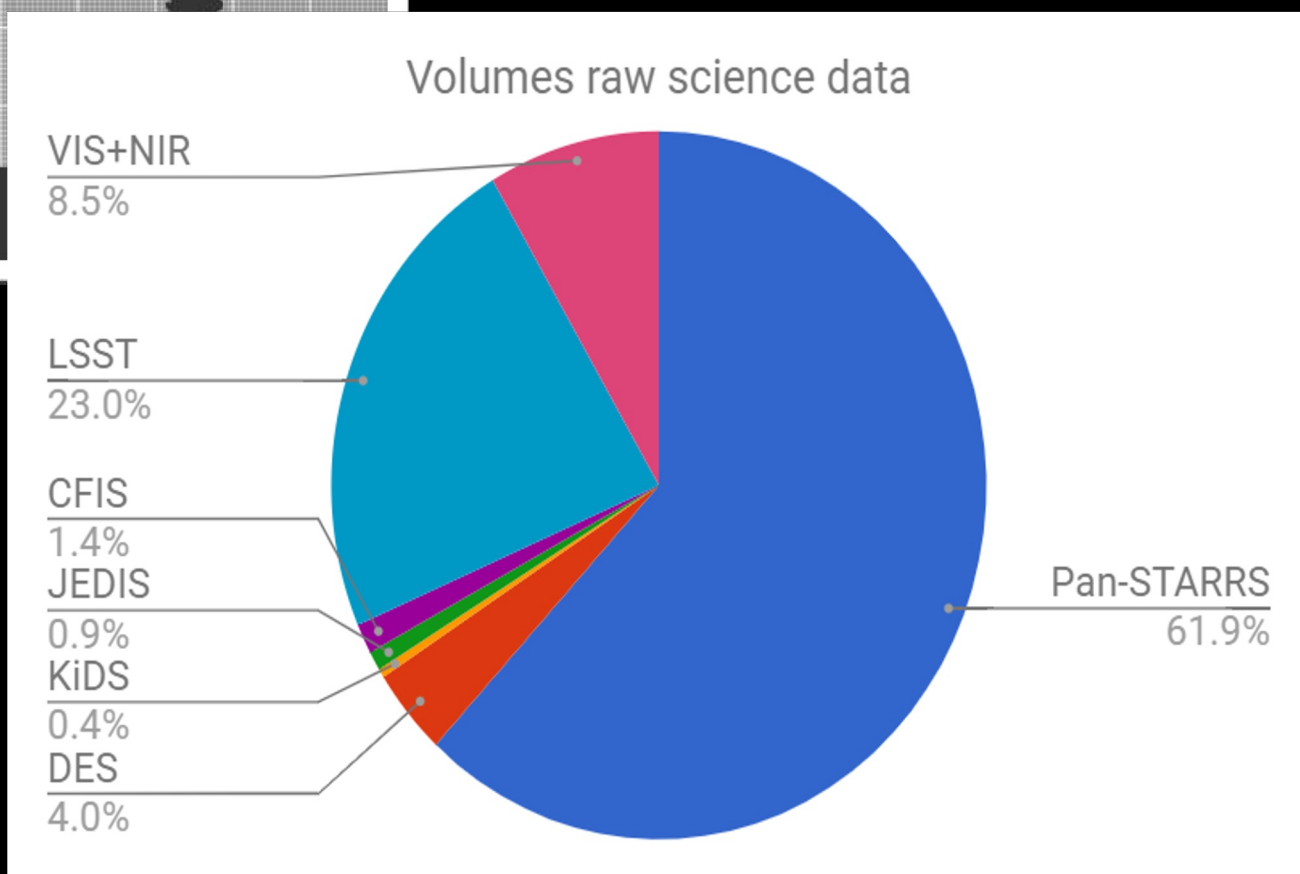
2 billion galaxies – distortions and distances

200.000 exposures (1.8 Million images) – 2.5 billion pixels /4.400 sec > 3.700 Tbyte

Costs of data handling ~ costs instruments



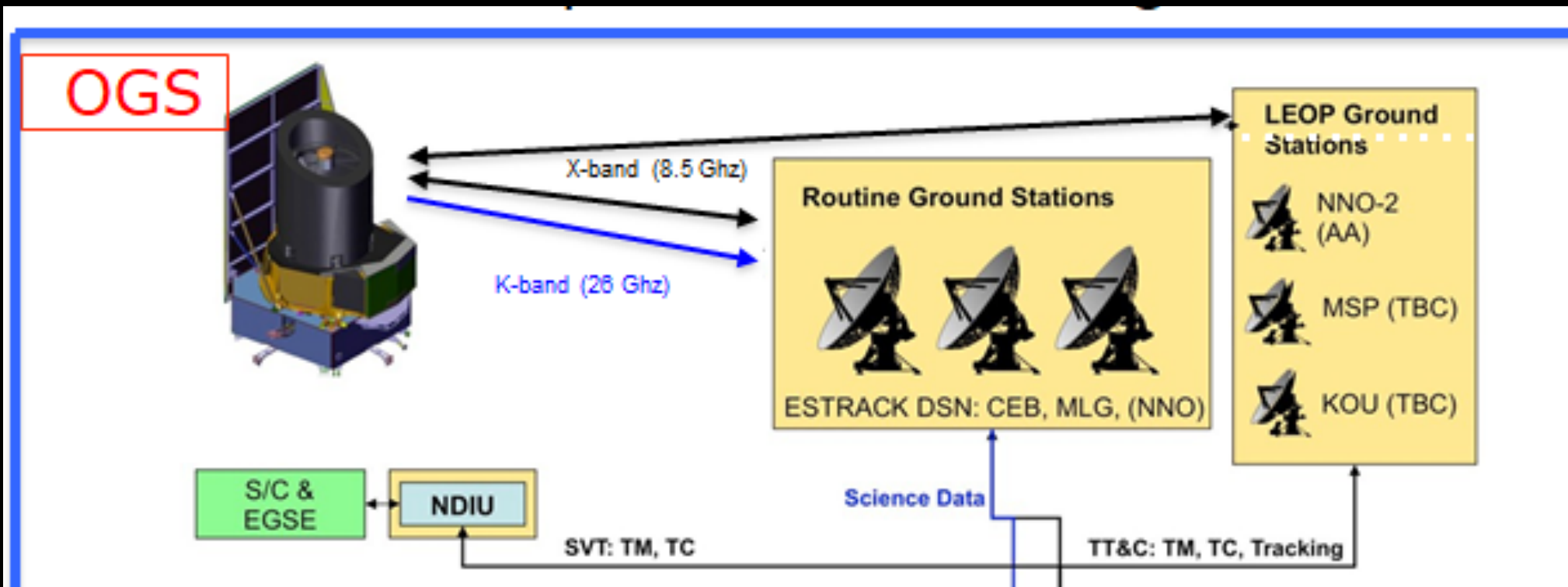
Euclid images > 3.7 Petabyte
 Photometric redshifts- ground based data
 > 40 Petabyte
 NL- OU EXT





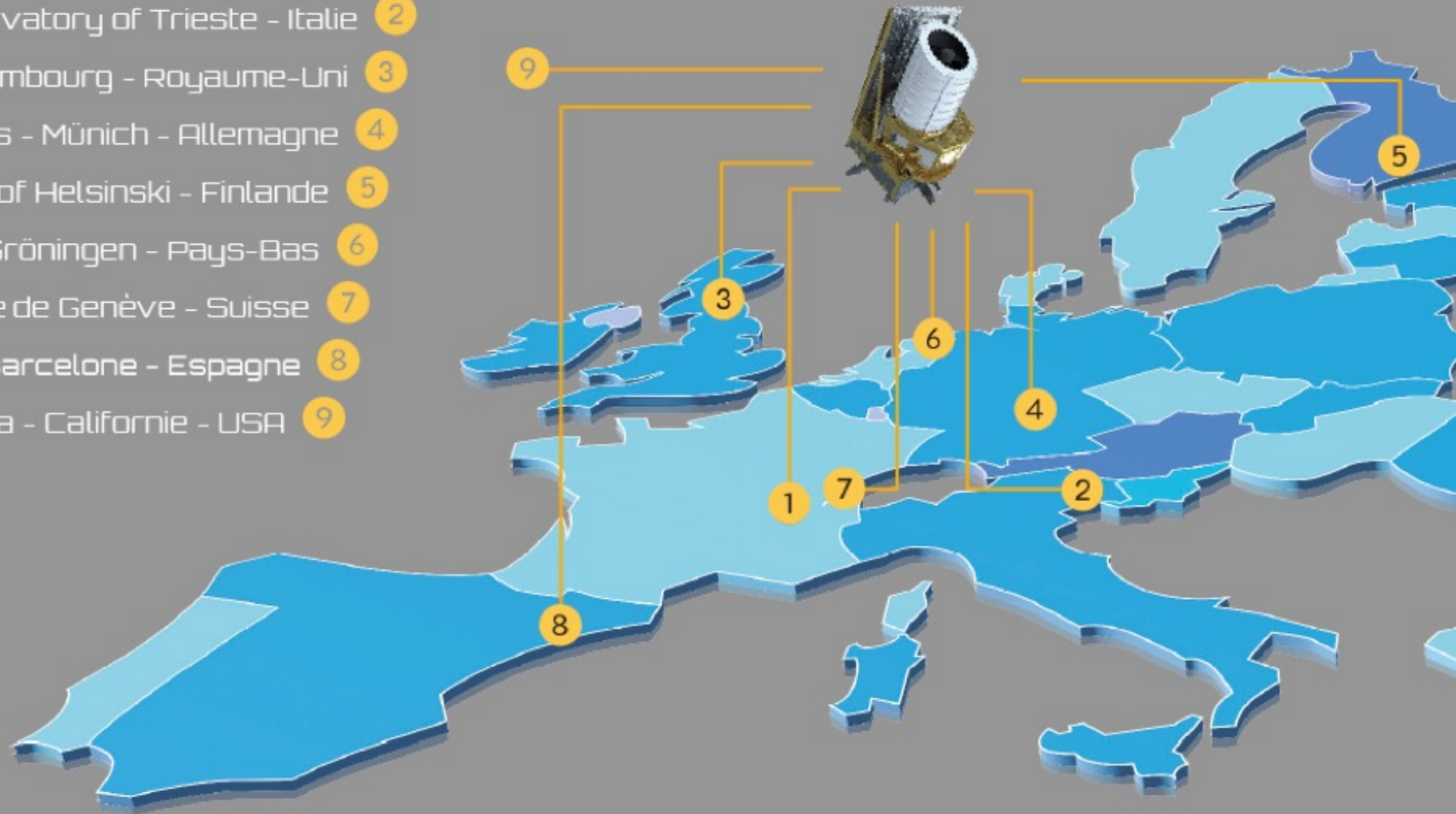
EUCLID EARLY COMMISSIONING TEST IMAGES





K-Band (26 GHz) - L2 - high gain antenna's groundstations

- Centre de Calcul de l'IN2P3 - Lyon - France ①
- Astronomical Observatory of Trieste - Italie ②
- Institute for Astronomy - Edimbourg - Royaume-Uni ③
- Max-Planck-Institute for Extraterrestrial Physics - München - Allemagne ④
- University of Helsinki - Finlande ⑤
- Ronald Smits Centrum voor Informatie Technologie - Groningen - Pays-Bas ⑥
- Département d'astronomie de l'université de Genève - Suisse ⑦
- Port d'Informació Científica - Barcelone - Espagne ⑧
- IPAC, Caltech, Pasadena - Californie - USA ⑨

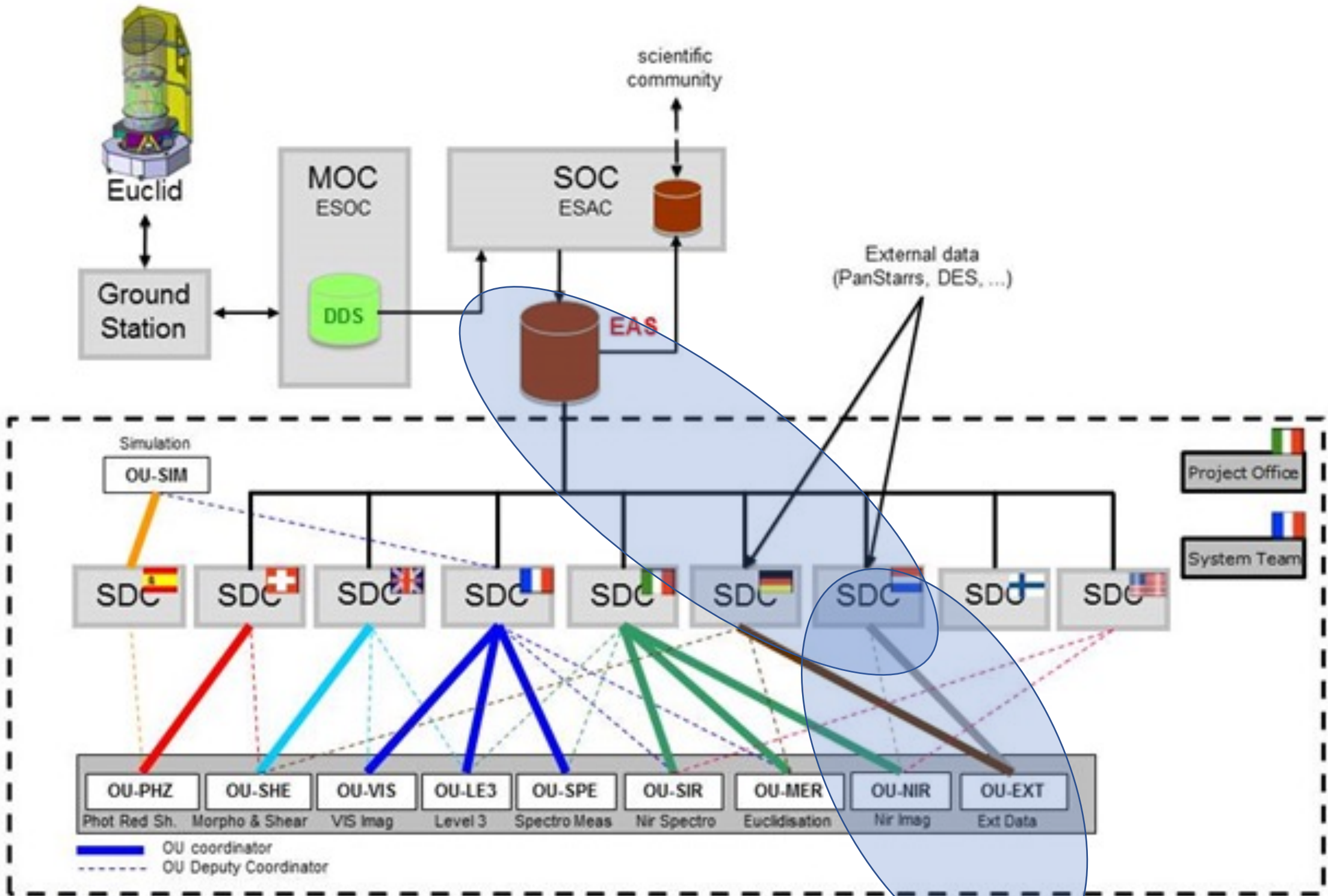


Network of 8 data centers in Europe – 1 in US

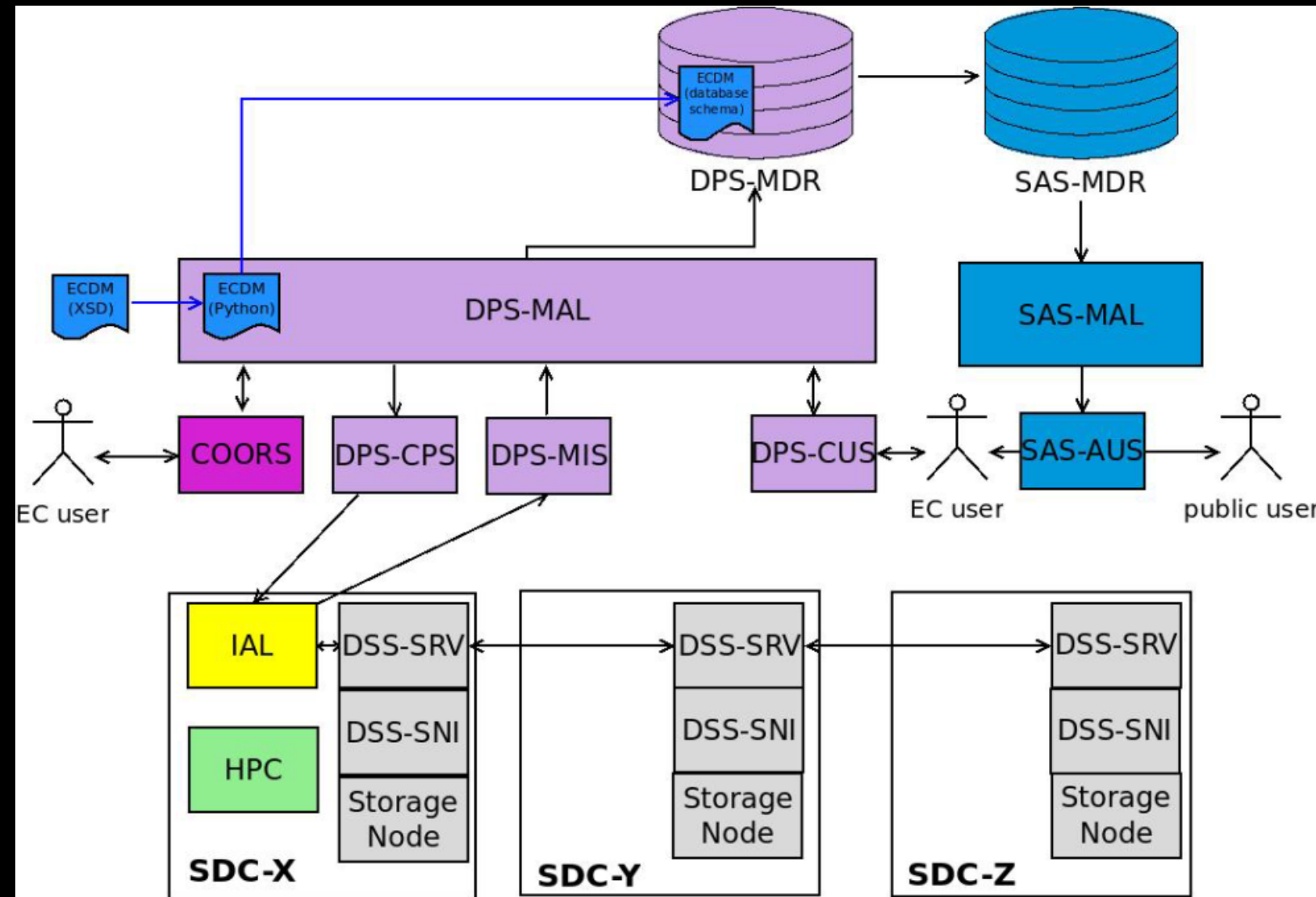
1 integrated system FAIR AstroWISE technology (2003, OmegaCAM, Lofar, +)

Formal NL contribution Euclid = datahandling:

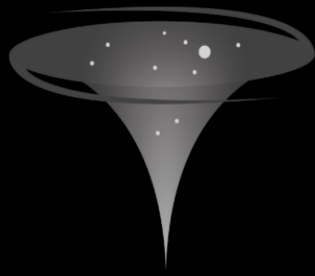
- Euclid Archive System = central database – distributed processing + storage
- Ground based data
- NIR pipeline



Euclid Archive System



- Central DB **F A I R (Reproducible)**
- Connect 9 datacenters
- AstroWISE 2003 – KiDS, LoWISE, MUSEWISE
- No dedicated hardware
- Many access layers
- **Data Processing System - DPS**
 - at the heart of Euclid data processing
 - Stores Metadata all files, including lineage (calibrations) and processing
 - Production system hosted by ESAC
 - Development system by SDC-NL
- **Distributed Storage System - DSS**
 - Stores the data files
 - Distributed across all national SDCs
- **Science Archive System – SAS** access to
 - external users -official data releases
 - internal users - internal data releases



DB view
EAS = DPS + DSS
db mirror at NL-SDC

DpdVisRawFrame

Order by: descending ascending

2nd Order by: descending ascending

Maximum number of rows:

Show only data within project: yes no

Show expanded attributes: yes no

When expanding attributes show empty joins: yes no

Expand desc list attributes: yes no

Show checkboxes: yes no

Export result as:

Submit

DpdVisRawFrame

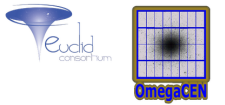
Query results for table DpdVisRawFrame

Shown: 100 rows out of 5200 entries, from only project EUCLID

[Send via SAMP](#) [Help SAMP](#)

ROWNUM	object_id	Lineage	+PRIVILEGES	project_id	Data.AxisNumber	Data.CompressionAlgorithm	Data.DataLength	Data.DataSize	Data.ImgNumber	Data.InstrumentMode	Data.ObservationMode	Data.OnBoardS
1	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
2	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
3	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
4	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
5	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	CALIBRATION	3.1.2
6	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
7	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
8	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
9	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
10	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	CALIBRATION	3.1.2
11	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	CALIBRATION	3.1.2
12	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
13	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
14	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
15	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
16	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
17	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
18	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
19	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
20	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
21	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
22	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
23	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
24	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
25	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
26	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
27	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	CALIBRATION	3.1.2
28	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
29	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2

empowered by



EAS:
~80 NL-EC
2.800 EC


later:
Public ESA - SAS

Calibration Files - timestamping

Calibration Timestamps x EUCLID Parametric Plots x Euclid Archive System x Euclid Archive System x Euclid Archive System x Launch Meeting - Zoom x Calibration Timestamps x

eas-dps-cus-ops.esac.esa.int/CalTS

Apps MijnOverheid Music Euclid Sailing Computers Food Media Proeftuin Travel Retail ECM Curacao Surgery Summary Report fo... Information Univers... ear All



Calibration Timestamps DEV

Contact
Willem-Jan Vriend

DB user
OWILLIAM

Help
getting started

Vis

DpdVisLargeFlatFrame
DpdVisPSFModel
DpdVisXTalk
DpdVisMasterBiasFrame
DpdVisMasterDarkFrame
DpdVisMasterFlatFrame

DpdVisFileContainer

config_ini
gain_model
ron_model
saturation_model
crosstalk_model
ghost_model
nonlinearity_model
starmask_model
dark_flagmap
flat_flagmap
calts_test

Nir

DpdNirLargeFlatFrame
DpdNirGeometricDistortions
DpdNirAbsPhotoCalibration
DpdNirMasterFlatFrame

Project: TEST Instrument: EUCLID Privileges: 1 private

Chip: <none> Filter: <none>

year: 2022 quarter: <none> month: 4 apr week: <none> custom period

Hide manual flagged data Hide unknown status Graph
 Hide quality flagged data Hide fully eclipsed data Table

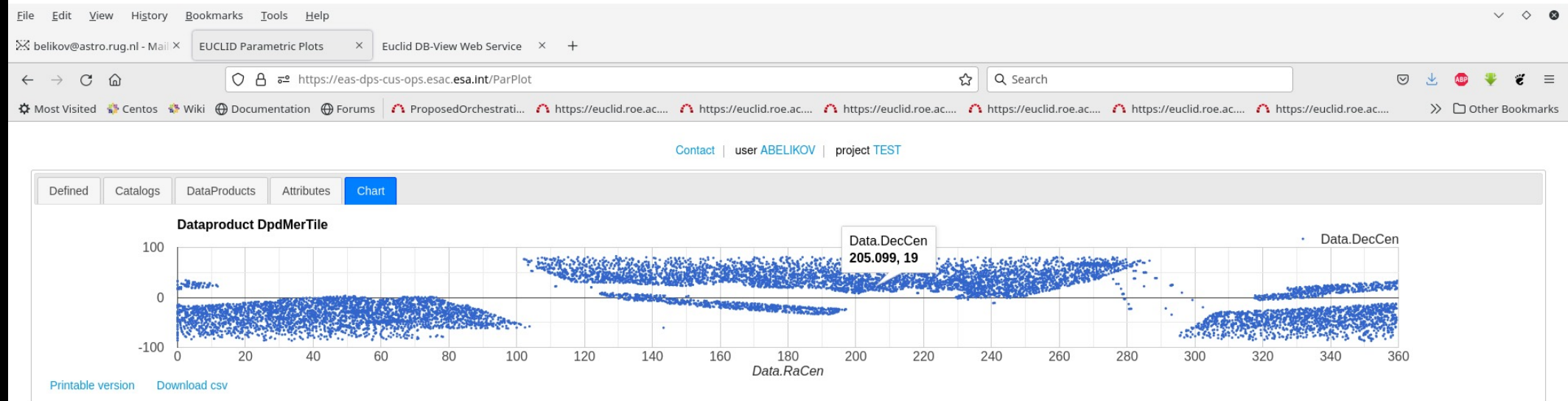
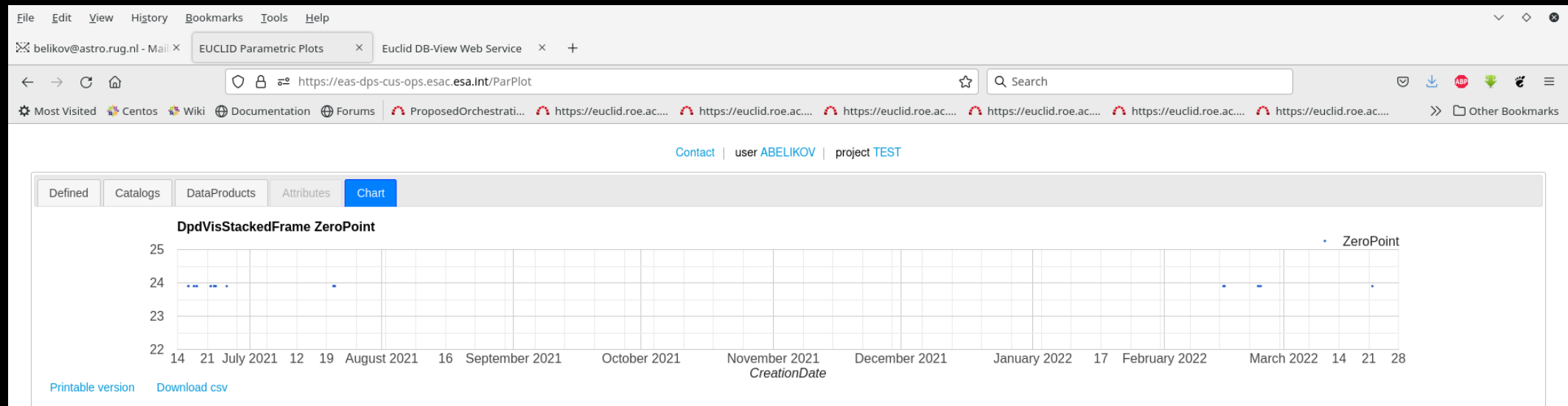
Total calts_test 25 (see in dbview)

Links	Timestamps; from 01 Apr 2022 to 30 Apr 2022	Creation
01 Apr 2022-29 Apr 2022 + D		25 Apr 2021
01 Apr 2022-02 Apr 2022 + D		01 Apr 2022
02 Apr 2022-03 Apr 2022 + D		02 Apr 2022
03 Apr 2022-04 Apr 2022 + D		03 Apr 2022
05 Apr 2022-06 Apr 2022 + D		05 Apr 2022
07 Apr 2022-08 Apr 2022 + D		07 Apr 2022
08 Apr 2022-09 Apr 2022 + D		08 Apr 2022
10 Apr 2022-11 Apr 2022 + D		10 Apr 2022
12 Apr 2022-13 Apr 2022 + D		12 Apr 2022
13 Apr 2022-14 Apr 2022 + D		13 Apr 2022
14 Apr 2022-15 Apr 2022 + D		13 Apr 2022
18 Apr 2022-19 Apr 2022 + D		14 Apr 2022
20 Apr 2022-21 Apr 2022 + D		18 Apr 2022
21 Apr 2022-22 Apr 2022 + D		20 Apr 2022
22 Apr 2022-23 Apr 2022 + D		21 Apr 2022
22 Apr 2022-23 Apr 2022 + D		22 Apr 2022
23 Apr 2022-24 Apr 2022 + D		22 Apr 2022
26 Apr 2022-27 Apr 2022 + D		23 Apr 2022
27 Apr 2022-28 Apr 2022 + D		26 Apr 2022
28 Apr 2022-29 Apr 2022 + D		27 Apr 2022
		28 Apr 2022

Legend:
valid data (green)
status unknown (grey)
eclipsed data (black)
quality flag set (purple)
valid flag set (red)

page generated 2022-04-29 11:27:35
empowered by

ParPlot e.g. trending



Jupyter Notebooks - Gitlab

Home Page - Select or create a x Untitled - Jupyter Notebook x +

https://apceulidccjupyter.in2p3.fr/user/tnutma/notebooks/Untitled.ipynb?kernel_name=eas-dps-env_0.8 120% ☆ ≡

jupyterhub Untitled Last Checkpoint: 4 minutes ago (unsaved changes) Logout Control Panel

File Edit View Insert Cell Kernel Help Trusted Python 3 (EAS-DPS 0.8) ○

Run Code

2021-07-20 17:29:48.271	2021-07-20 00:04:23.703	2021-07-20 02:09:53.747	23.9	0.01141	mag	VIS
2021-07-20 18:36:36.578	2021-07-20 03:40:47.094	2021-07-20 04:10:23.209	23.9	0.01317	mag	VIS

```
In [10]: # Plot the zero point vs creation time.  
df.plot(y='zeropoint', yerr='zeropointerror', alpha=0.4)
```

Out[10]: <AxesSubplot:xlabel='creation_date'>

In []: |

SDC-NL: Facilities at CIT

3 PByte storage

80 TByte SSD storage and two ORACLE database servers

8 servers Data Handling System - Euclid Archive System

640 cpu cores Habrok High Performance Cluster
+ **18,000 shared cores**



Contacts NL-Science Data Center

SDC-NL@list.rug.nl



Valentijn
lead



Saifollahi
Notebooks



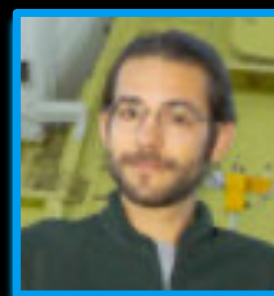
Verdoes Kleijn
External data



de Jong
External Data



Eduardo Balbinot
External Data



Pablo Corcho Caballero
External Data



Williams
GIT lead



Vriend
User Interfaces



Boxhoorn
Databasing



Dröge
HPC



Tsyganov
Databasing



Belikov
Euclid Archive



Bouwens- NISP



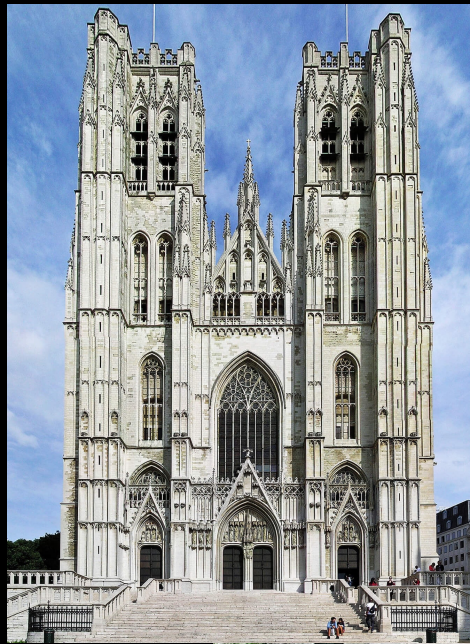
Venemans- NISP

Organizing distributed communities – Open Science

Cathedral

Database

Data model



Access layers

DPS MAL

DPS x

IAL

DSS x

Containers -data Meta data



Bazar

Hardware

processors

storage

Jupiter Notebooks

