Netherlands Institute for Radio Astronomy

Impacts of Operations Models on Computing:
Observatory and User Perspectives

Michael Wise
ASTRON

Second ngVLA Technical Workshop
Socorro, December 8, 2015

ASTRON is part of the Netherlands Organisation for Scientific Research (NWO)

ASTRON

Netherlands Institute for Radio Astronomy

Impacts of Operations Models on Computing: *Observatory and User Perspectives*

Talk Outline

Evolving Definition of Operational Computing

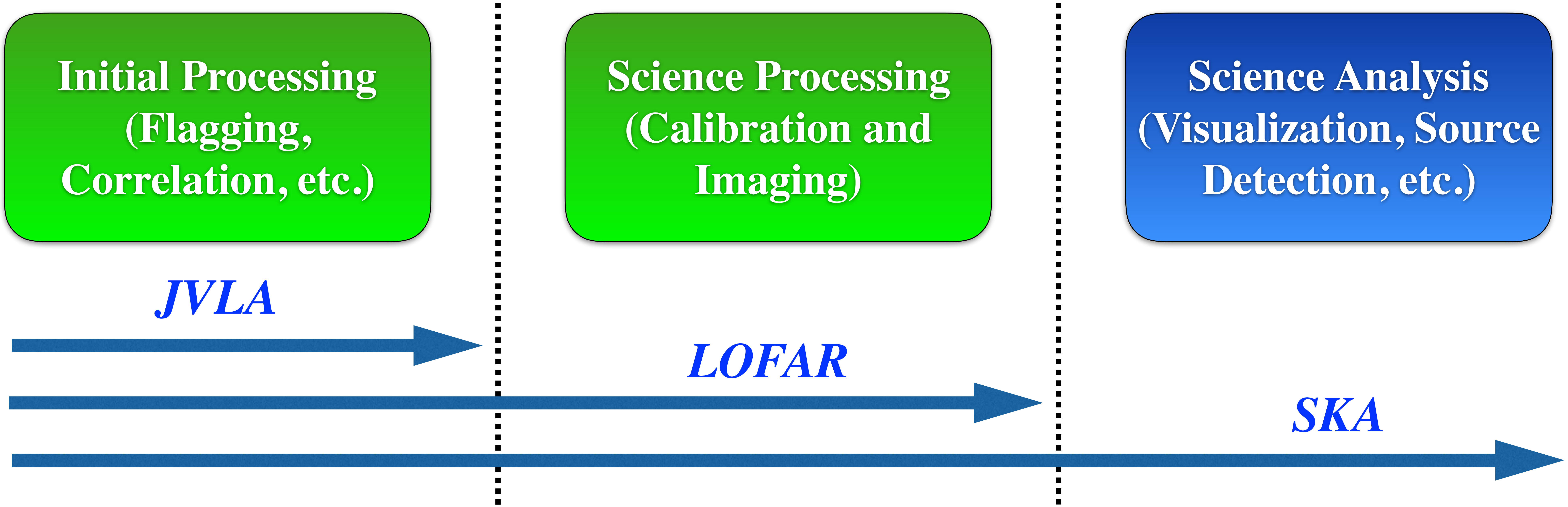
Operational Support for KSPs and Individual Users

Science Extraction from Laptop to Data Centres

LOFAR

ASTRON is part of the Netherlands Organisation for Scientific Research (NWO)

Full Cost of Science Extraction



→ ***What is covered by operational computing?***

→ ***What is left to the community to support?***

Increasing data scale pushes us to more operational integration

LOFAR System Data Flow

Station signals collected in the station cabinets

Signals sent to COBALT for correlation

Data sent to CEP2 for initial post-processing

CEP2

CEP3

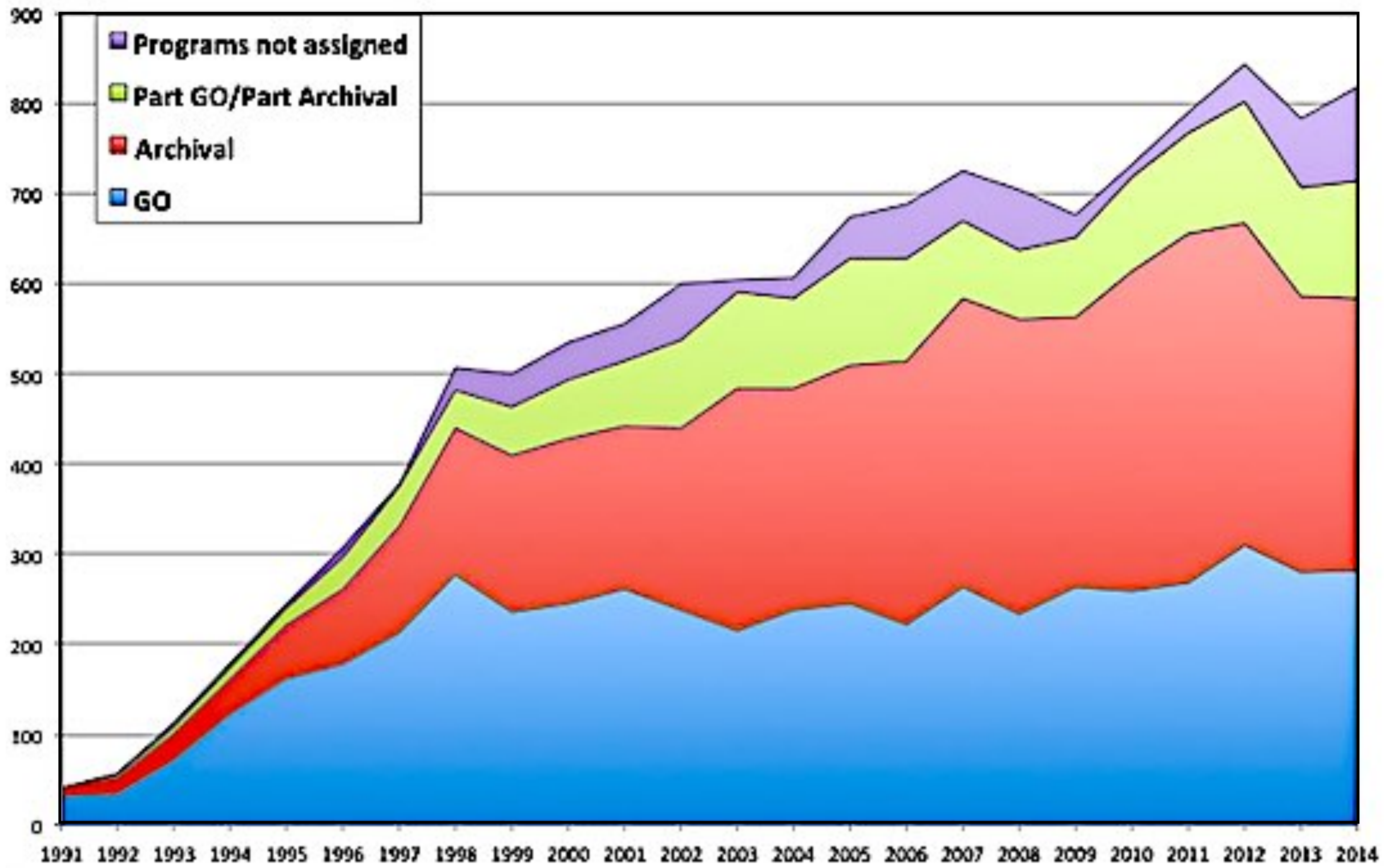
CEP3 cluster available for additional user processing

Products sent to the long-term archive

Entire process is overseen by Observatory
Data compression of ~200 before archiving
Does not yield “science-ready” products

Impact on Science Archives

Science archives are a multiplier for total science output



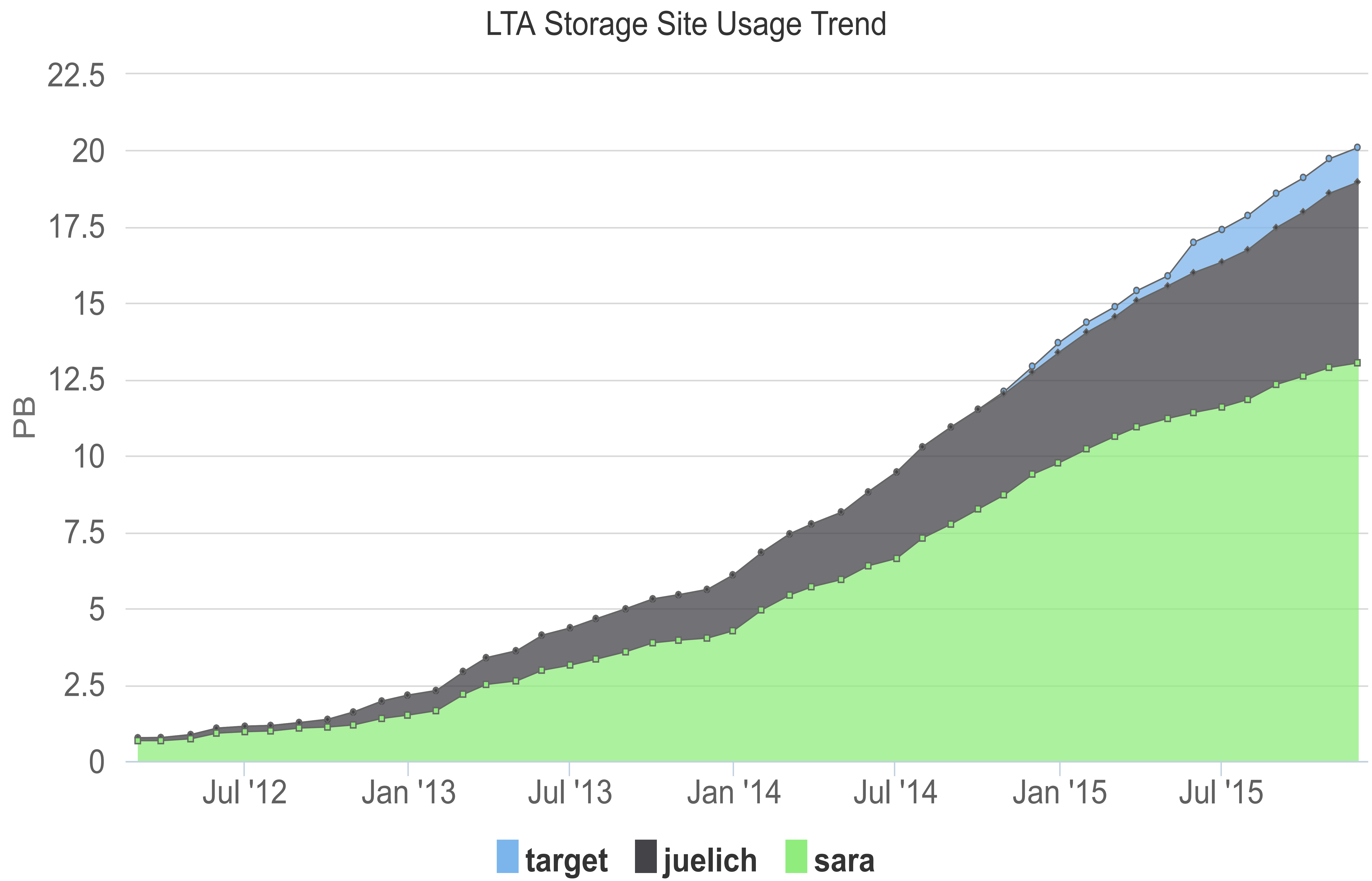
HST Publication Rate

- Assumes the archives are persistent and maintained
- Assumes archival data is open and accessible
- Assumes users retrieving data have resources to process to a science result
- Assumes data products stored are appropriate for general use

LOFAR Data Accumulation

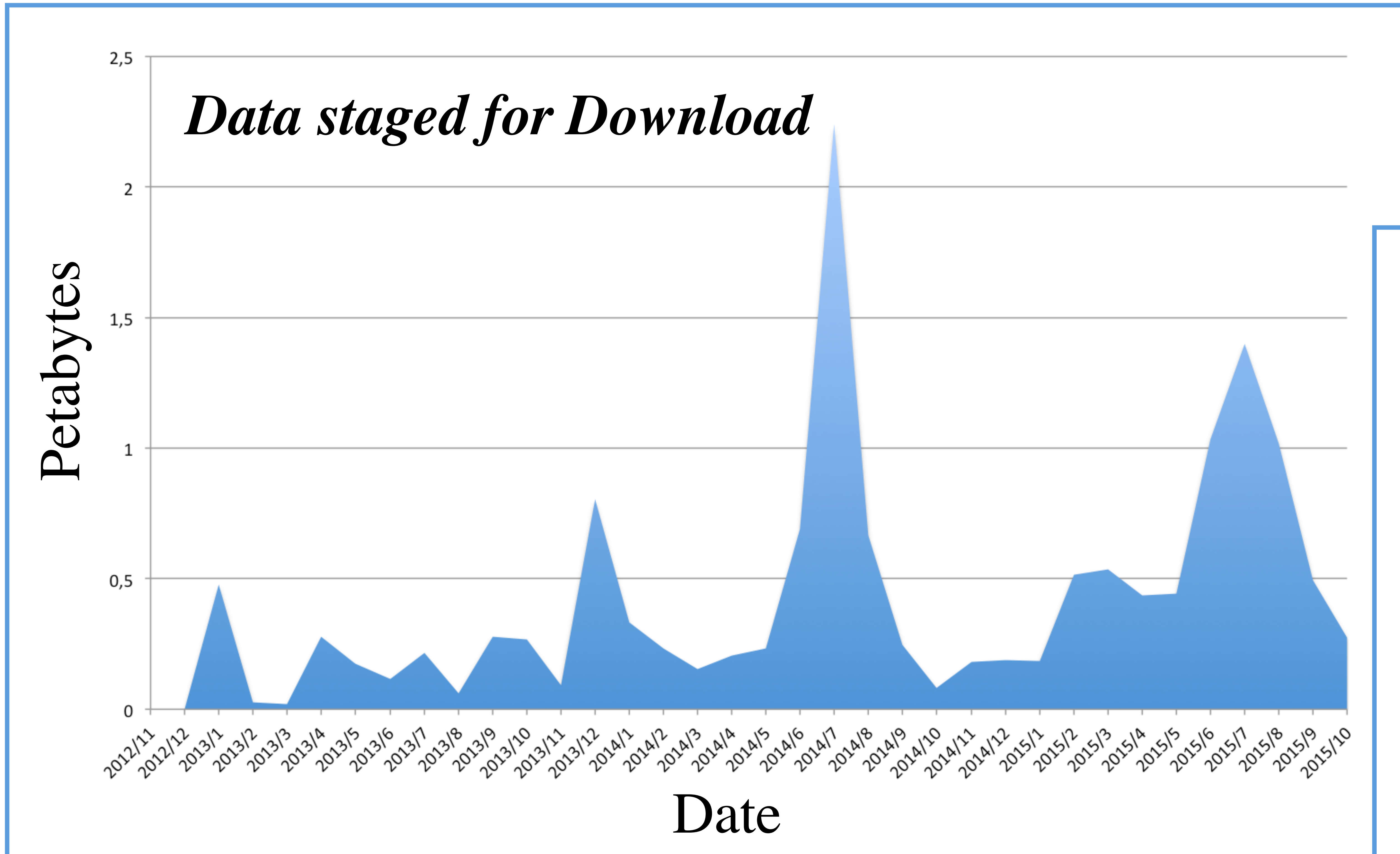
- Data Storage
  - *20.1 Petabytes*
  - *3 PB/yr growth*
  - *3 sites, 2 countries*
  - *300 TB/month ingest*
  - *100 TB/month staged*

- Contents
  - *Over 5x10^{6} products*
  - *10^{9} individual files*
  - *Visibilities, images, and BF data*
  - *Does not include raw visibilities*

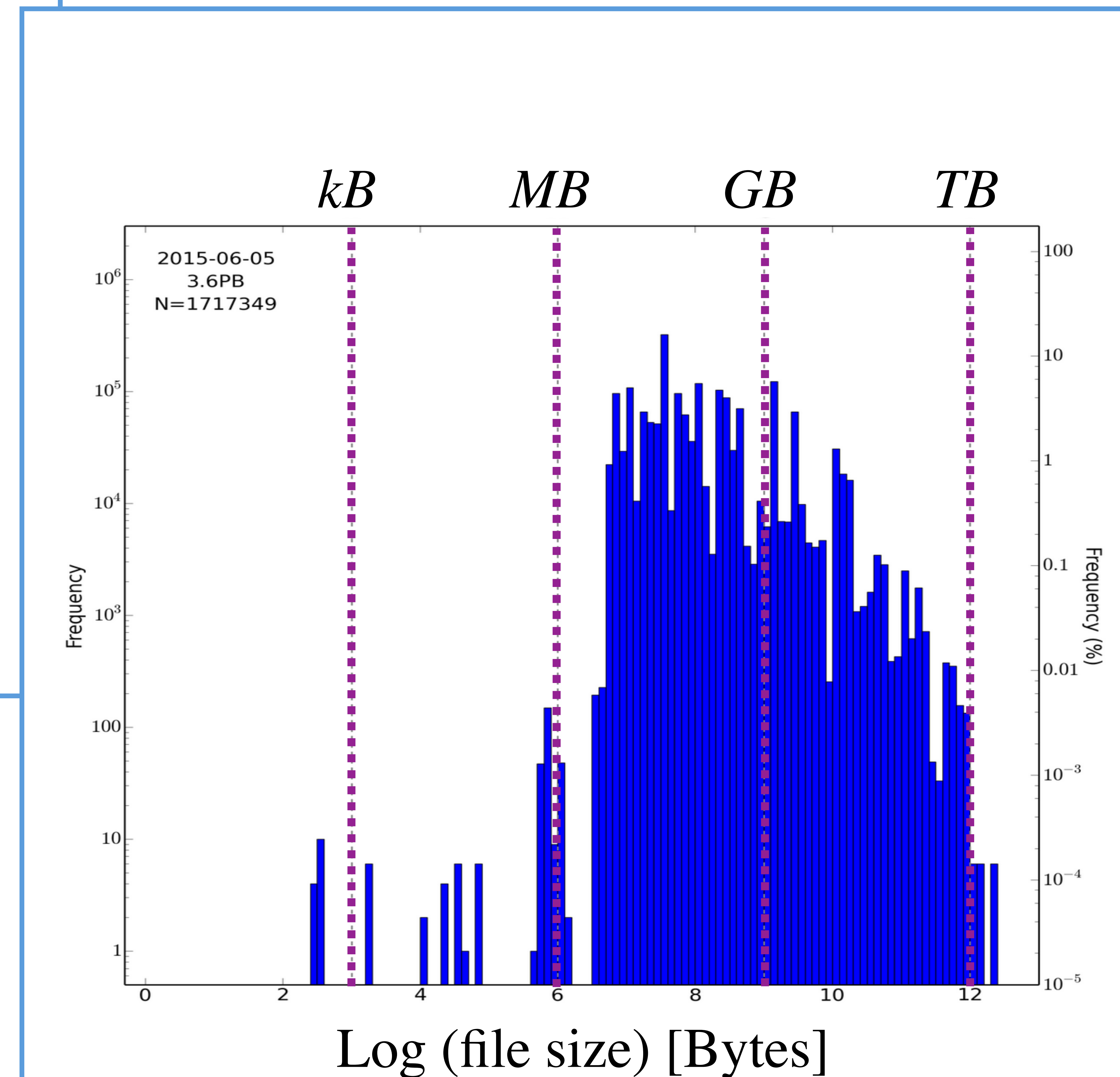


LOFAR LTA team: H.A.Holties, G.A.Renting, Y. Grange, J. Schaap, N.Vermaas, W.J.Vriend

Data Access Patterns



Typical data size is 10-100 Tb
Problematic for many researchers!



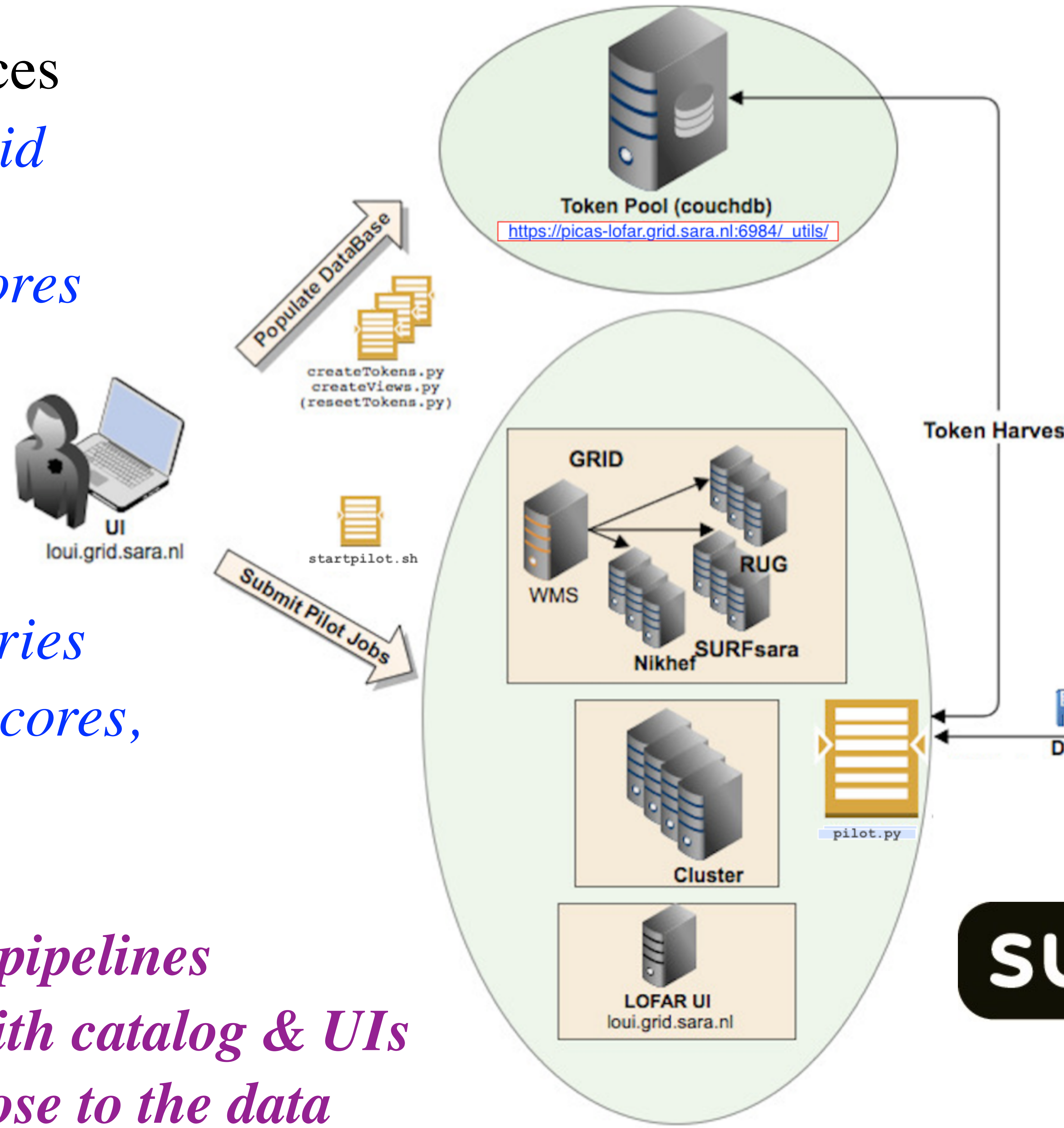
Typical 5 node clusters (320 Gb, 120 cores, 250 Tb) at individual research institutes are NO longer sufficient:

→ *Data transfer from archive to institutes too slow: ~ 10 Mb/s*

→ *Current P/O for a single observation too high: 10 - 100*

Processing in a Distributed Archive

- National NL GRID Resources
  - *14 data centres (3 large grid clusters, 11 smaller ones)*
  - *approx. 10,000 compute cores*
  - *12 PB disk, 170 PB tape*



- Global GRID Resources
  - *170 data centres in 36 countries*
  - *more than 330,000 compute cores,*
  - *500 PB disk, 500 PB tape*

For LOFAR ⇒ Standardized pipelines
Integration with catalog & UIs
Processing close to the data

EGI Federated Cloud for calibrating and analysing Radio-Astronomy data

Susana Sanchez-Expósito [1,*], José Sabater [2], Daniele Lezzi [3], Julian Garrido [1], Lourdes Verdes-Montenegro [1], José Enrique Ruiz [1], Pablo Martín [4]. Raül Sirvent [3], Rosa M. Badia [3], Antonio Ruiz-Falcó [4]

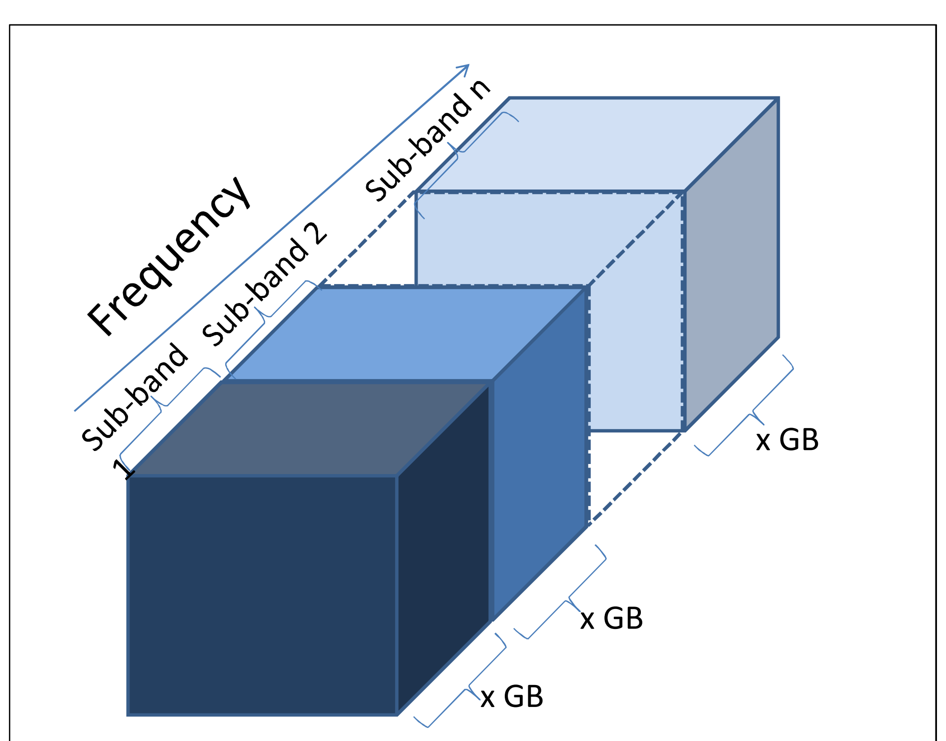
(1) Instituto de Astrofísica de Andalucía – CSIC
(2) University of Edinburgh
(3)Barcelona Supercomputing Center
(4) Fundación de Supercomputación de Castilla y León

Use Case 1: LOFAR data calibration

Fig.1. LOFAR datacube representation

An interferometer as LOFAR correlates the signals from several antennas, generating the so-called measurement sets. They are a kind of **datacubes** (3D data)**:** two Fourier spatial coordinate axes plus a spectral axis.

A datacube can reach several **terabytes,** depending on factors as the amount of involved antennas, the observation time, as well as the amount of observed subbands – i.e. frequency intervals-. LOFAR telescope allows **up to 488 subbands**, which can reach several GBs. **Each subband can be processed independently what allows the parallelization of the whole datacube calibration**.

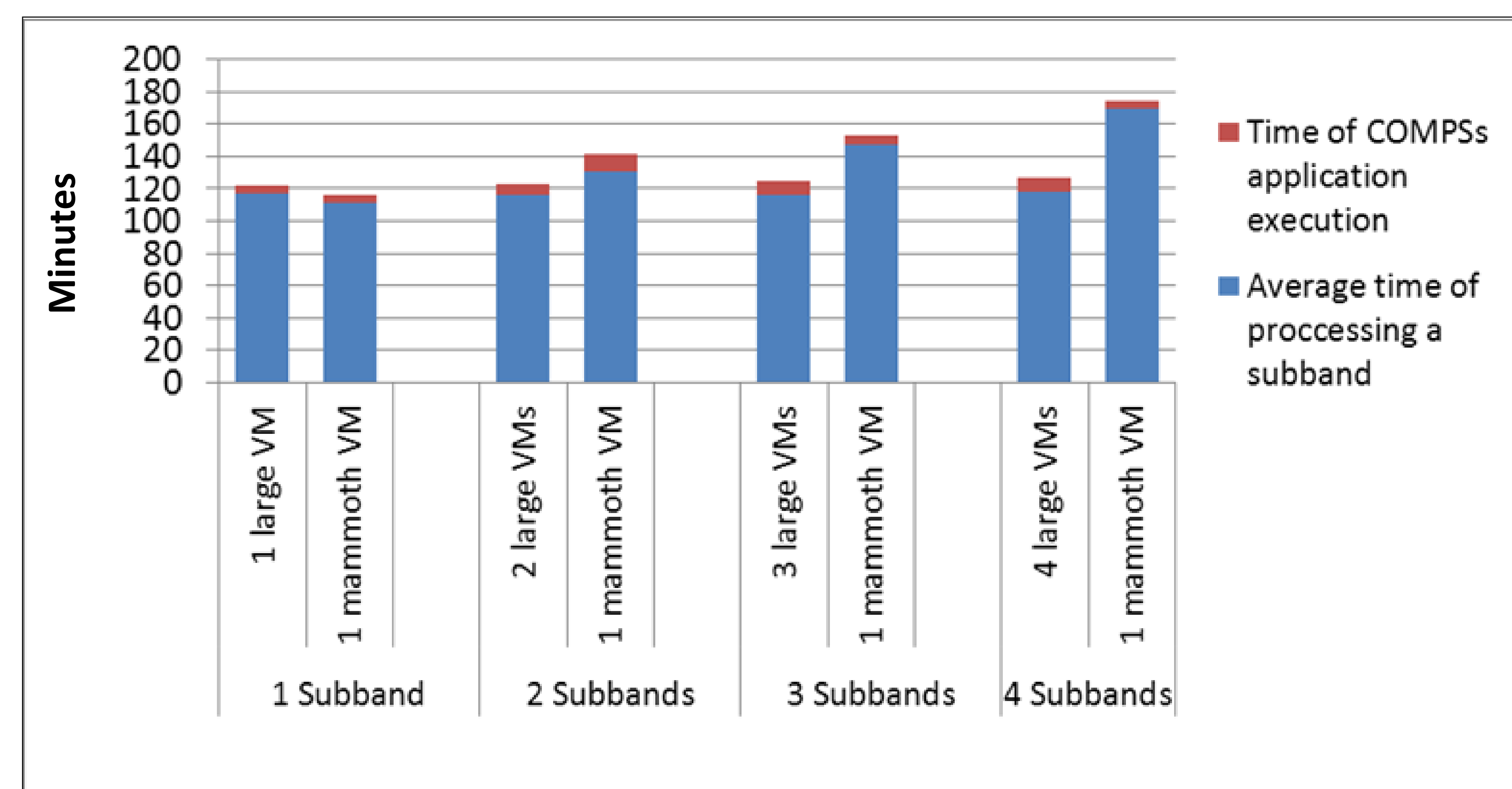



```
import subprocess
import sys
import os
from pycompss.api.task import task
from pycompss.api.parameter import *

@task( script_name = FILE)
def iter_calib(script_name):
   os.chmod(script_name,0744)
   subprocess.call(script_name)
   print "end executiong"
if __name__ == "__main__":
   args = sys.argv[1:]
   DATA_PATH=args[0]
   TEMPLATE_FILE=args[1]
   f=open(TEMPLATE_FILE,'r')
   content=f.read()
   f.close()
   list_f=os.listdir(DATA_PATH)
   for directory in list_f: # Iterate over the data inputs
      if os.path.isdir(DATA_PATH+"/"+directory):
        new_content=content.replace("INPUTDATAPATH",directory)
        script_name="job"+directory+".sh"
        f=open(script_name,"w")
        f.write(new_content)
        f.close()
        iter_calib(script_name)
```

Fig.2. COMPSs application. It iterates over the subbands, executing for each one a COMPSs task that calls the LOFAR software. Through a simple interface for describing the methods, COMPSs is able to analyse the dependencies among them, to match their requirements with the available resources and to orchestrate their execution on VMs.

Fig.3. Execution time. This figure shows the results of different tests in which the application has been configured to calibrate from 1 to 4 sub-bands, and its tasks have been distributed either on a high capacity VM (mammoth=32GB memory + 8 cores) or on several smaller VMs (large=8GB memory + 4 cores). Since each subband is processed in parallel, **the executions for calibrating one subband take approximately the same time than those for calibrating several subbands**. The results also reveal that **the execution time for the whole COMPSs application (in red) is slightly higher than for the tasks (in blue)**. Thus we can state that the time to start and contextualize the VMs is not significant. In addition, the time for the applications running on mammoth is higher than the applications whose tasks have been distributed on smaller VMs. This would mean that **distributing the tasks among several small VMs is more efficient than gathering them in a VM with high memory capacity and amount of cores.**



Distributing Operational Software

Docker:
- Overcome bad coding practices of astronomers
- Robust (runs on most platforms)
  - Cloud, cluster, laptop
  - Repeatable
- Seamlessly ship pipelines across platforms

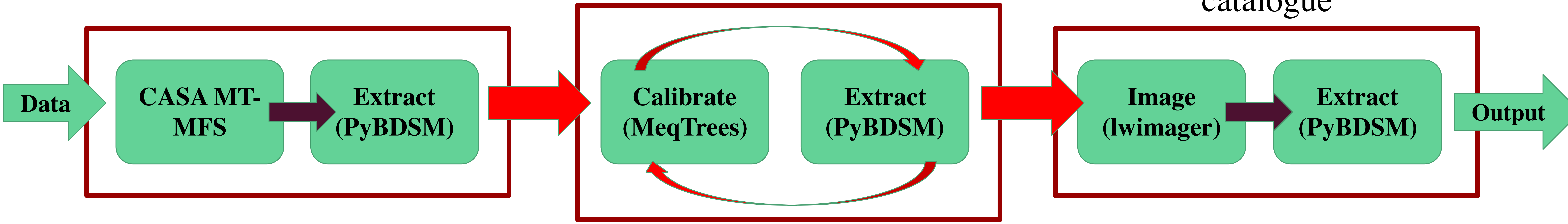
Build, Ship, Run

Dockerized task:
imaging, source finder, self-cal, etc.

Make initial model

Calibration loop

Make final image and/or catalogue



(Sphesihle Makhathini, Rhodes University)

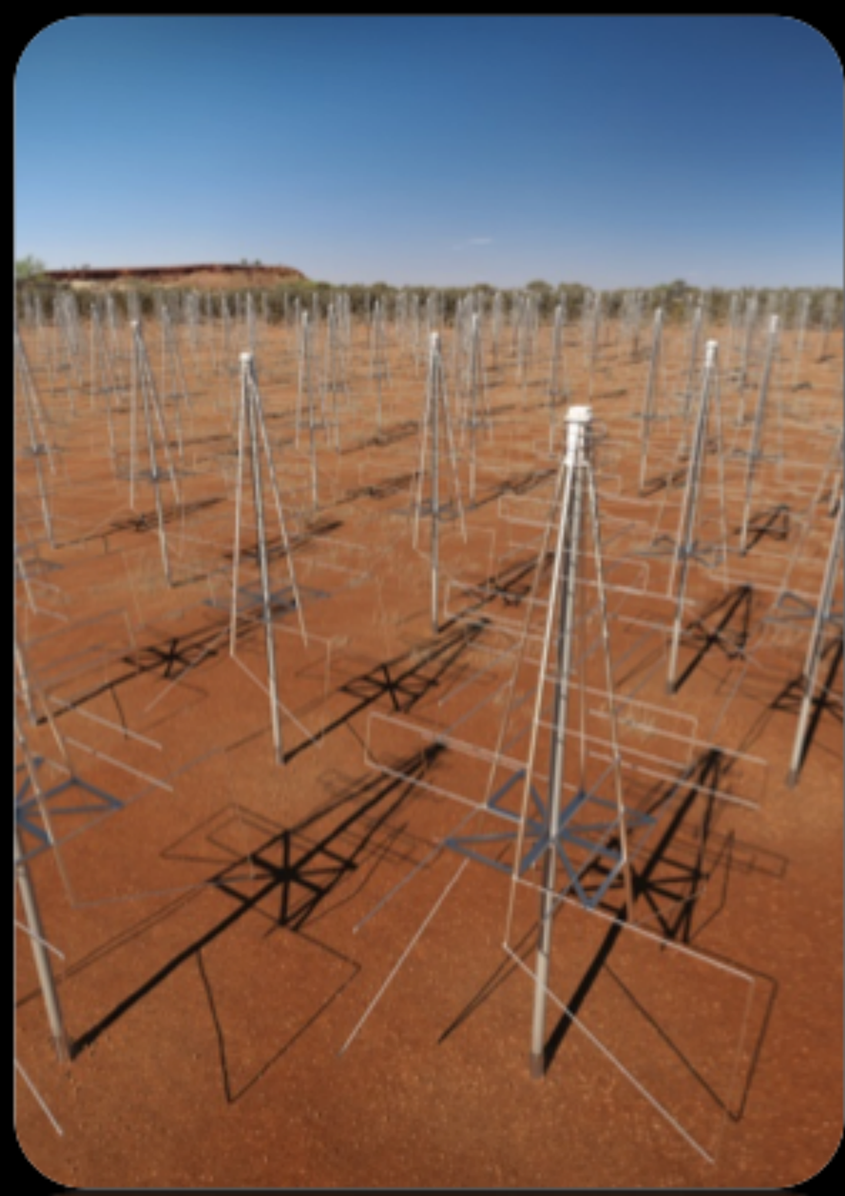
Tradeoffs of KSPs vs. General Users

	Pros	Cons
Key Science Projects	Reduced operational computing budget	Community must provide necessary computing resources
	Potentially smaller science archive	KSP-tailored data products could limit discovery space
	Reduced operational software overhead	Increased support for multiple external platforms required
Individual User Projects	Wider range of science and multiplicative factor for science from archived data	Increased operational and analysis computing budget
	Increased discovery potential for archival science data	Larger archive size
	Wider access to non-expert users	Increase in necessary user support

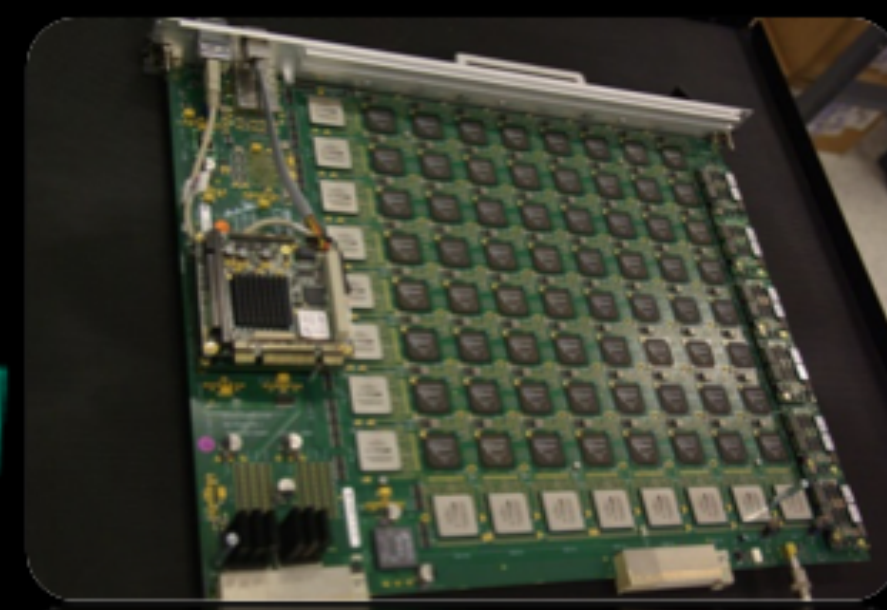
Data scale and network costs can undermine traditional operational advantages offered by KSP model

SKA System Data Flow

Antennas



Digital Signal Processing (DSP)



Transfer antennas to DSP
2020: 5,000 PBytes/day
2030: 100,000 PBytes/day

Over 10’s to 1000’s kms

To Process in HPC
2020: 50 PBytes/day
2030: 10,000 PBytes/day

Over 10’s to 1000’s kms

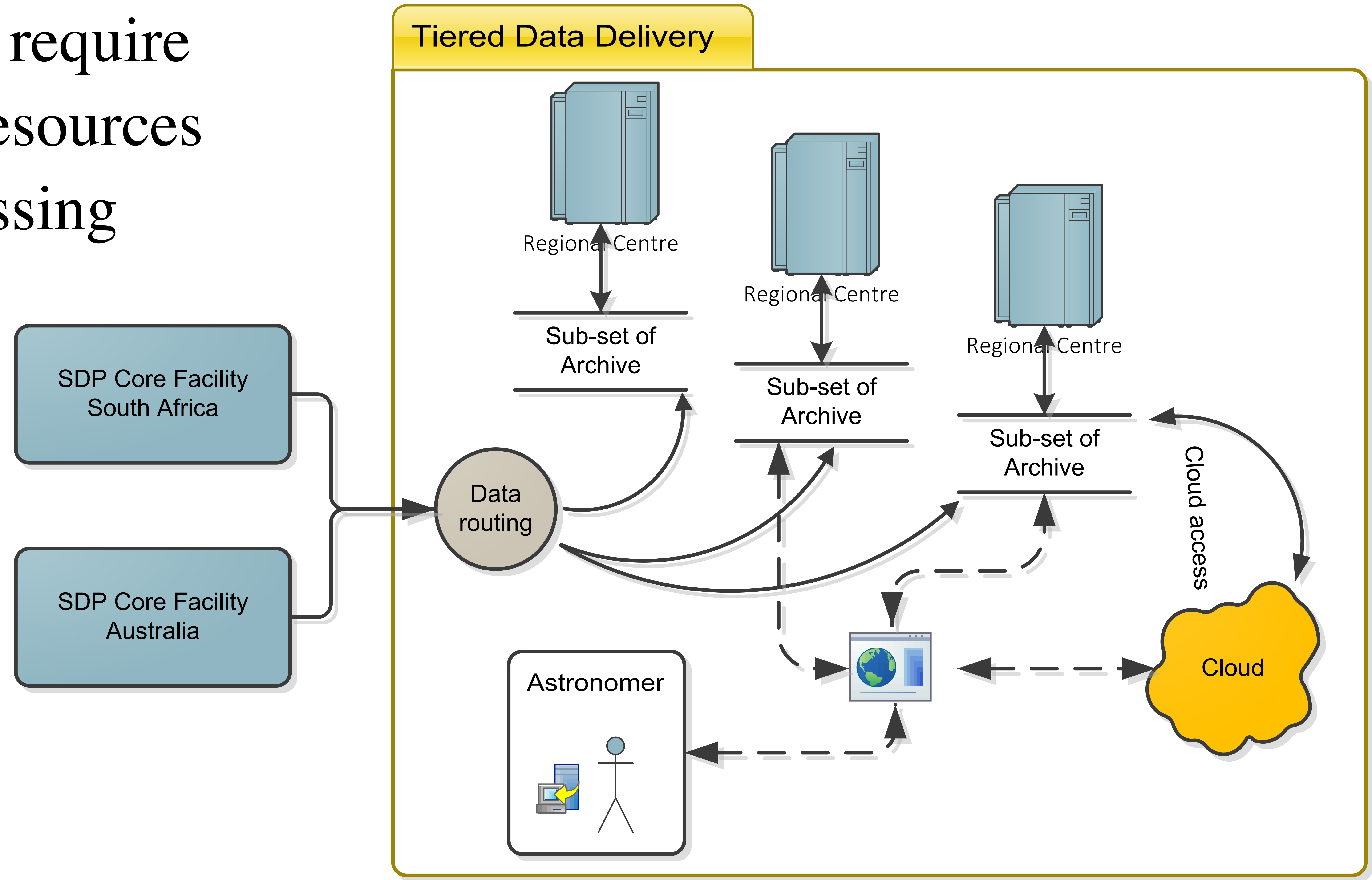
High Performance Computing Facility (HPC)

HPC Processing
2020: 300 PFlop
2028: 30 EFlop

30 – 300 Pbytes / year of fully processed science data products

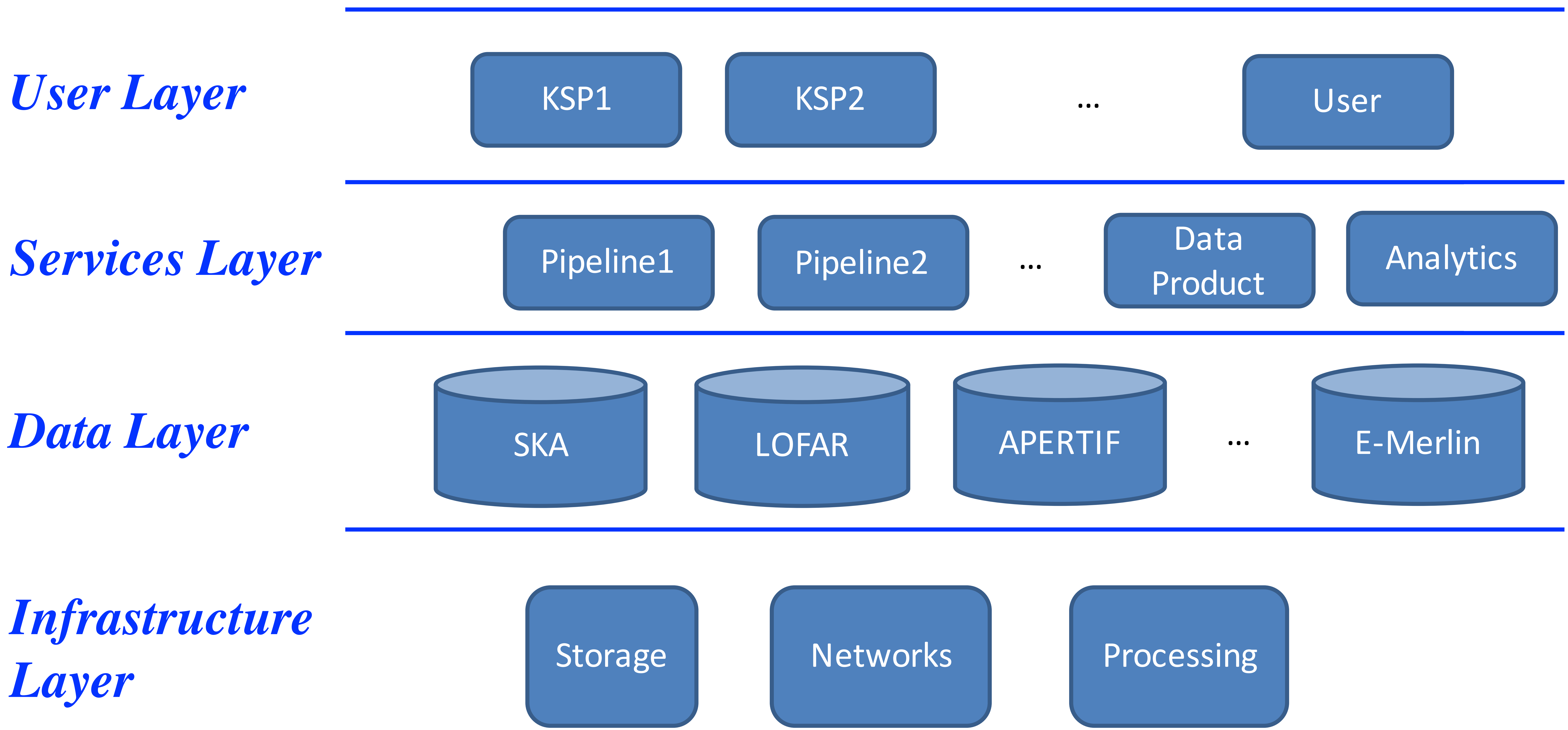
SKA Regional Science Centres

- Science extraction will require significant additional resources beyond standard processing
- Need for RSCs noted by SKAO board
- Likely there will be multiple RSCs to support community



RSCs will be the working surface for SKA science!

Regional Centres Structure

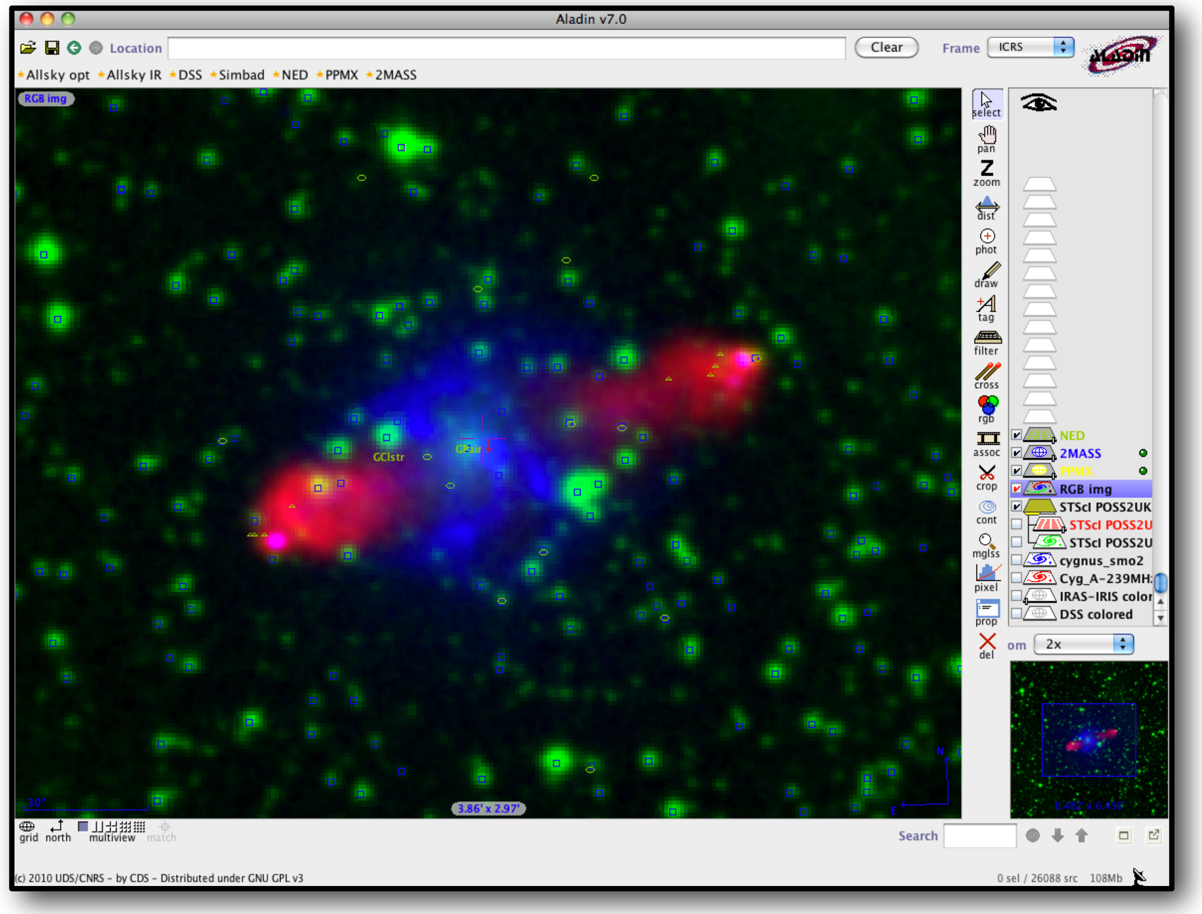


Also need an expertise layer!

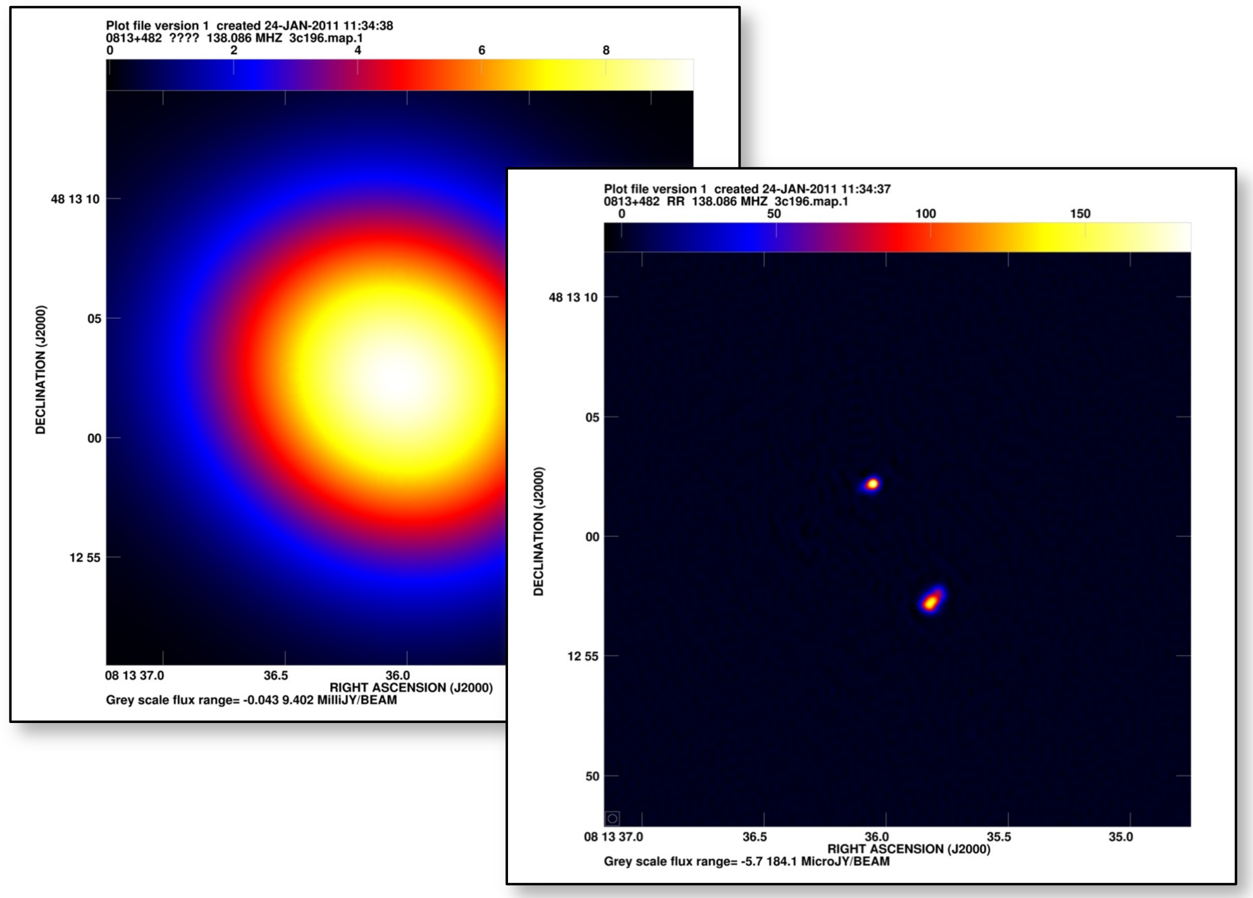
Data Centre Functionality

Data Discovery

- Observation database
- Associated metadata
- Quick-look data products
- Flexible catalog queries
- Integration with VO tools
- Publish data to VO



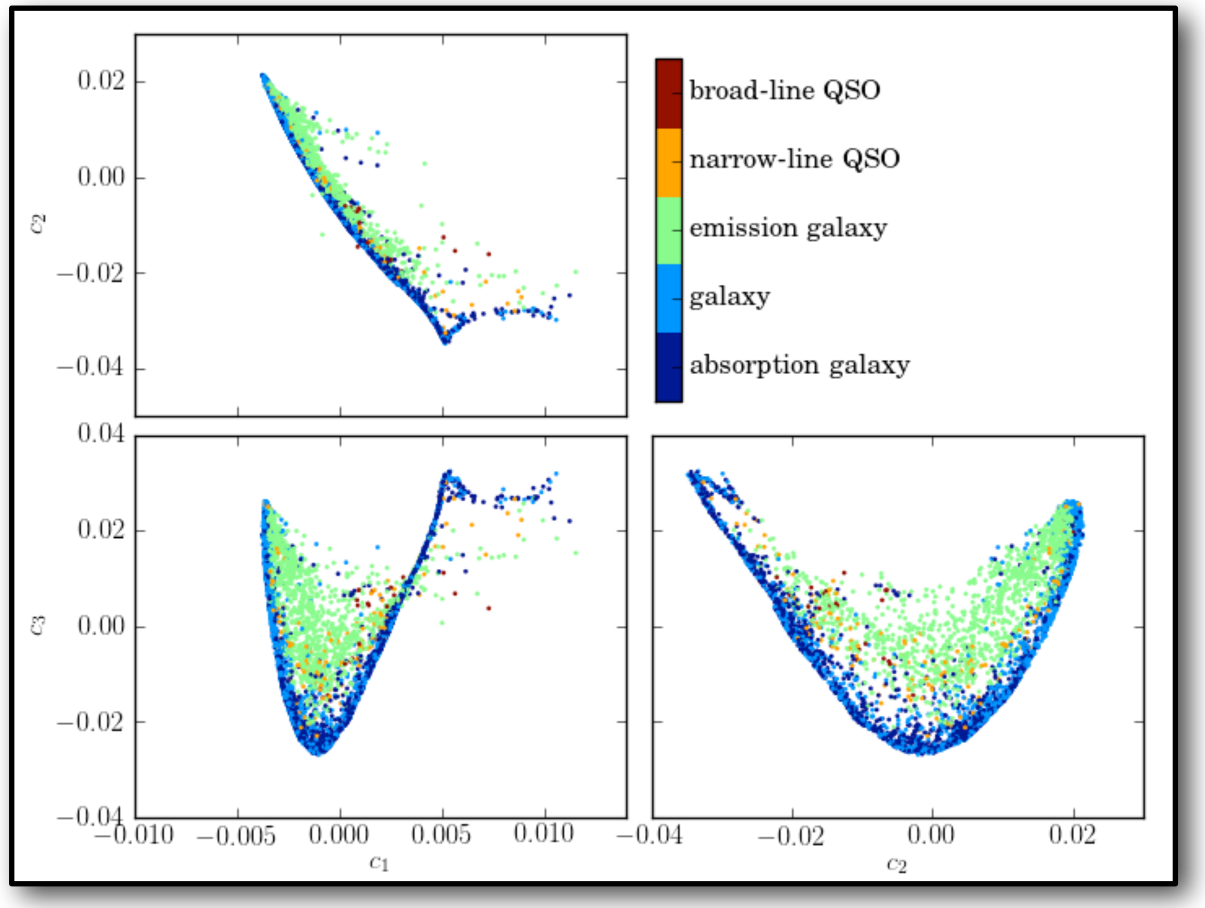
Data Processing



- Reprocessing and calibration
- High resolution imaging
- Mosaicing
- Source extraction
- Catalog re-creation
- DM searches

Data Mining

- Multi-wavelength studies
- Catalog cross-matching
- Light-curve analysis
- Transient classification
- Feature detection
- Visualization



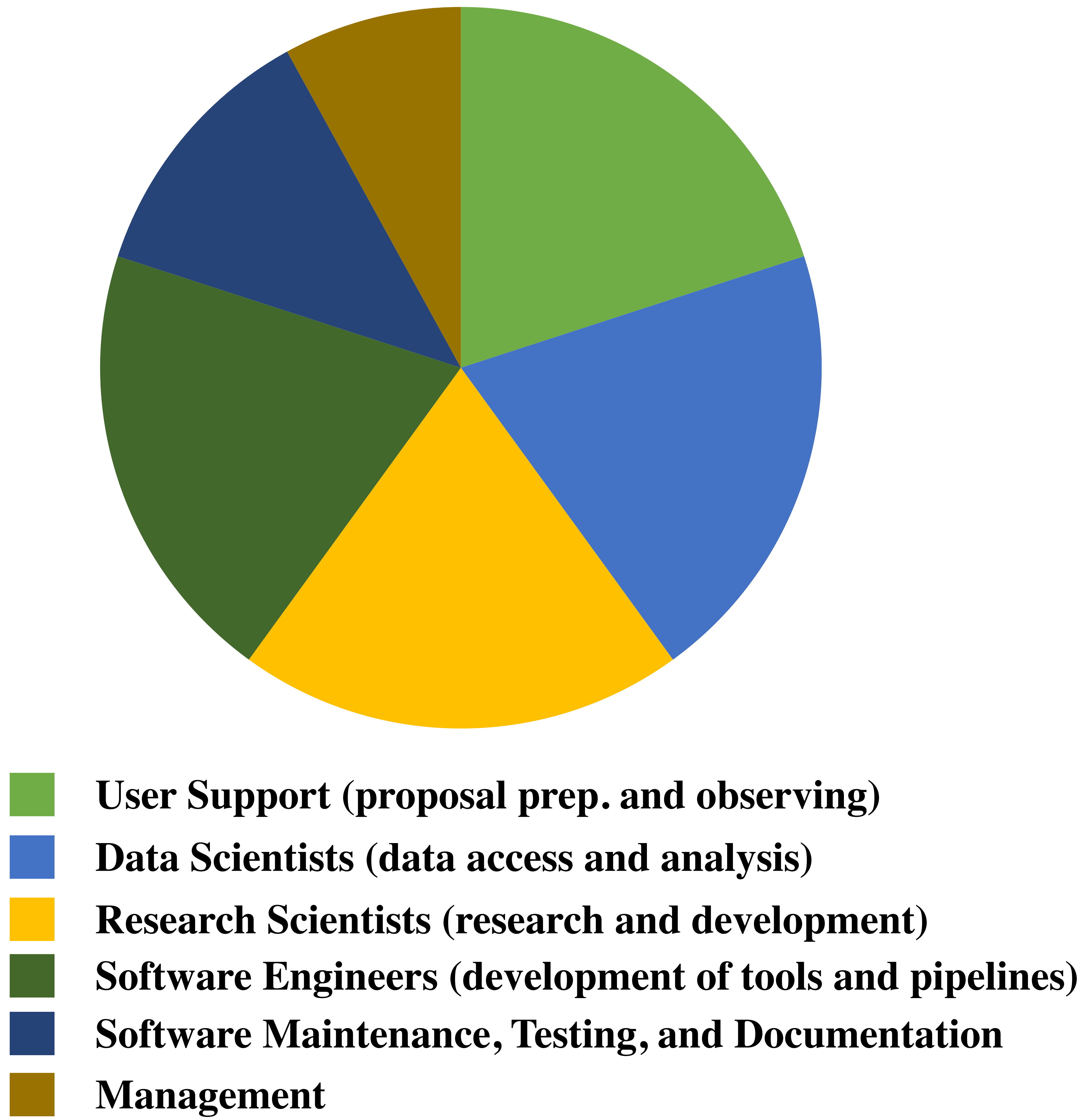
Regional Centres Capabilities

Data oriented operations:

- Data archiving and curation
- Data management, discovery, and access
- Automated processing and reprocessing
- Generation and storage of science products
- Continued pipeline development

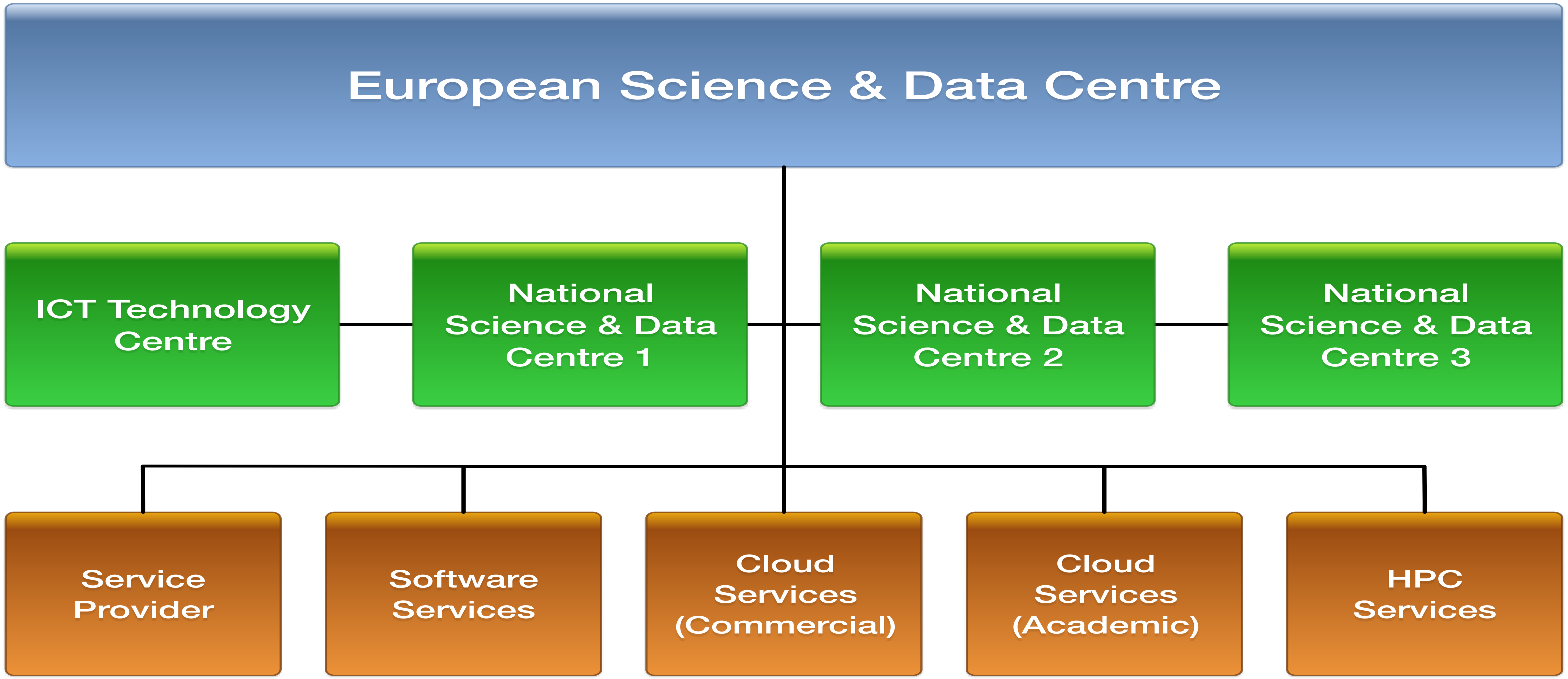
Science oriented operations:

- Portal-based data product access
- Interface to processing pipelines
- Interface to VO discovery and analysis tools
- Support for custom user analysis
- Development of new algorithms and tools
- End-to-end astronomer support
- Community education & outreach
- Face-to-face user support
- 24/7 help desk



Requires wider range of staff skills!

Working Group on European RSCs



- Create a European-scale federated Science Data Center for the SKA
- Coordinated engagement with national ICT communities, industry, and service providers
- Facilitate shared development, interoperability, accessibility and innovation
- Serve as coordinating entity to pursue EC and other funding
- European counterpart for engagement with other Science Data Centres internationally

Summary

- *Operational choices tightly coupled to computing costs*
- *Need to consider the full cost of science extraction*
- *Increasing data scales can reduce KSP operational advantages*
- *SKA scales will require distributed, large-scale data centres*