Obelics WP 3.3 D-INT LOFAR status

Tammo Jan Dijkema
Obelics F2F Madrid, 15 September 2016
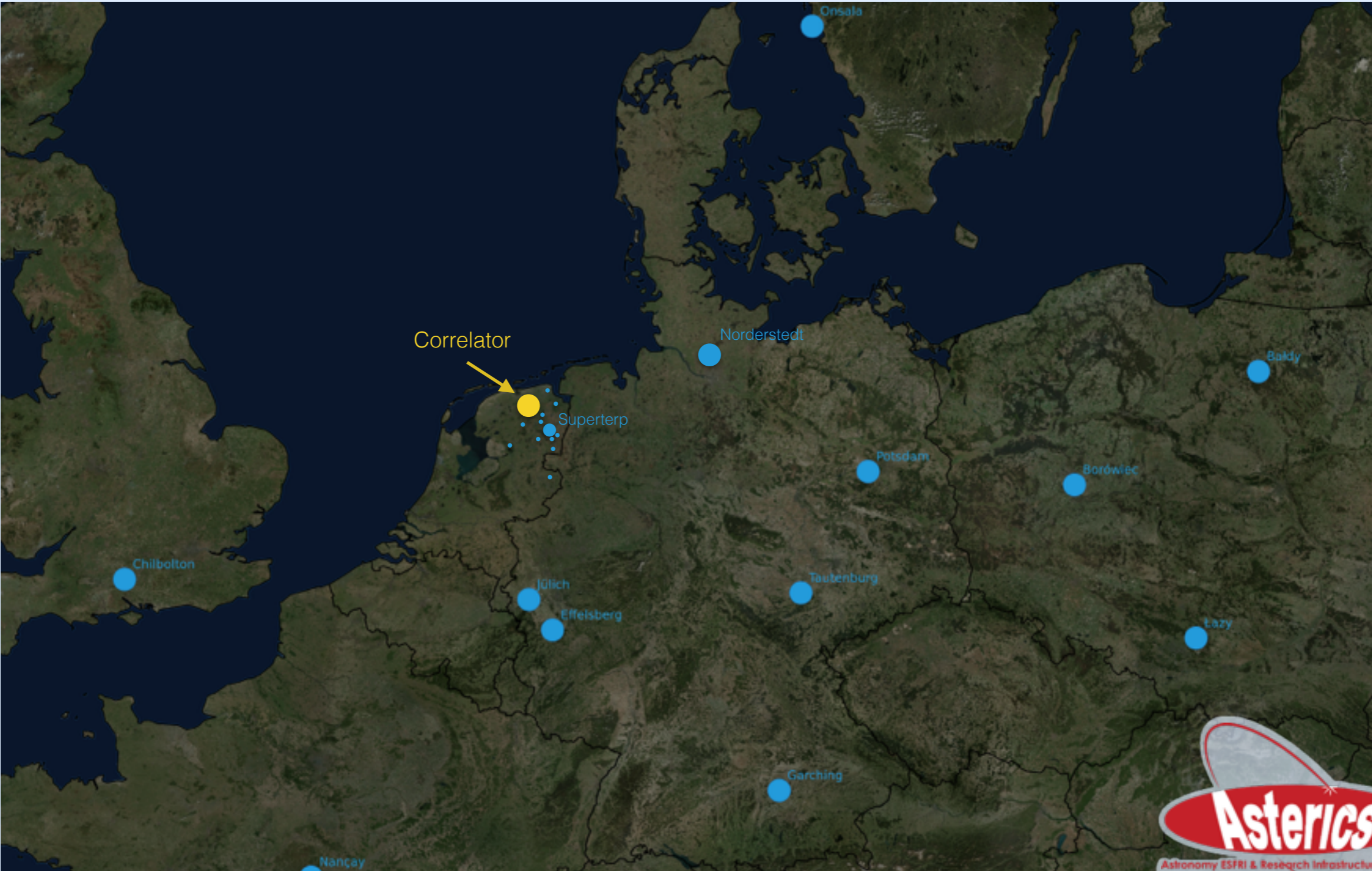
ASTRON

Netherlands Institute for Radio Astronomy

ASTERICS
653477

Asterics

Outline

- LOFAR overview

- LTA status

- Processing requirements

LOFAR overview

Processing overview

Antennas / Stations	Realtime System	Offline Processing	Long Term Archive
on-site reduction	8 node GPU correlator	~ 100 node CPU cluster	4 grid sites ~20 Petabyte

200 Gbit/s → **80 Gbit/s** → **2 Gbit/s** →

LOFAR data flow

Correlator

...

...

CEP4

LTA
(Archive)

CEP3

...

...
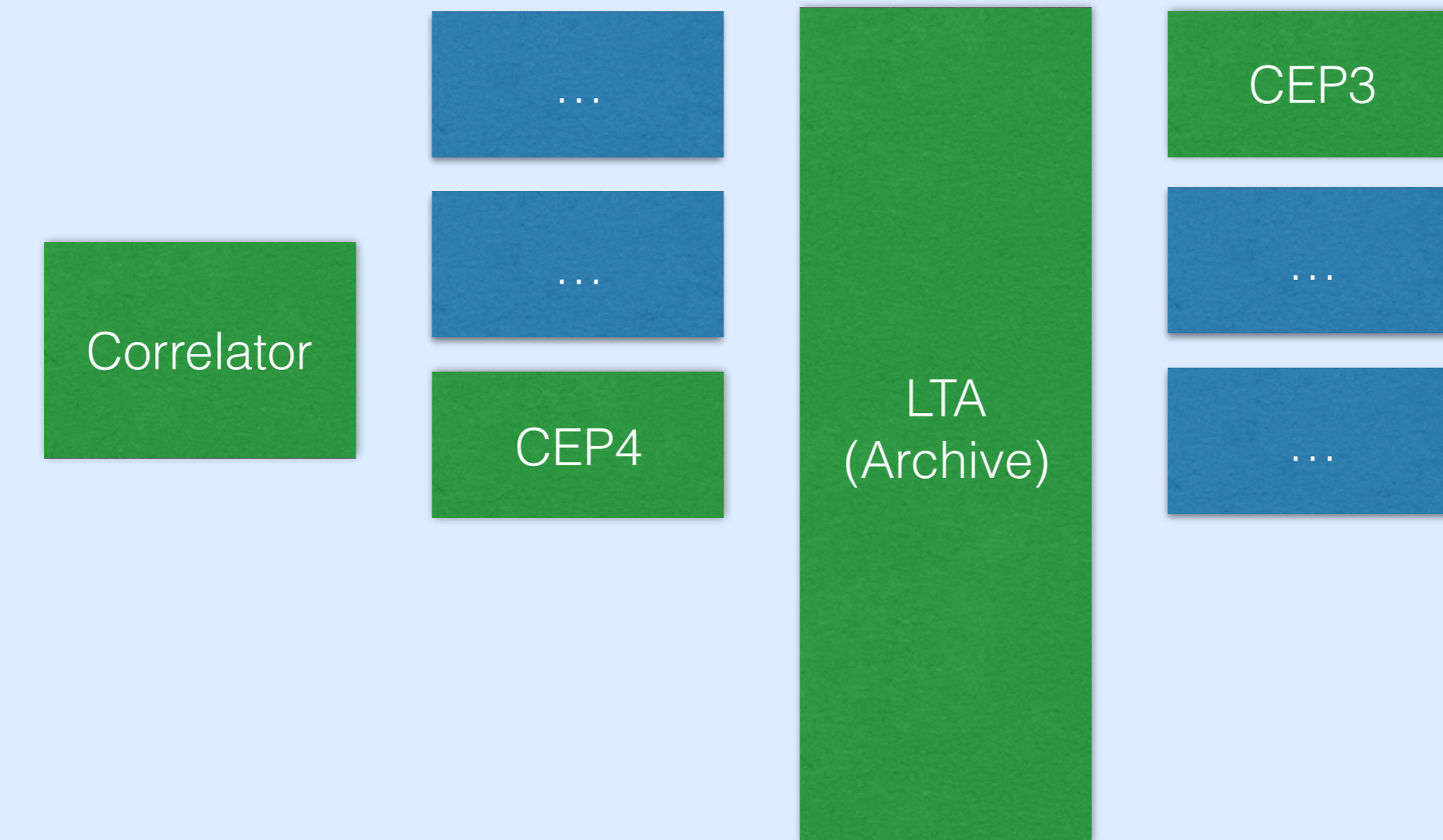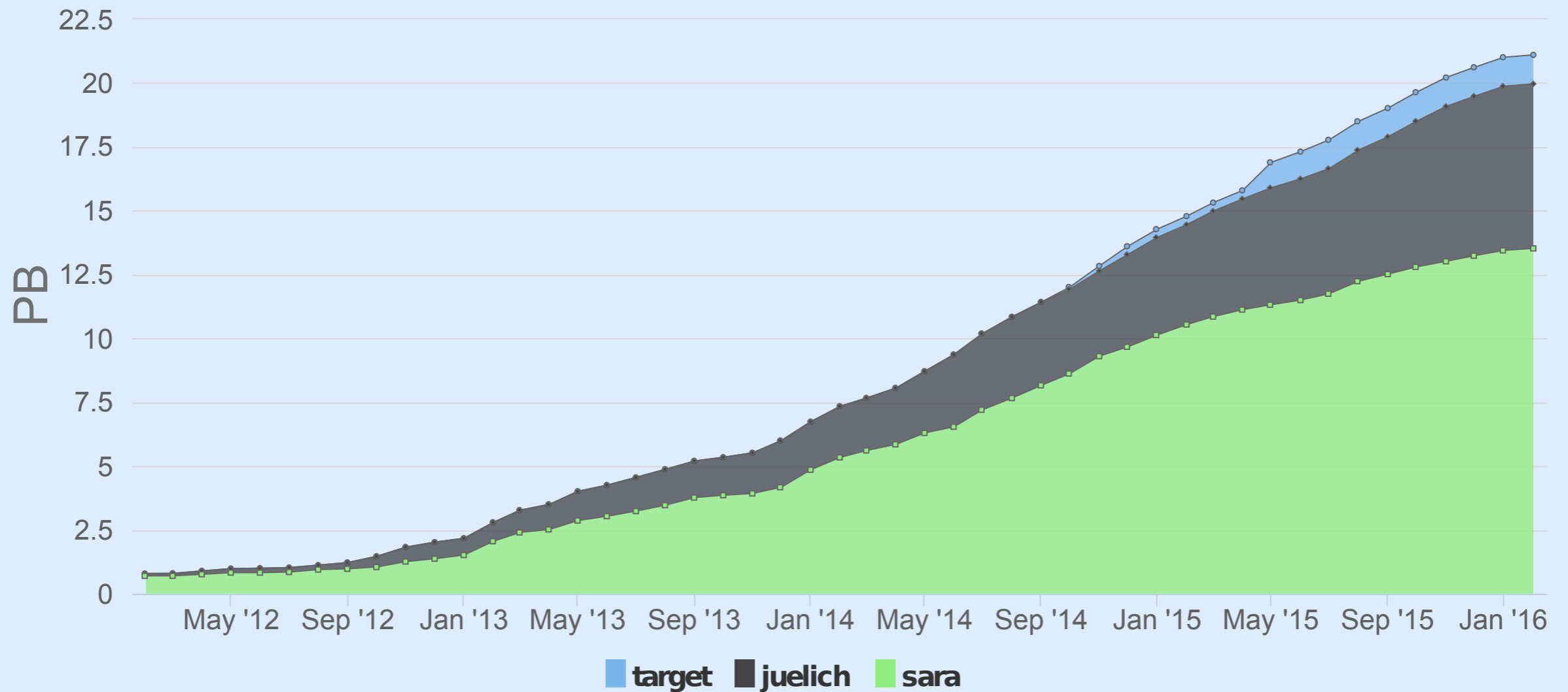
operated by observatory

operated by others

Long term archive status

LOFAR imaging pipelines

- Pipelines on raw data run on dedicated cluster CEP4
  - Averaging soon: less disk space and transfer time LTA

- Other pipelines can be run on dedicated cluster CEP3 **or** your own cluster → packaging issues
  - Software is heterogeneous

Processing overview

- Ideally, $\dfrac{\text{Processing time}}{\text{Observing time}} \lesssim 1$

- With the CEP4 100 CPU (×20 cores / CPU) cluster, this is not yet achieved for the stage up to LTA

  - Depending on observation

- Newest pipeline (FACTOR) takes about P/O = 10

- Need for more efficient algorithms

LOFAR pipelines on CEP4

- Pipelines run inside Docker container
    - Whole pipeline in one container
    - Docker used mainly for build management, reproducibility
    - Flexible to run almost any software

- Custom python pipeline framework
- Execution over ssh, local, slurm, mpi, …
- Generic format for specification

LOFAR Long Term Archive

- Authentication & Authorization through grid certificates

- Processing on the grid sites is planned, but hard

Questions?