

ESCAPE-ing LOFAR Data

Use Case - Prefactor



LOFAR



ASTRON

V.N.Pandey (ASTRON/Kapteyn Inst.)

Yan Grange (ASTRON)

ESCAPE WP2 Meeting (Remote)

May 20, 2020



**university of
groningen**

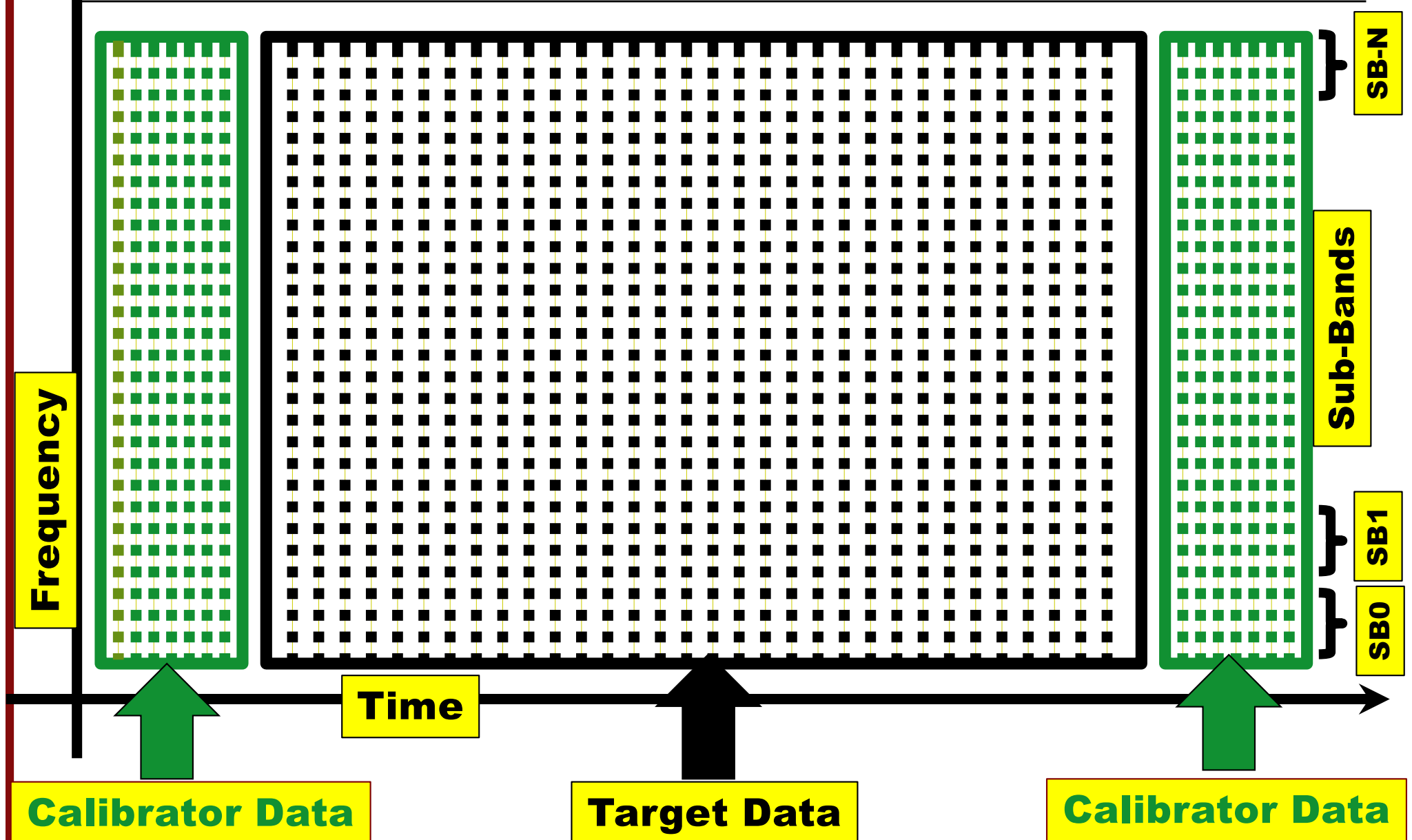
faculty of science
and engineering

kapteyn astronomical
institute

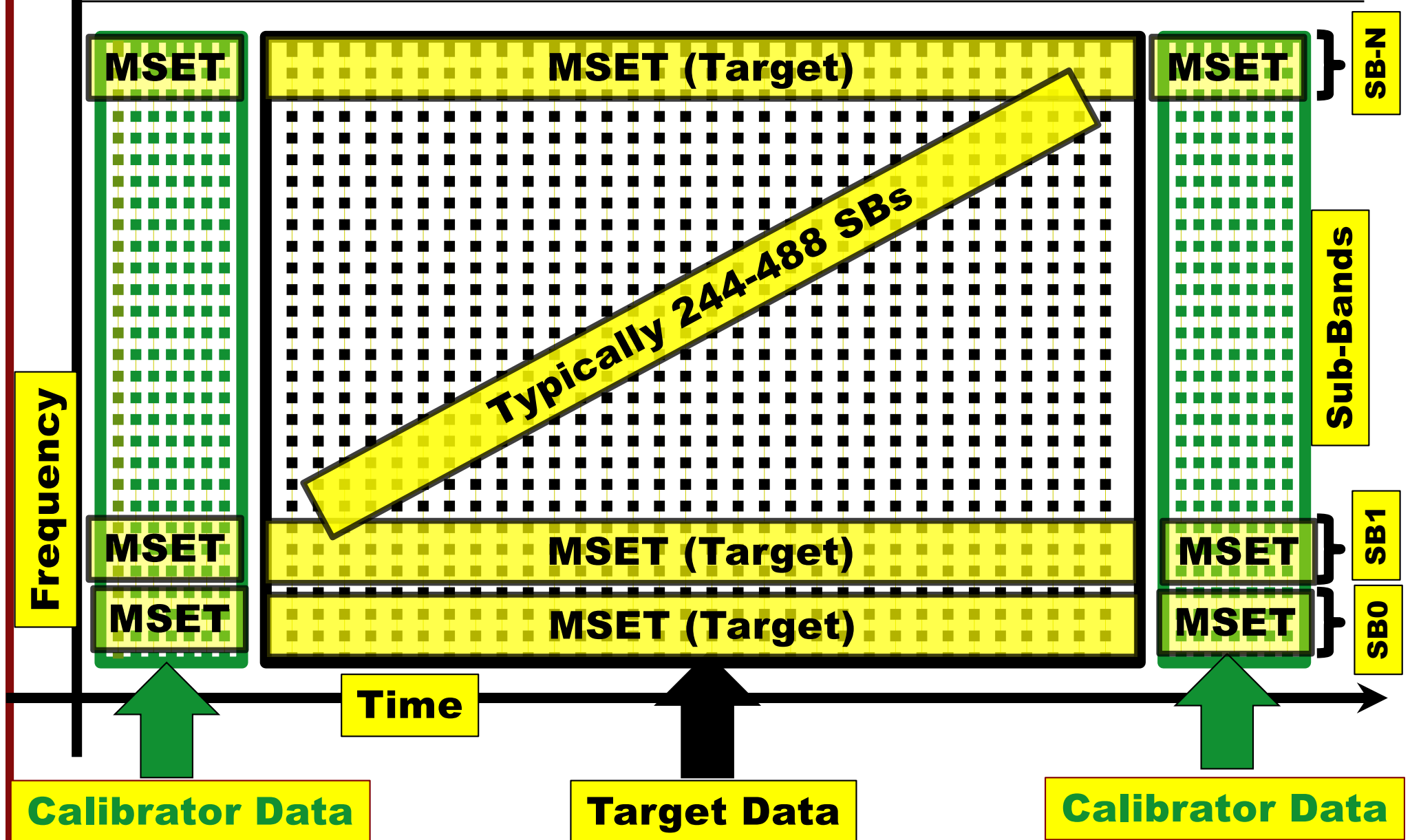
LOFAR - Pipelines

- ☐ Prefactor Pipelines (Imaging, sky surveys)
 - may contain several individual pipelines
 - ☐ Transients Pipeline (intermittent and time varying)
 - ☐ Long Baselines Pipeline (small ang-size sources)
- } CPU
-
- ☐ EoR pipeline (first stars/galaxies in the Universe)
 - ☐ Pulsar Pipeline (ex. rotating Pulsars/Neutron stars)
 - ☐ AARTFAAC Pipeline (streaming data, all sky)
- } GPU
-
- ☐ <https://www.astron.nl/citt/prefactor>

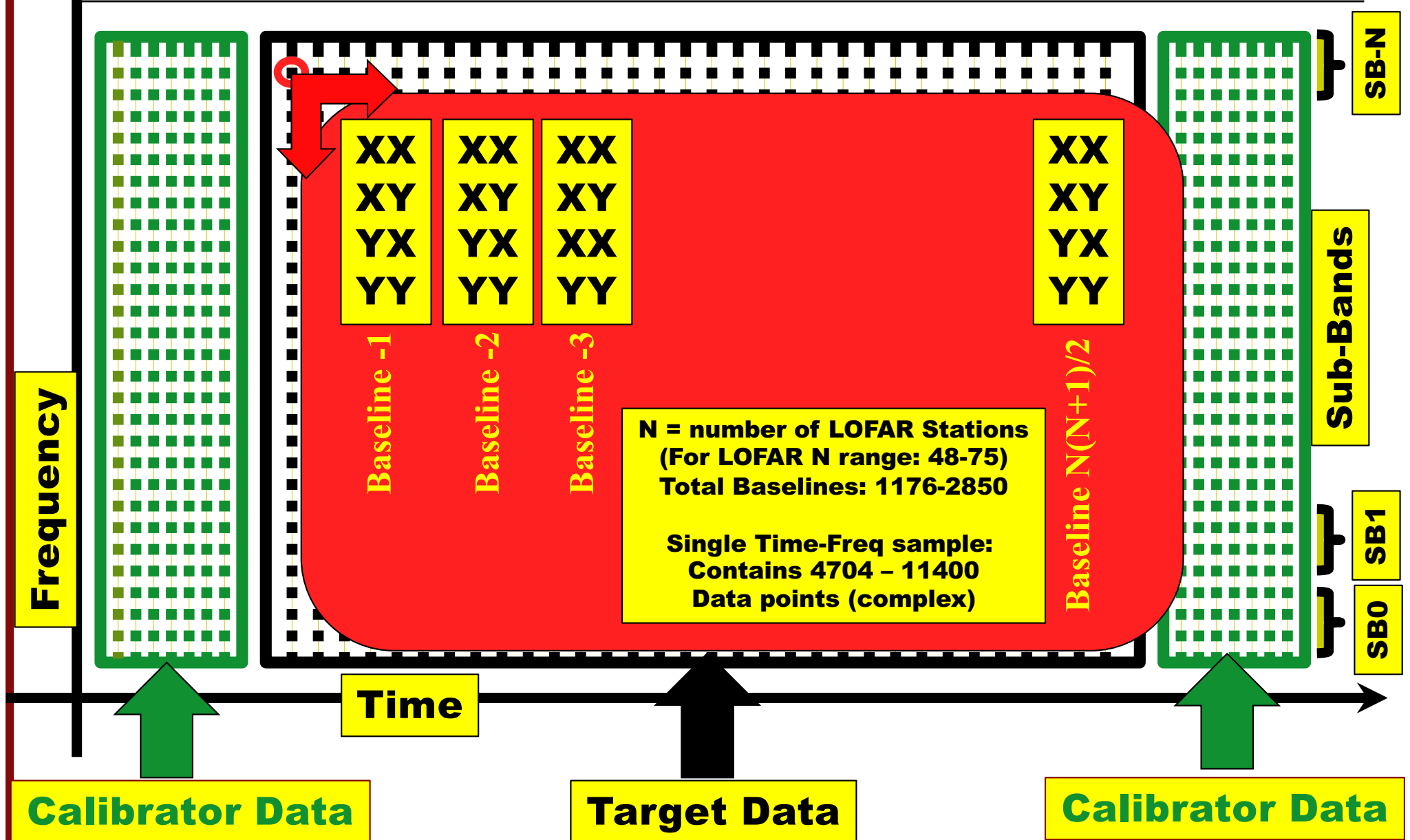
Prefactor data – bird's eye view



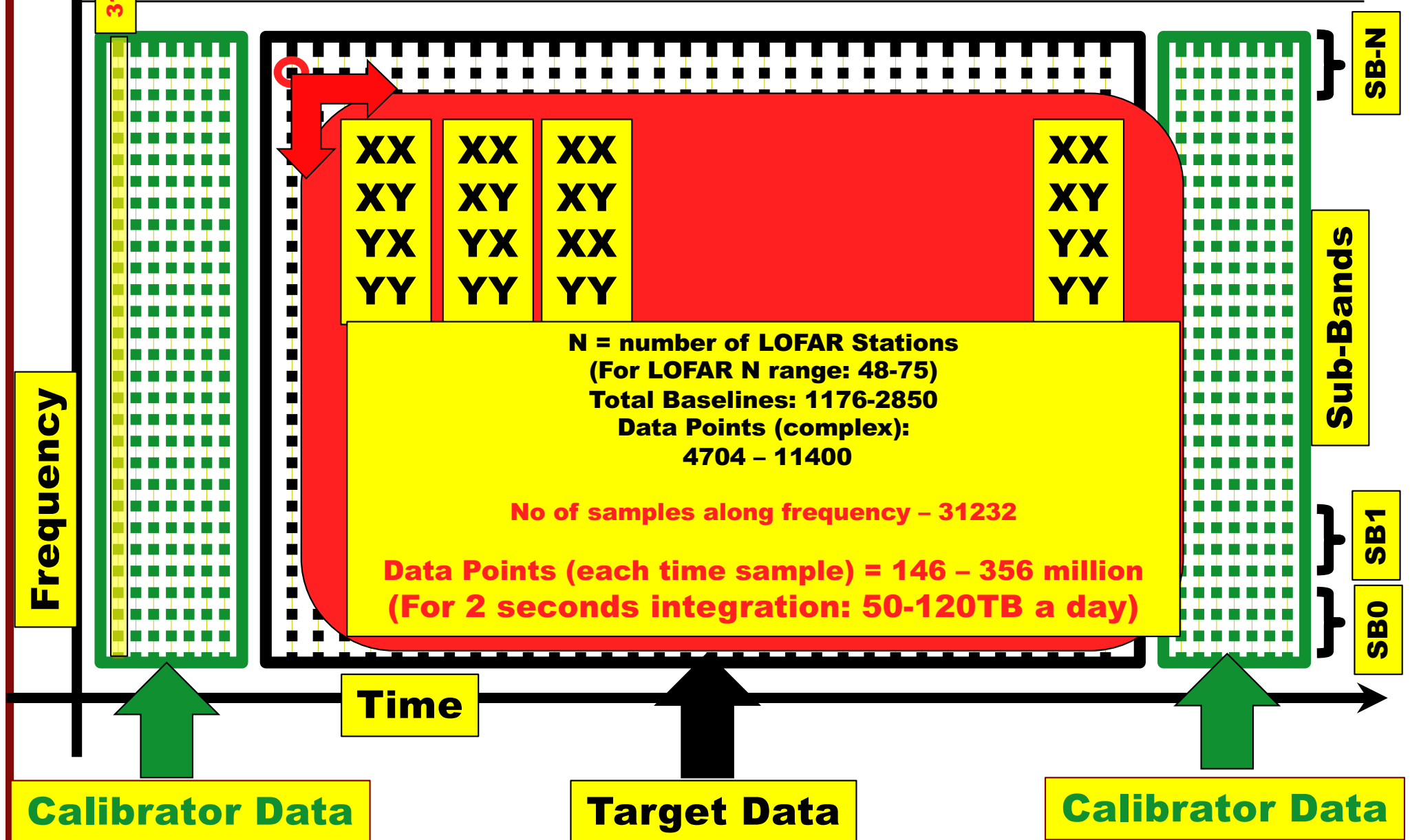
Prefactor data – organisation



Data sample – Zoom in



Data Size



Prefactor status - Software

- ❑ **Pipeline Software (needs LOFAR suite,)
encapsulated into self contained singularity Image**

`-rwxr-xr-x 1 pandey 1001 639M May 19 12:07 lofar.py2.simg`

- ❑ **Built using singularity 3.5.2 (>3.2 should be ok)
- (separate host/guest filesystem,..)**

- ❑ **Uses Python 2.7.3**

- ❑ **Available as tarball (includes LOFAR test data)**

`-rw-rw-r-- 1 pandey 1001 3.8G May 18 19:59 A_ESCAPE_001.tar`

- ❑ **Documentation (text based)**

`-rw-r--r-- 1 pandey 1001 12692 May 10 19:56 readme_A_ESCAPE_001.txt`

Prefactor - Configuration

☐ Pipeline configuration

```
-rw-r--r-- 1 pandey 1001 1266 May 10 16:25 pipeline.cfg  
-rw-r--r-- 1 pandey 1001 89059 May 10 16:22 Pre-Facet-Calibrator.parset  
-rw-r--r-- 1 pandey 1001 69463 May 10 16:23 Pre-Facet-Target.parset
```

☐ Can be used to set up parameters like:

- Data Paths
- No. of processes
- Threads per process
- Memory usage

☐ Steps - flagging bad data, Avg, calibration, interpolation, applying corrections, compression, imaging.

☐ Each MS processed independently but at the end of a step may involve information sharing (e.g. interpolation)

☐ Data combination (optional)

Prefactor - Status

- ❑ **Successfully experimented on several places (DESY, RUG, ASTRON)**
- ❑ **LOFAR Data for prefactor is on Data Lake (61GB, 244SBs total ~ 15TB)**

512 /pnfs/grid.sara.nl/data/escape/disk/LOFAR/557202/

61727488000 /pnfs/grid.sara.nl/data/escape/disk/LOFAR/557202/L557202_SB001_uv.MS_37db2ad0.tar

61727293440 /pnfs/grid.sara.nl/data/escape/disk/LOFAR/557202/L557202_SB002_uv.MS_d7283945.tar

- ❑ **ESCAPE wiki (documentation)**

https://wiki.escape2020.de/index.php/Use_cases:LOFAR

https://wiki.escape2020.de/index.php/WP2_-_DIOS#Use_cases

- ❑ **Small subgroup - to accelerate work progress**

Prefactor - Next Steps/Open Questions

- ☐ Update wiki documentation
- ☐ Easier pipeline configuration settings
- ☐ Recipes on wiki/singularity hub
- ☐ Execute pipeline on LOFAR data on data lake
 - POSIX data available in a mounted directory (?)
 - Software location? Data lake ?
 - Compute (Output is 3xoriginal data) ?
 - tar/untar format?

Need Inputs for pipeline execution

- ☐ Performance tuning & research (network, Xcache, QoS)
- ☐ Pipeline in Cluster Configuration/mode (time)
- ☐ Nomenclature/Naming Scheme

Questions/Discussion



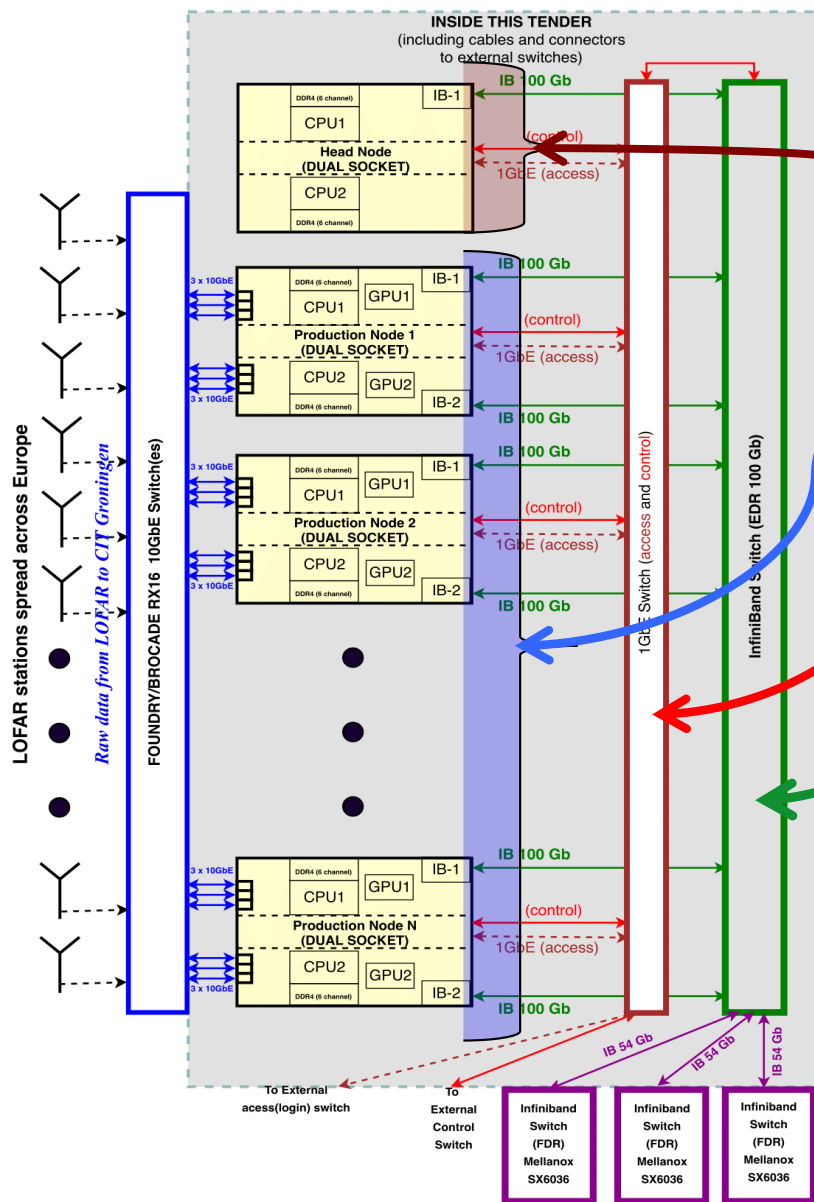


Figure 1: An overview of the COBALT 2.0 cluster layout. The networks are depicted with distinct colors. The raw data network (10 GbE) is shown in **Blue**, while the batch data network (100 Gbps EDR IB) is shown in **Green**. The external connections to the already existing IB FDR switches (54 Gbps) are shown in **Dark Magenta**. The user access (login) is shown in **Brown (dashed)** while the control (management) is shown in **Red** color.

Tender – BOSSERS & CNOSSEN (Dell Hardware)

- 1 - Head Node
Dell PowerEdge R440 Server
- 13 - Production Nodes
Dell PowerEdge R740 Server
- 1 - 48 port 1GbE switch
Dell S3048-ON
- 1 - 36 port 100Gb IB switch
Mellanox SB7800 EDR IB
- all required cables – included
- acceptance tests - required

❖ Hardware Delivered (Jan 2019)

❖ Installation & Acceptance (Feb 2019)

➤ Operating System Tuning & Software customization/tuning – Done

➤ Commissioning June 2019

LOFAR data – Zoom IN

