FP7- Grant Agreement no. 283393 – *RadioNet3*

Project name: Advanced Radio Astronomy in Europe

Funding scheme: Combination of CP & CSA

Start date: 01 January 2012 Duration: 48 month

Deliverable D10.11

Scientific publication with the results from the FTI demonstrator application on the selected hardware.

Due date of deliverable: 2015-02-28

Actual submission date: 2015-11-20

Deliverable Leading Partner: STICHTING ASTRONOMISCH ONDERZOEK IN NEDERLAND (ASTRON), THE NETHRELANDS

An European project supported within the 7th framework programme

1 Document information

Type	Demonstrator - Report
Title	Scientific publication with the results from the FTI demonstrator application on the selected hardware.
WP	10 (Hilado)
Authors	Ger van Diepen (ASTRON)

1.1 Dissemination Level

Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

1.2 Content

2 Fast Transient Pipeline

2.1 Introduction

The goal of the Fast Transient Imaging Pipeline (FTI) is to make in near real-time every N seconds an image of the sky for a few frequency ranges. It is outside the scope of FTI to find transient events in such images, but the data should be sufficiently flagged and calibrated to make transient detection possible.
The images produced by the FTI are used in the TraP pipeline. It searches for transients and variables in the images and stores their parameters in a light curve database for further analysis.

An article describing the results of the Fast Transient Imager will been published in the proceedings of the Astronomical Data Analysis Software and Systems conference XXV, October 2015, published in the ASP Conference Series of the Astronomical Society of the Pacific. A poster was presented at the ADASS XXV conference on October 25-29, 2015 as well.

2.2 Summary

The LOFAR radio telescope can operate in various modes, one of them being synthesis imaging. The data are stored in MeasurementSets for later processing, but can also be used for detection of slow transients by making an image per time sample. The paper describes the flagging, calibration and imaging steps to form such images, that can be fed into another pipeline (TraP) for transient detection. More accurate a posteriori processing is also possible using the stored MeasurementSets as the data source.

Some more work is required to make it a production system:

- Extra hardware has to be installed, because the current cluster has no spare capacity.
- The average time to process the observed data is slightly too high, mainly because the calibration takes too long. It has to be investigated if the calibration can be further optimized. Another possibility is to investigate if the calibration solutions are sufficiently stable to do the calibration every few seconds only.
- Because multiple sub bands are combined, the imaging time as such is not an issue. However, it introduces a latency of a few seconds which might have drawbacks for a possible future mechanism to generate VOEvents. It has to be investigated if the imaging can be sped up.

Copyright

© Copyright 2015 RadioNet3
This document has been produced within the scope of the RadioNet3 Projects.
The utilization and release of this document is subject to the conditions of the contract within the 7th Framework Programme, contract no, 283393

LOFAR Transient Imaging Pipeline

Ger van Diepen,[1]

[1]*ASTRON, Dwingeloo, Netherlands;* `diepen@astron.nl`

Abstract. The LOFAR radio telescope can operate in various modes, one of them being synthesis imaging. The data are stored in MeasurementSets for later processing, but can also be used for detection of slow transients by making an image per time sample. This paper describes the flagging, calibration and imaging steps to form such images, that can be fed into another pipeline (TraP) for transient detection. More accurate a posteriori processing is also possible using the stored MeasurementSets as the data source.

1. Introduction

LOFAR (van Haarlem et al. 2013) is an innovative radio telescope based in the Netherlands. It consists of stations in the Netherlands and some stations in various European countries. The Dutch stations are divided into core stations with a maximum baseline length of 3.6 km and remote stations with a maximum baseline length of about 110 km. LOFAR can observe in the low frequency band (LBA, 10-80 MHz) and high frequency band (HBA, 110-240 MHz).

The signals received by the dipoles or tiles in a station are beamformed at the station and sent to a GPU-based software correlator in Groningen. The dump time of the correlator is 1 sec resulting in a typical data stream of 1 GByte/s (2000+ baselines, 4 correlations, and 244 subbands of 64 channels), but can produce more data for other configurations. The output of the correlator is sent to a large cluster (80+ nodes) where the data of each subband are stored in a separate Casacore MeasurementSet (van Diepen 2015) to make parallel processing easily possible. Once all data of an observation are stored, pipelines are scheduled to do the standard data calibration and imaging.

Next to the standard data processing, the data can also be used in 'piggyback' mode for transient detection purposes. AARTFAAC (Prasad & Wijnholds 2012) is an all-sky radio monitor using the signals of the individual dipoles and tiles in the LOFAR core stations. The Transient Imaging Pipeline described in this paper, is another such application for the detection of slow transients. The pipeline can be run while an observation is done to form an image of some subbands for each correlator dump. The resulting images can be sent to the TraP system (Swinbank et al. 2015) to detect transients and lightcurves.

If needed, the data stored in the MeasurementSets, can be reprocessed to achieve more accurate results or to process the subbands that could not be handled at observation time.

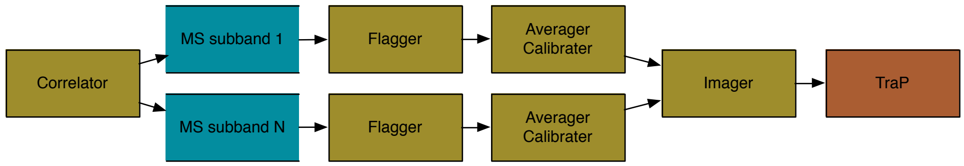


Figure 1.: Overview of FTI pipeline

2. Transient imaging pipeline

The pipeline consists of a number of processing steps as shown in Figure 1.

It is based on the LOFAR DPPP framework [1]. The visibility data of correlator dump are sent through a pipeline in which a series of processing steps are executed. Each subband has its own pipeline for flagging and calibration, but in the imaging step subbands are combined. The cadence of the pipeline (in principle one set of images per second) and the number of subbands to combine are defined by pipeline parameters.

2.1. Reading the MeasurementSets

The first step in the pipeline is reading the data from the MeasurementSet while it is being filled. The Casacore Table Data System (van Diepen 2015) used by the MeasurementSets, makes it possible to do this without any file locking. The advantage of this approach is that i) the real-time correlator software does not need to be changed and ii) the MeasurementSet acts as a buffer in case the processing of data takes longer than expected. It should be noted that reading the data does not impose extra IO operations because normally the OS will keep the data in its file buffers.

The baselines with European stations are filtered out as well as the baselines above a given length. Not only does it give a wider field of view, but it also reduces the amount of processing needed.

2.2. Flagging

The RFI environment at the LOFAR stations is such that data have to be flagged for RFI. The regular LOFAR data processing uses the AOFlagger (Offringa 2012) to detect radio interference and flag the affected data. It operates per baseline per subband on an as large as possible time window to achieve the best results. However, a large time window cannot be used in the transient imaging pipeline, otherwise extra latency would be introduced. For this reason a special version of AOFlagger has been developed using a sliding window over the data. The data are inspected and flagged using the data of some previous time slots, giving good flagging results. The flagging parameters are set such that transient data are not flagged.

2.3. Calibration and averaging

Heald et al. (2011) describe that the visibility function of LOFAR is strongly influenced by strong sources like CasA outside the field of view. They show that the demixing method (van der Tol et al. 2007) can calibrate for this problem. It solves for the gains

[1] http://www.lofar.org/operations/doku.php?id=public:user_software:ndppp

in the direction of those sources and subtracts them. The maximum number of solve iterations is limited to ensure real-time behaviour. For sufficient accuracy it has to be done at high resolution, whereafter the data can be averaged.

Calibration for the overall gain can be done with the *gaincal* step of DPPP, which uses the fast StefCal method (Salvini & Wijnholds 2014) to do the non-linear solve. It uses a source model created from the sources in the field of view found in LOFAR's Global Sky Model, which contains sources from catalogues like VLSS, NVSS, and WENSS. Using the *subtract* step of DPPP it is possible to subtract the sources from the visibility data.

2.4. Imaging

During the imaging stage multiple subbands can be combined to obtain better uv-coverage and a higher signal-to-noise ratio. By default one clean cycle is done.
A few imagers have been compared. All of them use W-projection and are highly parallellised.
- *wsclean* (Offringa 2014); a fast wide-field imager.
- The CASA imager; a wide-field, wide-band imager. It can correct for direction-dependent effects varying with parallactic angle.
- The LOFAR *awimager*; a wide-field imager capable of applying direction dependent corrections like beam and ionosphere that vary with time, frequency, and baseline.

Imagers often have substantial initialisation overhead when used on very short observations. The tests showed *wsclean* is much faster than the other imagers, while giving comparable image quality.

3. Results

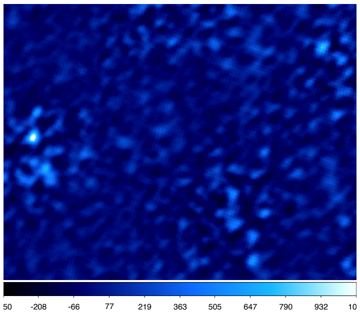
Pipeline runs have been done on the LOFAR CEP3 cluster using a node with four 2.2 GHz Intel Xeon processors, each having ten cores. Ten simultaneous pipelines were run, each using four cores. Table 1 shows the average timings of 10 simultaneous flagger and calibration pipelines and imager steps combining 10 subbands into a single image. The two columns show the results for runs on simulated data (containing transients) and on observed data.

Step	Simulated data	Observed data
Read + filter	0.04	0.04
Flagger	0.15	0.14
Calibrater + averager	0.21	0.72
Imager	3.09	3.11

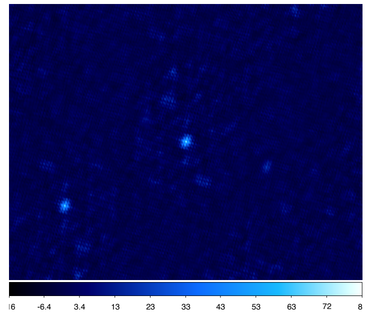
Table 1.: Pipeline timings on simulated and observed data

Note that the imager takes about 3 seconds, but since it combines 10 subbands, the average imaging time per subband is 0.3 seconds. However, it does introduce latency. The resulting images are augmented with meta data for correct handling by the TraP system, which has found the simulated transients.

Demixing. The importance of the demixing step is shown below. Images of 4x4 degrees are made from a simulated visibility data set containing 2 sources. The much stronger CygA source is about 8.5 degrees from the field center and contaminates the data. Subtracting CygA using the demixing method is essential.



(a) Without demixing



(b) With demixing

3.1. Conclusions and Future work

The Transient Imaging Pipeline make it possible to form images of groups of subbands for each second of LOFAR data. Some more work is required to make it a production system:

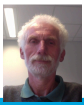
- Extra hardware has to be installed, because the current cluster has no spare capacity.
- The average time to process the observed data is slightly too high, mainly because the calibration takes too long. It has to be investigated if the calibration can be further optimized. Another possibility is to investigate if the calibration solutions are sufficiently stable to do the calibration every few seconds only.
- Because multiple subbands are combined, the imaging time as such is not an issue. However, it introduces a latency of a few seconds which might have drawbacks for a possible future mechanism to generate VOEvents. It has to be investigated if the imaging can be sped up.

Acknowledgments. The work has received funding from the European Commission Seventh Framework Programme (FP/2007-2013) under grant agreement No 283393 (RadioNet3).

References

Heald, G., et al. 2011, Astronomy and Astrophysics, 32
Offringa, A. R. 2012, Ph.D. thesis, University of Groningen
— 2014, MNRAS, 444, 606
Prasad, P., & Wijnholds, S. J. 2012, in Proceedings of New Windows on Transients across the Universe: A discussion meeting of the Royal Society (London)
Salvini, S., & Wijnholds, S. J. 2014, Astronomy and Astrophysics, 571, A97
Swinbank, J. D., et al. 2015, Astronomy and Computing, 11, 25
van der Tol, S., Jeffs, B. D., & van der Veen, A. J. 2007, IEEE Transactions on Signal Processing, 55, 4497
van Diepen, G. N. J. 2015, Astronomy and Computing, 12, 174
van Haarlem, M. P., et al. 2013, Astronomy and Astrophysics, 556, A2

LOFAR Transient Imaging Pipeline Framework



Ger van Diepen (diepen@astron.nl)

ASTRON

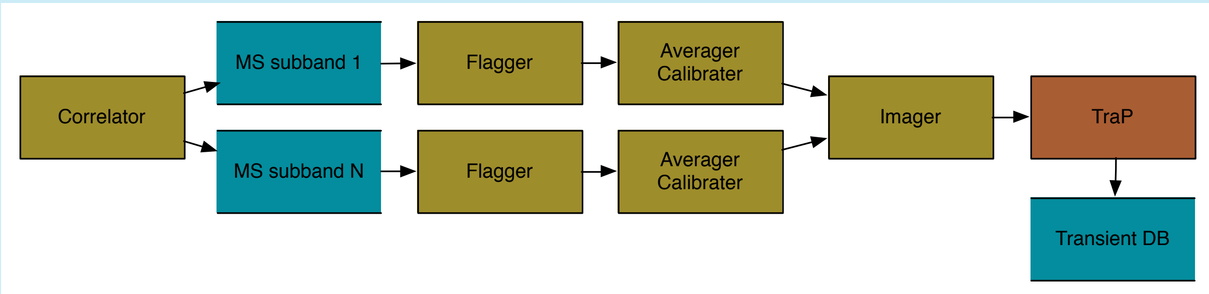
www.astron.nl

1. Introduction

The correlator used by the LOFAR telescope dumps visibility data at a rate of 1 sec resulting in a typical data rate of 1-4 GByte/s. Next to using the visibility data in the standard data processing pipelines, these data can also be used in 'piggyback' mode for transient detection purposes. AARTFAAC (Prasad, 2012) is one such application resulting in an all-sky radio monitor using the signals of the individual dipoles and tiles in the LOFAR core stations.

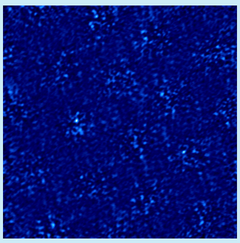
The Transient Imaging Pipeline described in this poster, is another such application for the detection of slow transients. The pipeline can be run during an observation to form an image of a few subbands for each time stamp. The resulting images can be sent to the TraP system (Swinbank, 2015) to detect transients and lightcurves. If needed, the data stored in the Casacore MeasurementSets, can be reprocessed to achieve more accurate results or to process the subbands that could not be handled at observation time.

2. Pipeline overview

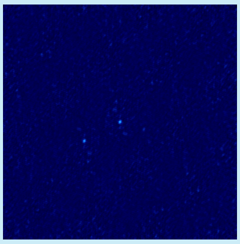


The flagger step flags for RFI using Andre Offringa's AOFlagger. Typically 5-8% of LOFAR data have to be flagged.

The calibrater contains the so-called demixing step. The signal from strong sources like CasA enter through the sidelobes of the beam and have to be subtracted at high resolution. The difference can be seen in the images below. Only the right one shows the two sources.



without demixing



with demixing

The imager can combine multiple subbands to get a better SNR. Typically, only the shorter baselines (<10 km) are used to get a wider field of view.

The final images are sent to the TraP sytem (University of Amsterdam) to extract transient and lightcurve info from the images.

The pipeline is based on LOFAR's DPPP pipeline concept. The steps in the pipeline are given in a so-called parset text file. An extract is given below.

```
demixer.demixfreqstep=64
demixer.demixtimestep=1
demixer.freqstep=64
demixer.skymodel=ateam.sdb
demixer.subtractsources=[CasA,CygA]
```

3. Results

A simulated data set and a real data set have been processed with the pipeline on a 40-core machine using 10 cores. The table shows the times to process a single time sample for a single subband. The imager performs one clean cycle.

	simulated	real
Read + filter	0.04	0.04
Flagger	0.15	0.14
Calibrater + averager	0.21	0.72
Imager	2.58	2.59

The imager takes more than one second, but because multiple subbands are combined, there is sufficient processing capacity. The only problem left is the latency of > 1 sec.

ASTRON Netherlands Institute for Radio Astronomy

RadioNet

ASTRON is part of the Netherlands Organisation for Scientific Research

NWO Netherlands Organisation for Scientific Research