SF2A 2022
J. Richard, A. Siebert, E. Lagadec, N. Lagarde, O. Venot, J. Malzac, J.-B. Marquette, M. N'Diaye, D. Briot (eds)

NENUFAR DATA CENTRE AND EARLY SCIENCE RESULTS

B. Cecconi[1, 2], A. Loh[1,2], J. N. Girard[1], E. Tremou[3], B. Lamiré[4], P. Hamy[4], P. Murphy[1], C. Dumez-Viou[2] and the NenuFAR builders[5]

Abstract. NenuFAR is a world-class low frequency (10-85 MHz) radio telescope built in Nançay, France. It is an official SKA pathfinder, and will remain complementary to SKA in the lowest frequencies. A several petabytes-scale data centre (NenuFAR-DC) is prepared to host and share the NenuFAR data, based on a cloud architecture (distributed storage and computing).

Keywords: subject, verb, noun, apostrophe

1 Introduction

NenuFAR (New Extension in Nançay Upgrading LOFAR) Zarka et al. (2015) is a giant low-frequency station of the LOFAR array van Haarlem et al. (2013). It is an SKA (Square Kilometre Array) pathfinder focusing on the lower end of the radio spectrum (10-85 MHz). It can produce more than 10 TB of raw data per hour, hence requiring advanced data reduction and data management solutions. In this paper, we present an overview of the current early science activities, as well as the planned design of the NenuFAR data management components.

2 Current building state and performance of NenuFAR as of Fall 2022

2.1 Current status

NenuFAR is a SKA-LOW pathfinder working at low frequencies between ∼10 and 88 MHz, around the current FR606 station of the International LOFAR Telescope. NenuFAR was designed originally as a dense core of 400 m in diameter made from the optimal grouping of 96 small analogue phased arrays of 19 antennas (hereby coined "mini-array"), each acting as a sensitive equivalent 25-m aperture dish. NenuFAR can act as: i) a single pixel telescope, when signals from all mini-arrays are electronically and coherently summed (a.k.a. beamformed) toward a direction of interest, ii) as an interferometer when each pair of mini-array signals are cross-correlated together to form visibility measurements, and iii) a sensitive LOFAR station as part of the International LOFAR Station (a.k.a., "LOFAR Super Station" – LSS mode).

The design of NenuFAR was optimized from top to bottom. The mini-array elements are LWA-like linear crossed-dipoles with optimized Low Noise Amplifiers (LNA) in the 10-90 MHz Dallier et al. (2017). The distribution of elements within a mini-array are arranged in a hexagon to facilitate the analogue phasing while keeping the mini-array beam patterns optimized for the field of view. The distribution of mini-arrays, at the whole core level, are pseudo randomly distributed to give a gaussian-shaped snapshot (u,v) coverage, thus improving the image quality even if the integration time is short.

Complementary to the core, 6 "remote" mini-arrays are planned to provide between 800m to 2 km baselines from the core. As of summer 2022, 77 mini-arrays are built in the core and 3 remote mini-array are left to build or commission.

[1] LESIA, Observatoire de Paris, Université PSL, Sorbonne Université, Université Paris Cité, CNRS, 92190 Meudon, France
[2] USN, Observatoire de Paris, Université PSL, Univ Orléans, CNRS, 18330 Nançay, France
[3] NRAO, Socorro, NM, USA
[4] DIO, Observatoire de Paris, Université PSL, CNRS, 75014 Paris, France
[5] L. Bondonneau, C. Briand, F. Casoli, D. Charrier, A. Coffre, I. Cognard, S. Corbel, R. Dallier, L. Denis, C. Ferrari, J.-M. Grießmeier, G. Hellbourg, A. Konovalenko, L. Koopmans, L. Martin, F. Mertens, B. Semelin, C. Taffoureau, M. Tagger, C. Tasse, G. Theureau, P. Tokarsky, O. Ulyanov, R. Vermeulen, V. Zakharenko, & P. Zarka

The current performances of NenuFAR in terms of collecting area 1 and angular resolution 2 can be retrieve from the NenuFAR website `https://nenufar.obs-nancay.fr/en/astronomer/`.

Collecting area (m^2)	**Frequency**	**Current**	**Complete**
Core only	15 MHz	69000 m^2	83000 m^2
	50 MHz	18000 m^2	22000 m^2
	85 MHz	6300 m^2	7500m^2
Core + remote	15 MHz	72000 m^2	88000 m^2
	50 MHz	19000 m^2	23000 m^2
	85 MHz	6600 m^2	800m^2

Table 1. Theoretical Effective area

Angular resolution (°)	**Frequency**	**Current**	**Complete**
Core only	15 MHz	2.39°	2.11°
	50 MHz	0.72 °	0.63 °
	85 MHz	25.2 °	22.2°
Core + remote	15 MHz	21.5’	19.3’
	50 MHz	6.5’	5.8’
	85 MHz	3.8’	3.4’

Table 2. Theoretical angular resolution

In Super-station mode, NenuFar is connected to the LOFAR FR606 backend, so that the recorded signal can be sent to the LOFAR central correlator. Each mini-array then acts as a Low-band LOFAR antenna. In Stand-alone mode, the two main modes are the *Beamformer* (e.g., for the dedicated `nenufar-pulsar` pipeline Bondonneau et al. (2021), or the `nenufar-tf` mode) and the *Imager* (for the `nenufar-nri` mode). The full list of NenuFAR acquisition pipelines is provided in Table 3.

`nenufar-bst`	Beam Statistics (low resolution)
`nenufar-pulsar`	Pulsar pipeline
`nenufar-nri`	Radio Imager
`nenufar-tbb`	Transient Buffer Board
`nenufar-tf`	High Resolution Spectro-Temporal mode
`nenufar-sst`	Extended Statistics
`nenufar-xstms`	Correlator Mode

Table 3. NenuFAR acquisition pipelines

2.2 NenuFAR studies during the Early Science phase

NenuFAR started its commissioning phase in the scope of several “Early Science” projects carried by the french and international low frequency radio community. From Cosmic Dawn studies(Mertens et al. 2021) to a pulsar survey (Bondonneau et al. 2021), NenuFAR covers a wide range of science case in the low frequency window that remained fairly unexplored with enough sensitivity (Zarka et al. 2015). One can also mention plasma physics in our solar system (Briand et al. 2022) or at selected exoplanet candidates associated with Star-Planet interactions potentially emitting in the NenuFAR window (Turner et al. 2022). Early Science project also come with pilot studies on dedicated targets, such as the prospective study on cosmic filaments and magnetism in the Coma cluster (Bonnassieux et al. 2021), or the monitoring of the solar activity Carley et al. (2021). A transient working group is currently at work to help building the default pre-processing pipeline of NenuFAR imaging data (from L0 to L1) and post-processing pipeline (Calibration and imaging) in order to plug in a dedicated transient detection pipeline.

3 NenuFAR Data Management

During the NenuFAR commissioning, the ES teams have developed data pipelines to process the raw NenuFAR data (level 0, L0) into reduced data products (level 1, L1). The L0 data products (up to ∼14 TB/hr) are stored temporarily on a 3 PB storage facility at CDN (Centre de Données de Nançay). The L0 to L1 data processing is run on a set of computing nodes (called *nancep1*, *nancep2*, etc, nodes), also managed by CDN. Each data product is identified by four main metadata, as listed in Table 4.

Keyword	**Description**	**Example**
`team-id`	Team ID	`ES11`
`instr-id`	Acquistion pipeline	`nenufar-tf`
`obs-id`	Observation ID	`20220130_112900_20220130_123100_SUN_TRACKING`
`proc-level`	Processing Level	`L0`

Table 4. Main metadata keywords for NenuFAR data product identification.

In general, the L0 data are not kept after the L1 data is produced, so they are erased when the L1 data are produced. However, the L0 observation directory is kept. During the L0 cleaing process, the actual data files are deleted, but the configuration, pointing, or instrument log files are kept for the record. The trace of the deleted file is kept in a file (*epitaph.log*), including the reason of the data removal (e.g., *reduced data*, *no science signal*, or *corrupted observation*.

The final L0 and L1 data products are then meant to be transferred to NenuFAR-DC (NenuFAR Data Centre), which is currently being designed and will be hosted in the Région Centre-Val-de-Loire datacenter (at BRGM in Orléans).

3.1 NenuFAR Data Center and its Datalake

NenuFAR-DC is designed with a *cloud* architecture: scalable storage (Object storage, based on CEPH), cloud computing infrastructure (based on Openstack), science platform interface (based on JupyterLab), etc. We also plan to use or implement an AAI (Authorzation and Authentication Infrastructure) in a second stage. The data transfer from CDN to NenuFAR-DC will be managed with IRODS or RUCIO, two distributed data management systems. The object storage index of NenuFAR-DC will be the main metadata index for NenuFAR collections. It is fed when data are pushed to NenuFAR-DC. It will be used by the user and interoperable interfaces to access the data.

Virtual observatory interfaces (from astronomy, planetary and heliophysic) will be implemented. Datasets will be published openly after an embargo period (six months to three years), and citable with a DOI. Users will be encouraged to feed higher level data products back into NenuFAR-DC. Each team will have to prepare a data management plan explaining how the data is processed, what higher level data is available, what software is used, etc., with the goal of following the FAIR principles (Findable, Accessible, Interoperable and Reusable).

NenuFAR-DC will provide preconfigured and maintained software containers (based on Singularity), providing the user community to up-to-date and adapted software suites for NenuFAR data processing. Deployment and operation tests have conducted with the computing facilities at Observatoire de Paris, as well as with Jean-Zay at IDRIS. After the tests, the Singularity containers are considered to be a satisfactory option for deploying containers on heterogeneous computing infrastructure. Note the in this framework, we favor portability to optimization.

The NenuFAR-DC local computing capabilities will be limited due to cost constraints. A Datalake infrastructure is thus envisaged to provide additional capabilities to the NenuFAR user community. Several mesoscale data centers (e.g., MesoPSL in Meudon) agreed to participate to the NenuFAR datalake, as well as other facilities like, e.g., cold storage backup hosted on EOSC (at MPCDF, Germany). Figure 1 shows the overall planned architecture for NenuFAR-DC and its datalake.

4 Conclusions

NenuFAR is a very promising instrument. While still in commissioning phase, NenuFAR already provided many science results that have been published in peer-reviewed journals. Several community workshops have been organised to train the community.

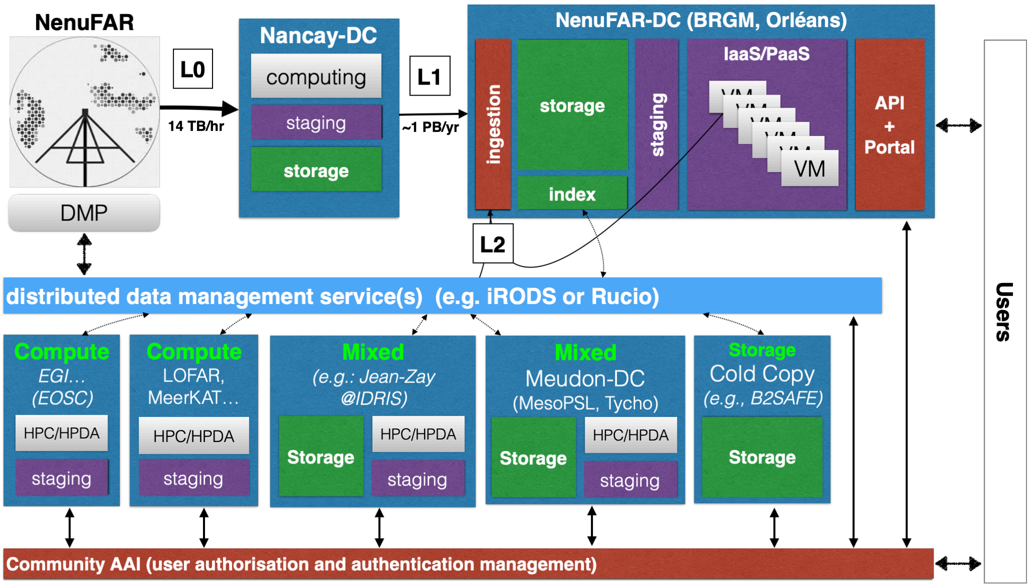

Fig. 1. Planned architecture for NenuFAR-DC and its datalake

On the data management side, NenuFAR-DC can also be considered as a pathfinder for the future SKA regional centres. Even though the NenuFAR data products sizing doesn't scale with the planned SKA ones, its overall distributed and heterogeneous architecture is illustrating the challenges, which will have to faced and solved for the SKA regional centres.

The development of NenuFAR and its data centre is supported by personnel and funding from: Observatoire Radioastronomique de Nançay, CNRS-INSU, Observatoire de Paris-PSL, Université d'Orléans, Observatoire des Sciences de l'Univers en région Centre, Région Centre-Val de Loire, DIM-ACAV and DIM-ACAV+ of Région Ile-de-France, Agence Nationale de la Recherche. NenuFAR-DC is hosted by Bureau des Recherches Géologiques et Minières in Orléans. NenuFAR is also supported by the EOSC-DICE project, which received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no 101017207.

References

Bondonneau, L., Grießmeier, J. M., Theureau, G., et al. 2021, A&A, 652, A34

Bonnassieux, E., Tremou, E., Girard, J. N., et al. 2021, Galaxies, 9, 105

Briand, C., Carley, E., Cecconi, B., Reid, H., & Zarka, P. 2022, in EGU General Assembly Conference Abstracts, EGU General Assembly Conference Abstracts, EGU22–2920

Carley, E. P., Cecconi, B., Reid, H. A., et al. 2021, ApJ, 921, 3

Dallier, R., Carduner, H., Charrier, D., et al. 2017, in 35th International Cosmic Ray Conference, Vol. ICRC2017, Busan, South Korea, 417

Mertens, F. G., Semelin, B., & Koopmans, L. V. E. 2021, in SF2A-2021: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, ed. A. Siebert, K. Baillié, E. Lagadec, N. Lagarde, J. Malzac, J. B. Marquette, M. N'Diaye, J. Richard, & O. Venot, 211–214

Turner, J. D., Zarka, P., & Grießmeier, J.-M. 2022, in Bulletin of the American Astronomical Society, Vol. 54, 404.02

van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, A&A, 556, A2

Zarka, P., Tagger, M., Denis, L., et al. 2015, in 2015 International Conference on Antenna Theory and Techniques (ICATT), 1–6