ASTRON
Netherlands Institute for Radio Astronomy

LOFAR LTA Site Requirements

	Organisatie / Organization	Datum / Date
Auteur(s) / Author(s): John D. Swinbank Hanno Holties Roberto Pizzo	 ASTRON ASTRON ASTRON	 2025-07-29
Controle / Checked: Michiel P. van Haarlem	 LOFAR ERIC	 2025-05-08
Goedkeuring / Approval: ...	 ...	 ...
Autorisatie / Authorization: ... **Handtekening / Signature:**	...	...

This technical note expresses the opinions of the authors on the date of writing.
It is not binding on the SDC Program, and may not accurately reflect future developments.

Document History

Revision	Date	Description
2.0	2025-05-09	Incorporate improvements from Michiel van Haarlem
1.0	2025-05-08	First version

List of Abbreviations

API Application Programming Interface.

CPU Central Processing Unit.

ERIC European Research Infrastructure Consortium.

GPU Graphics Processing Unit.

L2LP LOFAR2.0 Large Programme.

LOFAR LOw Frequency ARray.

LTA Long Term Archive.

SRCNet SKA Regional Centre Network.

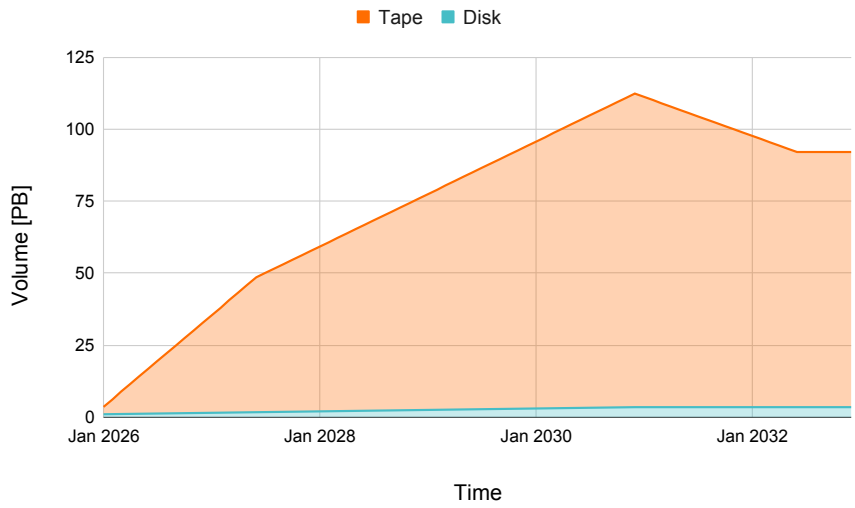

Figure 1: Total storage required for LOFAR2.0 through 2032 based on a five-year observing programme. Beyond this point, the amount of data stored remains constant indefinitely in this model; this would clearly not be the case if observations are continued.

1 Introduction

This document outlines the requirements for LOFAR LTA sites in the era of LOFAR2.0. It is intended to act as a guideline and high-level reference to set expectations during the early stages of establishing an LTA site; full requirements and implementation details should be further established on a site-by-site basis.

2 Total Capacity

The resource requirements described in this section represent current best estimates based on current LOFAR2.0 Large Programme (L2LP) proposals and an appropriate extrapolation to future open-skies calls. They are expected to continue to evolve over the course of preparing for and executing the LOFAR2.0 observing programme.

2.1 Storage

As shown in Fig. 1, total amount of storage required to support a five-year LOFAR2.0 observing programme across the whole LTA system increases to around 110 PB in 2031, ultimately dropping back to about 90 PB for long-term preservation. Some of this storage—increasing gradually to 3.5 PB across the whole archive—must be "fast" (disk); that is, users should immediately be able to access it. The remainder can be "slow" (tape), from which data is staged on user request for later access. Each site must provide fast storage; the use of slow storage is an implementation decision for the site.

The total amount of storage provided by each site is flexible, as long as the total capacity across the network is adequate. A *minimum pledge of at least 1 PB* is required for the site to be regarded as viable.

2.2 Compute

Current estimates are that 1.5 billion (1.5×10^9) CPU hours, distributed equally over the five year operational lifecycle of the telescope, will be required to generated advanced data products from LOFAR2.0. A *minimum pledge of at least 10 million CPU hours* is required for the site to be regarded as viable.

In future additional data analysis services will be provided through LTA sites that may add to this total (see §4); on the other hand, pipeline efficiency is steadily increasing, which will drive the total down.

3 Technical Requirements

3.1 Storage

- Provide a dCache-managed storage environment.
- Provide file access using WebDAV.
- Provide file management using the dCache API.
- Support automatic migration of data to slow storage (if used).
- Provide sustained average speed between fast and slow storage (if used) of at least 4 Gbps.

3.2 Compute

- Provide a Linux-based compute cluster.
- Support Singularity/Apptainer based containers on both head and compute nodes.
- Support Python-based daemonic services on head nodes.
- The SLURM workload manager is preferred; use of others is subject to evaluation & test.
- Provide a shared (POSIX) filesystem, accessible from all nodes.
- Provide at least 30 x86-64 compute cores per node.
- Provide at least 8 GB RAM per core.
- Provide at least 200 GB fast local “scratch” storage per core.

3.3 Network

- Support a dedicated connection from LOFAR central processing supporting a data rate of at least 10 gigabit/second.
- Provide a connection of at least 10 gigabit/second between storage and processing systems located at the same site.
- Provide data access to other LTA sites and general users supporting a total of at least 10 gigabit/second transfer rate.

3.4 Support

- A commitment of resources for a period of at least five years is recommended.
- In the event of termination (early or scheduled) of site operations, the site must support migration of data to another site. This includes agreement of a migration plan with LOFAR ERIC and allowing at least one year for the transfer to take place.

- Service availability of 95% is expected, as is response to issues within one working day.

4 Future Evolution

The situation described in §§2 & 3 is current as of the date of writing. However, as with any complex system, we anticipate that the technological landscape will continue to evolve over the lifetime of the LTA. For example, we anticipate:

- An increasing use of GPUs and (perhaps) other accelerator technologies in core LOFAR data processing pipelines;
- An evolution of the LTA technology stack to converge with development being undertaken in the context of the SKA SRCNet;
- More user-driven interactive analysis functionality taking place within the archive;
- An increased emphasis on high-throughput, low-latency processing of data as it arrives at the archive.

Of course, this list is extremely unlikely to be complete.

We therefore request that LTA sites:

- Signal their willingness to evolve their technology stack in tandem with, and under the direction of, ASTRON and LOFAR ERIC.
- Assign an appropriate level of staffing to participate in working groups, technology forums, and similar activities to help guide and steer the future technical direction of the LTA.