LOFAR: first of a new generation of radio telescopes

Harvey R. Butcher
LOFAR: First of a New Generation of Radio Telescopes

H. R. Butcher
on behalf of the LOFAR development team at the
Netherlands Foundation for Research in Astronomy (ASTRON)
P.O. Box 2, 7990 AA Dwingeloo, Netherlands

ABSTRACT

Aperture plane detection of cosmic radio signals can confer several important advantages. The mechanical signal processing of conventional parabolic dish antennas can be replaced by signal processing purely in electronics and software, thereby placing aperture plane array instruments in a regime with decreasing future costs. Aperture arrays using all-sky element antennas also allow multiple, independent, full sensitivity telescopes to be formed in software. The Low Frequency Array (LOFAR) project aims to realize these advantages in a previously largely unexplored frequency domain, 10–240 MHz. Consisting of up to eight shared aperture telescopes operated independently over internet, it will explore the epoch of re-ionization, open the transient radio Universe to systematic study, record ultra-high energy cosmic particles and add a new dimension to solar and space physics research.

Keywords: Radio telescope, aperture array, software telescope, low frequency, cosmic rays

1. INTRODUCTION

The scientific drivers for a new generation of radio telescopes argue for roughly two orders of magnitude increased sensitivity, which can only be achieved by increasing the total flux collecting area, together with correspondingly increased angular resolution. The conventional design paradigm for radio telescopes employs large parabolic dish antennas, which unfortunately are too expensive to scale as required. To reduce total cost as well as to optimize flexibility, migration to designs consisting of large numbers of low mass antennas together with increased reliance on signal processing in electronics and software is unavoidable. At frequencies substantially above a GHz, telescope designs based on many hundreds of small, mass-produced, movable dish antennas appear to be a realistic approach. At a GHz and below, it is now feasible to employ very large numbers of simple, all-sky antennas. In the latter case, in analogy with the software radio concept, one may speak of a software telescope in which pointing, beam forming, frequency selection etc are realized in embedded software. Indeed, many independent software telescopes may be formed from the same set of input signals. LOFAR is a first demonstration of such a telescope in an operational environment.

The LOFAR radio telescope is an aperture synthesis array consisting of a large number of low-cost, broadband, all-sky antennas whose signals are connected by data links and hierarchical digital signal processing chains to a powerful central processor. This architecture is generic to a large class of wide area sensor networks and allows not only great flexibility of implementation but also provides prospects for including sensors other than RF antennas. Indeed, the current LOFAR project embraces in addition to RF antennas, arrays of seismic geophones, weather stations, microphones and agricultural sensors. Only the radio telescope implementation will be discussed here.

Central to the telescope project is the provision of operational access on-line to geographically distributed user groups. The project thereby explores extension of the GRID concept to include the coupling of on-line operations of a major research facility with distributed computational facilities and distributed databases.

2. SCIENCE GOALS

The five key areas of astrophysical science driving the design of the LOFAR radio telescope are the following.

- High Redshift Universe. Deep surveys at radio frequencies lower than conventional radio telescopes record promise to reveal populations of objects in the early Universe that are luminous enough to permit detailed study. At later cosmological epochs one expects to encounter important populations of fossil radio sources.
- **Epoch of Re-ionization.** The first luminous objects that formed in the early Universe must have re-ionized the surrounding neutral medium. This universal phase change marks the last completely unexplored cosmological epoch. The detection of its global signature and the mapping of its structures are thought best to be done at radio frequencies in the LOFAR band. The telescope is therefore being designed to attempt its detection.

- **Particle astrophysics.** The origin and distribution of cosmic rays in galaxies and the physical mechanisms responsible for their acceleration to exceedingly high energies are important astrophysical questions to be addressed by LOFAR. Low and medium energy cosmic ray electrons can be mapped for the first time in 3D in our Galaxy, while ultra-high energy particles impinging on the Earth can be detected and studied through the Askaryan effect (producing Giga-Jansky radio pulses that are readily detected by LOFAR antennas).

- **Bursting and Transient Universe.** The LOFAR signal processing chain is conceived to enable surveys and long term study of variable and transient phenomena – bursts from Jupiter-like planets, merging and interacting compact objects, pulsars and so on.

- **Solar-Terrestrial Relationships.** LOFAR will generate excellent maps of the solar wind using interplanetary scintillation of background radio sources, and of coronal mass ejections. It promises to provide essential information on magnetic fields frozen in the plasma, which is of prime importance to the strengths of the geomagnetic storms they can cause. Additional research in this category includes ionospheric tomography and the physics of the outer solar atmosphere.

### 3. SYSTEM DESIGN

The LOFAR system design aims to realize an aperture synthesis radio telescope based on a very large number of low-cost, stationary, all-sky antennas ~ 15000 initially, to be increased in future as science dictates and funding allows. The design then relies on a distributed, hierarchical, digital signal processing chain to implement the majority of its functionality in software. The resulting instrument has extreme agility in time, in pointing direction and in observing frequency, as well as multiple independently pointable beams and several simultaneous modes of operation. In this section is described the signal processing chain that achieves these functions.

![Figure 1](https://www.spiedigitallibrary.org/conference-proceedings-of-spie)  
*Figure 1. Prototype element antennas deployed at the initial LOFAR test site in Drenthe (Nov. 2003). Each element antenna here is a dual polarization, crossed dipole installed in a pyramidal mechanical construction made of inexpensive PVC pipe having the first stage, analog processing at the apex. This cluster of 100 elements is connected to receiver units and a small cluster computer in the container for design testing.*

The element antennas are deployed in clusters (cf. Figure 1), which we call stations, that in turn are laid out in a centrally concentrated, five arm log-spiral configuration extending over an area initially ~ 100 km in diameter. The center of the configuration is located in a thinly populated agricultural area in the Province of Drenthe in the northeastern Netherlands. Stations are located along the arms such that the number of stations is constant in consecutive elliptical distance shells incrementing by a factor of 1.7 in radius from the array center. Good Fourier plane (u,v) coverage results for full hour angle tracks and a minimum allowed observing elevation of 30 degrees. The FWHM beam at 150 MHz is about 6 seconds of arc.

The configuration is oriented such that future extensions toward the city of Bremen to the east and Bonn to the south might be realized. Even longer baselines making use of broadband data links to the observatories of the European VLBI Network are feasible in principle.
Two types of antennas are planned, each receiving signal from the whole sky and having a beam pattern exhibiting a null at the horizon. The first is a dual polarization dipole optimized for operation in the frequency range 30-80 MHz. The design retains sensitivity also to lower frequencies, which will be made available for observing at night during periods when interference is low (~ 80% of the time). It features in addition a sharp notch filter above 80 MHz to eliminate commercial FM-radio band signals. The total effective collecting area of these antennas is ~ 400000 m² at 20 MHz, and the sky-limited system temperature ~ 45000 K. The second type of antenna is a compound element consisting of 4x4 dual polarization dipoles optimized for the 120-240 MHz band, together with a single beam RF beam-former per polarization. At 150 MHz the total collecting area is ~ 130000 m² and the system temperature ~ 250 K.

The subsequent digital signal processing design relies on the expected technological evolution of the commercial IT industry at the projected moment of procurement. This approach allows substantial savings in time and development cost while introducing a certain level of risk. Specifically, the development of the systems for data transport, station level digital processing, central digital processing, and monitoring and control are occurring in close cooperation with industry.

This signal processing chain consists of five major subsystems: Remote Stations, Compact Core, Wide Area Network, Central Processor, and Monitoring & Control. Here we treat each of these in turn.

3.1. Remote Stations

Element antennas are clustered in 45 Remote Stations, each situated on 4 ha of ground and containing 100 elements. The signals are pre-filtered and amplified at the antenna, and transported in analog form over coaxial cable to a receiver unit located in a nearby, shielded hut. All digital hardware is contained in this hut to maintain EMC specs, to prevent common mode signals on the coax cables from the antennas, and to minimize the risk of damage by lightning.

The receiver unit selects for the antenna type, then filters and amplifies the signal, in the process converting the single ended signal into a differential signal prior to A/D conversion. For digitization, a direct-conversion architecture has been adopted that employs 12-bit sampling at rates of either 200MS/s or 160MS/s, which allows for optimization of bandwidth positioning in radio frequency. Data are then formed into complex numbers 16+16 bits wide.

Following A/D conversion, a poly-phase spectral filter bank splits the band from the antenna into 512 equidistant sub-bands, each 195 kHz or 156 kHz wide depending on the clock rate. Specific ones of these sub-bands may then be selected (e.g. to be free from unwanted signal) across the full 100 MHz or 80 MHz band, to yield a maximum total effective bandwidth of 32 MHz per polarization. This bandwidth is matched to the design capacity of the Central Processor.

To synthesize beams on the sky, the sub-band signals are combined in independent, pipelined, beam-forming processors. One fully steerable beam can be formed having the full 32 MHz bandwidth, eight independent beams at 4 MHz bandwidth each, or the combinations in between. The weights necessary for the beam-formers are calculated in a local control unit and are downloaded to the beam-formers. After beam-forming, the results have a bandwidth to be transported of about 2 Gb/s and are sent to the Central Processor via 10 GbE data links on the Wide Area Network. Additionally, statistical measurements are performed at the station and together with other monitoring information are forwarded to the Monitoring and Control system.

In parallel with filtering and beam-forming, raw all-sky data can also be stored in a cyclical RAM buffer at each Remote Station. Upon an internal or external trigger these data can be read out for processing, allowing the telescope to be pointed post-facto. Initially, these buffers will be able to store data for one second at the full sensitivity of the Station, or for longer periods at reduced sensitivity. Application will initially focus on particle event detection and analysis.

Timing signals at the Remote Stations are generated locally. A rubidium-disciplined oscillator maintains the short-term time base and is integrated with a global positioning system receiver for absolute timing and long-term stability. This approach also obviates the need for a centrally distributed clock for the LOFAR array.

3.2. Compact Core

Within an area of 320 ha (~ 2 km in diameter) are situated 3200 element antennas of each type (41% of the total). This area is elliptical having an axis ratio of 1.3 and is extended in the north-south direction. The signals from all the
3.3. Wide Area Network

A Wide Area Network (WAN) transports data from the Remote Stations and Compact Core antennas to the Central Processor. To minimize cost this subsystem will be based on 10 GbE digital optical fiber communications technologies, which are projected to become commercially available at the planned time of procurement.

The LOFAR WAN will consist of a mix of existing dark fiber and proprietary fiber laid by the LOFAR project. The primary backbone to the Central Processor will initially have a bandwidth approaching 1 Tb/s but will have sufficient capacity that it might be upgraded in future to allow the signals from all element antennas to be transported for central processing.

To minimize cost to the project and to promote good community relations the WAN implementation will be organized in cooperation (e.g. by sharing fiber ducts) with the roll-out of commercial broadband internet in the surrounding towns and countryside.

3.4. Central Processor

Final signal combination and generation of data products occurs in the Central Processor, which implements the conventional interferometric correlation function among all pairs of stations. In the LOFAR context it also does much more. For aperture synthesis imaging, LOFAR implements a distributed FX correlator, with the F-part largely implemented in filter-banks at the Remote Stations and the X-part implemented in a flexible Central Processor. But the Central Processor also performs all filtering and beam-forming functions for the Compact Core element antennas, as well as processing for the Tied Array, Transient Detection and UHECR Detection modes (cf section 5.1).

Several aspects of the LOFAR problem have influenced the choice of specific implementation of the Central Processor. The data stream into the Processor will initially be about 700 Gb/s and will require a total computing power of about 25 T-ops. Together with a need for great flexibility and substantial bit-depth, the Processor's architecture must combine the overall features of future general-purpose supercomputers with enhanced I/O capabilities.

The IBM company has proposed that these requirements might be accommodated in a specially built version of their BlueGene/L research computer. The /L is the first in a planned series of research processors that within a decade should yield a multi-petaFLOP/s supercomputer. Its architecture is ideal for the streaming processing of vast amounts of complex data, while its implementation is based on standard low-power components and on the LINUX operating system, facilitating both development and maintenance. The LOFAR version will employ 12288 processors but will consume only 150 kW of power.

This processor is to be embedded in a conventional LINUX computing cluster that provides for control, data storage and off-line processing. Of particular concern are the large volumes of data to be produced. An 8-beam 4-hour synthesis observation, after integration for 1 sec and over 10 kHz, generates 6 TB of raw visibility data, such that a month of observing in this mode results in a PB of data. It is therefore foreseen that data storage will be confined to several days of observation and archiving will be limited to final data products. On the other hand, transport of a single beam, 4-hour synthesis observation over dedicated 1 Gb/s network will take roughly 100 minutes, which will permit the export of selected observations to user sites, even on a regular basis.

3.5. Monitoring & Control

The Monitoring and Control subsystem (MAC) ensures that all parts of the LOFAR system work together coherently and that failures in the hardware, software or signal transport are detected. It is built around a modern commercial SCADA system, namely PVSS-II from the ETM company, augmented by adaptive capabilities relevant to LOFAR requirements. Specifically, PVSS-II provides a visualization and process control environment able to handle up to a
million or more control points among highly distributed and heterogeneous systems. It has also been adopted at CERN for control of the Large Hadron Collider. The LOFAR MAC includes a Specification and Scheduling subsystem for assisting users to prepare and execute their observing programs and for providing for the storage and retrieval of observational records.

3.6. RFI mitigation

LOFAR is among the first radio telescopes in which mitigation of the effects of unwanted man-made signals (RFI: radio frequency interference) forms an integral part of the system design. The occupation of such signals in time and frequency space in the LOFAR bands is found to be less than 10% at the site in Drenthe and there are sections of spectrum that are essentially free of interference. Nevertheless, it has been judged prudent to provide for the mitigation of such unwanted signals in the system design.

At Remote Station level, selection of the output sub-bands of width 195 kHz or 156 kHz from the poly-phase filter bank can be accomplished very rapidly, with the result that only those bands with minimal unwanted signal are retained. In the subsequent beam-forming pipeline, two strategies are available. First, chromatic beam squint is minimized by forming narrow spectral bands ~ 1 kHz in width; those narrow bands that are contaminated may then be flagged and blanked. Second, deterministic nulling (minimizing sensitivity in the direction of a stationary RFI source) may be applied in the beam-former. And finally, any remaining low level of RFI is expected to integrate away, both because it derives from different directions at the different parts of the array, and because it is either stationary or rapidly moving relative to the array while the sky revolves slowly overhead.

4. CALIBRATION

To calibrate the gains and phases of individual element antennas two approaches have been examined. The signal from each dipole may be compared to the sum of the signals of all dipoles in a Remote Station or the Central Core of which it is a part. Or the signal at each dipole from a commercial television, radio or other RFI source may be used as a standard source of known properties to stabilize the dipole. In practice the latter seems to work sufficiently well that it is currently the preferred calibration method.

Calibration of the imaging behavior of the system confronts several problems that require extension of the conventional self-calibration approach used with existing aperture synthesis radio telescopes. In general terms these problems may be summarized as follows.

- The array is high sensitivity requires very high imaging dynamic range, on the order of a million, to be able to address the proposed scientific goals. Sources in the far side lobes must be taken explicitly into account.
- The synthesized measurement beams vary in shape as a function of direction as well as system phase stability.
- The sky is very crowded with discrete sources and exhibits much extended emission at LOFAR frequencies. Imaging algorithms relying on the presence of relatively empty sky will no longer suffice.
- The ionosphere causes large phase errors at LOFAR frequencies that vary across the array on timescales of 10 sec. See Figure 2 for a typical progression of such errors in a given direction during a night.

Finding solutions to these problems represents an important research goal in itself as well as the greatest risk factor to realizing the ambitious scientific goals of the project.

The calibration approach adopted relies on the Hamaker-Bregman-Sault formulation of the detection process, which makes use of a so-called Measurement Equation. This Equation describes in a series of matrices all known effects and multiplying the matrices in the appropriate order yields the finally measured quantities, including both atmospheric and instrumental effects. Ideally, one fits the parameters describing each effect iteratively to arrive at a global solution that yields the fluxes, shapes, positions and polarizations of all sources in the field-of-view, as well as the state of the instrument and atmosphere. At the lower LOFAR frequencies, however, accounting for the problems noted above requires an exceedingly large number of parameters be fit on a very short timescale. Hence an approximate, step-wise approach developed by Noordam is being implemented. Called peeling, to distinguish it from the conventional CLEAN algorithm, the approach treats the brightest sources in the field first, removing their effects from the data, then proceeds to fit and remove progressively fainter sources. But it also involves additional steps and involves fitting many
more parameters, including position dependent phase errors and time dependent beam shapes. A fundamental
assumption made to reduce the number of parameters is that nearly all quantities including time dependences may
adequately be described piecewise by low order polynomials.

Briefly, the peeling scheme proceeds as follows. The brightest several sources in the field-of-view are identified from a
global sky model and are fit simultaneously using the Measurement Equation formalism restricted to their specific
directions and ignoring all other sources. These solutions allow one to acquire phase-lock across the field, free of \( 2\pi \)
ambiguities, and repeated fits permit tracking the phase on the required 10 sec timescale. Longer integrations on fainter
sources, up to several hundred, are now possible and in turn yield a model for the time variable ionospheric phase screen
as well as a description of, for example, the time dependent shape of the synthesized station beam. This information is
now available with full time-frequency resolution (1 sec integration time, 1 kHz radio frequency bandwidth) and may be
used to correct the incoming observational data (via a modest extrapolation of the phase trends) and integrate for
extended periods.

Such real-time conditioning of the data will be required for most applications even when imaging is not the main goal
(e.g. pulsar processing or transient detection). The necessity to acquire and track the phase may, however, limit the
rapidity with which wide field beam slewing can be accomplished in practice at the lowest observing frequencies.

These procedures will be integrated into the Central Processor computation stream. Their implementation makes
extensive use of aips++ library routines\(^8\) together with a new, optimized numerical fitting engine. It is scheduled to be
tested using new low-frequency front-end receivers on the Westerbork radio telescope, which become operational late in
2004, and will be available on LOFAR from its initial operations.

![Graph of ionospheric phase errors](image)

**Figure 2.** Measured ionospheric phase errors at 74 MHz as a function of time during a night. The
upper and lower curves show the errors on baselines of 10 km and 4.7 km respectively. The calibration
algorithms must track and remove these multi-radian errors continuously on a 10 sec timescale. Based
on data provided by Perley and Bust (private communication).

### 5. OPERATIONS

#### 5.1. Modes of Operation

The initial LOFAR system will have four main modes of operation:

- **Imaging mode**, in which station beams are correlated and visibility data pre-calibrated, integrated and post-
  processed to yield images with a manageable data-rate for archiving and export. The Central Processor is

---

542 Proc. of SPIE Vol. 5489
dimensioned such that Imaging mode data can be fully calibrated in nearly real-time. The final data-products of the processing pipeline in this mode are: (i) source data statistics, calibration data and quality assessment metrics; (ii) calibrated visibilities for all sources established during the peeling process (visibilities will not be archived but exported to SOCs; (iii) updates to the Global Sky Model (GSM); and (iv) images (following the full peeling process and including a residual image).

- **Tied Array mode**, in which signals from element antennas and/or stations are coherently added to form sets of pencil beams, each generating a time-series voltage sum. The mode is aimed at the long-term monitoring of individual sources, applications in VLBI, or possibly in special cases continuous calibration. Each tied array beam generates a raw data rate proportional to the bandwidth (e.g. 8 MSamples/sec for a 4 MHz bandwidth) independently of any spectral channelization of the data. For modest bandwidths of a few MHz, such beams can be exported either electronically or via physical media. Some applications may require a substantial number of such beams, however, which may consume significant resources at the central processing facility.

- **UHECR Detection mode**, in which the signals from a sub-set of all-sky element antennas is monitored for the signature of a cosmic ray event, following which the antenna-based data buffers are frozen and read-out for full-array near-field processing at high time resolution. These events are typically a nanosecond long and therefore require spatio-temporal synthesis processing.

- **Transient Detection mode**, which uses one or more Tied Array beams to search for transients of possibly unknown characteristics. The mode is similar to the UHECR Detection mode but usually requires full-array sensitivity and longer data sequences than can be accommodated in the antenna-based data buffers.

These modes may run in parallel and be initiated or reconfigured by a remote user, a pre-defined observation schedule, or by internal or external triggers.

### 5.2. Operations Centers

Its uniquely flexible capabilities and the very broad range of science LOFAR will be capable of addressing have led the project to consider an innovative operational model. A central Engineering Operations Center (EOC) will care for routine maintenance and the general health of the LOFAR system, while multiple geographically distributed Science Operations Centers (SOCs) will have direct on-line access to LOFAR resources for the scheduling and processing of scientific observations.

The SOCs in this concept share the responsibility for the successful scientific operation of the facility, both in carrying out specific scientific research programs that requiring development of specialized expertise, and in supporting a broader scientific community. The teams at these Centers will include expert users who are qualified to assess the instrument's behavior and the quality of the data generated, at least for the specialized application area of their SOC. That is, it seems appropriate that each SOC specialize in and be responsible for specific classes of observations of local research interest, but will also support part of the broader LOFAR user community. Block-allocation for a specified period is probably the most efficient way to allocate resources to research groups in this situation. Some projects will require all resources simultaneously and continuously for a well-defined period of time. But alternatively, resources might be allocated at various levels (e.g. eight 4 MHz beams allocated to eight different research groups for eight different projects). Such block-allocation access in a multi-institute context will require a joint Time Allocation Panel, but within the allocated blocks, the SOCs would have authority to schedule observations and manage data acquisition.

### 6. PROJECT TIMELINE

The LOFAR system described here is currently (June 2004) in the prototyping phase. Observations with prototype hardware and software are now being made remotely via a web-based interface. Final designs of some subsystems are in preparation, production software is being developed and tested on cluster computer simulators, while those sub-systems having long delivery times are already in procurement. A final design review is planned for mid-2005, following which final element antennas and much of the embedded digital hardware will be procured, installed and commissioned.

Initial observing operations are planned for late in 2006, during the minimum in solar and geo-magnetic activity, and final operational capability is foreseen for mid-2008.
Institutes potentially interested in hosting a Science Operations Center are invited to contact the LOFAR team, in the expectation that experimental observations can begin with an advanced prototype LOFAR during 2005. Extension of the facility beyond that described here is foreseen as additional partners and financing become available.

ACKNOWLEDGEMENTS

The LOFAR project has profited from the contributions of a large number of people and organizations. Of particular importance early in the project has been cooperation with the Haystack Observatory of the Massachusetts Institute of Technology and with the Remote Sensing Division of the U.S. Naval Research Laboratory, Washington D.C. Among the projectís commercial partners, several have been instrumental in helping define the facility. IBM Research in general and Bruce Elmegreen at the Thomas J. Watson Research Center in particular have been outstandingly creative in helping define the Central Processor system. Lucent Technologies Nederland has helped design the Wide Area Network. Ordina Technical Automation has researched, defined and prototyped the Monitoring and Control system, including adaptation of the PVSS-II SCADA sub-system. Science[&]Technology has analyzed the LOFAR system and proposed an automatic scheduling and hardware configuration management system including an all-important health management architecture to identify likely impending (sub)system failure. Dutch Space has made important contributions to the projectís system engineering. The RFI mitigation approach adopted has benefited greatly from many years of cooperation with Prof. A.J. van der Veen and his research group at Delft University of Technology. Design of the station level digital signal processing system has been aided by the structured design methodologies developed at LIACS of the Leiden University.

The idea for a radio telescope having a LOFAR-like architecture was proposed at ASTRON in the mid-1990ís by ir. Jaap Bregman.

REFERENCES

1. The Allen Telescope Array, described at URL: http://www.seti.org/seti/our_projects/Allen_telescope_array/
3. The BlueGene project is described at URL: http://www.research.ibm.com/bluegene/
4. The PVSS-II system is described at the web site: http://www.pvss.com/
5. Why PVSS-II was chosen for the LHC and for CERN projects in general is described at URL: http://www.manufacturing.net/ctl/article/CA206047/
7. J.E. Noordam, ñPeeling the Visibility Onion: the optimum way of Self-Calibrationí, internal project memo, LOFAR-ASTRON-MEM-078, 17 March, 2003; see presentation by Noordam at this symposium.
8. The aips++ system is described at URL: http://www.astron.nl/aips++/docs/user/Reffman/Reffman.html