Mem. S.A.It. Vol. 80, 509

Memorie della

Lofar information system design

E. Valentijn and A.N. Belikov

Kapteyn Astronomical Institute Landleven 12, 9474 AD, Groningen, the Netherlands
e-mail: [valentyn;belikov]@astro.rug.nl

Abstract. The Lofar Information System is a solution for Lofar Long Term Archive that is capable to store and handle PBs of raw and processed data. The newly created information system is based on Astro-WISE - the information system for wide field astronomy. We review an adaptation of Astro-WISE for the new tasks and propose a generic solution for the publishing of the data from the Lofar Information System to Virtual Observatory.

Key words. Astronomical data bases: miscellaneous – Surveys – Grids – Interoperability

1. Introduction

This paper discusses in detail a design for an adaptation of Astro-Wise (information system for wield filed astronomy) for the Lofar Long Term Archive.

We define "an information system" as a software suite (including user interface) accompanied with the data storage facilities. The software system allows the full circle of the data processing (both optical and radioastronomical) with the storage of the whole data chain from raw data to the final catalog. This data processing is realized in Astro-WISE, and the task is to extend distributed storage capacities of Astro-WISE as well as user interfaces, data classes, processing units to requirements of Lofar data storage.

Astro-WISE was selected as a base for developing of Lofar information system because of its high scalability, similarity in tasks, and an ability to easily adopt Astro-WISE for new data types without major changes in the system.

Send offprint requests to: E.Valentijn

2. Astro-Wise

Astro-Wise is a product of the OmegaCEN astronomical data center. Initially it was built as a storage and data processing system for OmegaCAM optical wide field imaging surveys (KIDS) to accommodate raw data for thousands of nights of observations with more than PB data storage, but in the recent years additional functionality was implemented which made Astro-WISE an universal tool for astronomical data processing, storage, and archiving.

While Astro-WISE was originally conceived to support modern optical wide field imaging experiments, the system has attracted additional communities and experiments that wish to propel their research using this infrastructure:

* ESO decided to initiate large public imaging surveys for its two new telescopes at Paranal (Chile), the 2.6 meter VST and the 4 meter Visual Infrared Telescope (VISTA), by launching a peer reviewed announcement of opportunity for the majority of the available time at the telescopes.

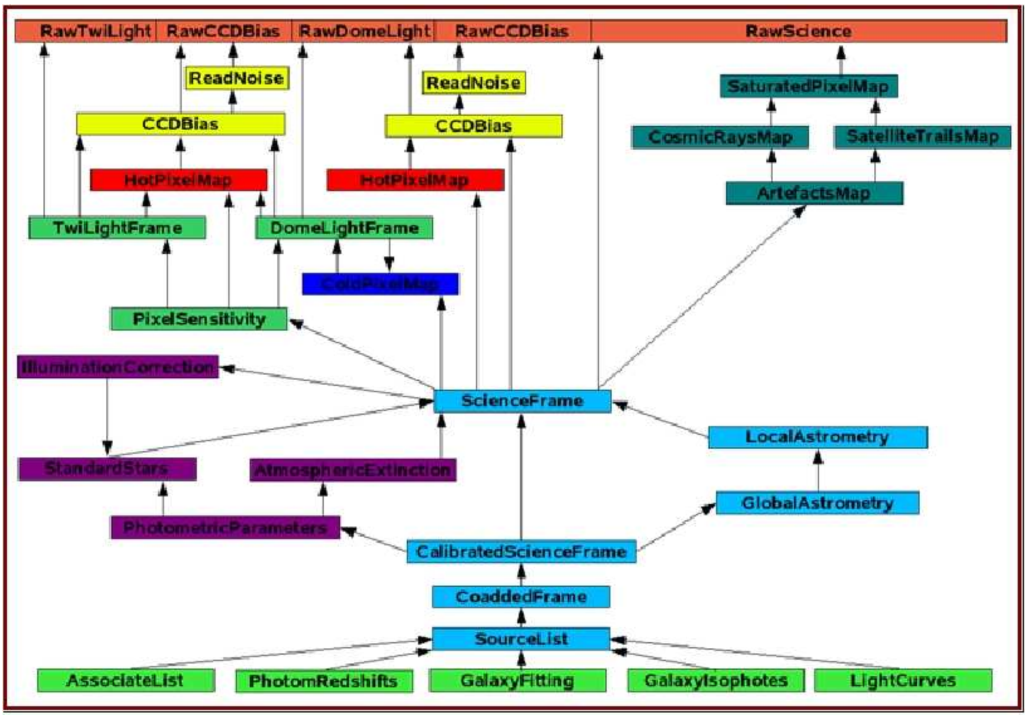

Fig. 1. Hierarchy of classes in Astro-Wise. Classes are defined according to the data processing chain.

The production, processing and analyses of these large surveys will be handled by selected international consortia, which in turn deliver the results to ESO. Most of the big VST public surveys and some of the VISTA public surveys intend to use the Astro-WISE infrastructure for their work.

* the Coma Cluster Legacy Survey, a large imaging project with the Hubble Space Telescope using the Advanced Camera for Surveys (ACS), is using Astro-WISE for data dissemination and detailed galaxy structure parameter derivation and analysis.
* the European partner (Max-Planck-Institut für extraterrestrische Physik - MPG) for the Panoramic Survey Telescope & Rapid Response System (Pan-STARRS) plans to use Astro-WISE for the production and eventual dissemination of the huge wide field imaging survey carried out with the Pan-STARRS telescope 1 (PS1) at the University of Hawaii's Institute for Astronomy. The data when delivered through Astro-WISE will serve a wide range of science goals.
* several other wide field imaging facilities, such as ESO's Wide Field Imager at the 2.2m telescope at La Silla, Chile (WFI2.2m), the wide field imager (WFC) at the UK-NL-Spain Observatorio del Roque de los Muchachos (Canarias, La Palma), the Suprime-Cam at the Japanese 8 meter Subaru telescope are supported by the system and attract new users who wish to reduce or analyse the enormous data archives.
* the artificial intelligence research group at the University of Groningen, together with the Dutch National Archive (Cultural Heritage Collections group) plans to use Astro-WISE to access kilometers of paper archives (30 Terabyte of scanned pages) by means of intelligent computing. Handwriting recognition code is run on the IBM Blue Gene supercomputer to

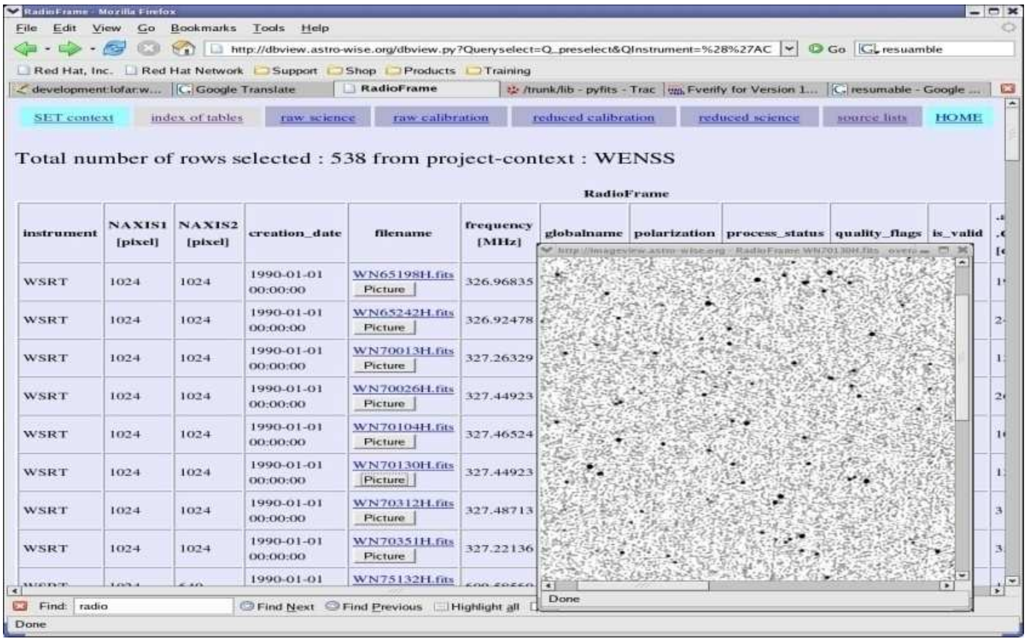

Fig. 2. Main user interface of Astro-Wise - dbview. The list of selected items is shown.

interpret the archives of the Kabinet van de[n] Koning[in] (Cabinet of the Queen), next to intense usage of the LINUX cluster. Similarities with astronomical survey IT issues are remarkable.

For all these projects, pioneer studies have taken place during the last year ensuring the technical feasibility of operating these projects in the Astro-WISE information system.

Astro-WISE is built on the Python language adding classes and libraries as well as programming environment. Relationships between classes are based on the data processing chain (see Fig. 1). New data processing chain and new classes can be implemented. The user has a standard set of procedures operating on Astro-WISE classes (for example, Sextracting objects from the image) which he can use on his local host or send for an execution to a computing element of Astro-WISE or EGEE cluster.

The user of Astro-WISE is supplied with a number of web-based interfaces to monitor, select or process data entities (images and catalogs). The database viewer (dbview) is the most popular tool and a general interface for all data entities stored in Astro-WISE (see Fig. 2). Dbview allows to select entities according to parameters set by the user and make a list of selected items. In the case of an image the user can inspect it with an additional viewer.

In addition to the browser (dbview) the user is accomplished with a command line interface (awe, see Fig. 3). The CLI allows user to call functions of Astro-WISE as well as to build a Python programs of his own. The user is allowed even to modify standard Astro-WISE code.

The Astro-WISE user is able to select data stored in Astro-WISE by tools and software of the Virtual Observatory; Astro-WISE realizes a VO Data Access Layer. The example of such an interaction is shown on Fig. 4.

The user can process data in Astro-WISE not only with the CLI but with the web-based interface as well (target processor, see Fig. 4). The target process shows already processed entities.

The last interface is a visualization of an important feature of Astro-WISE: an ability to

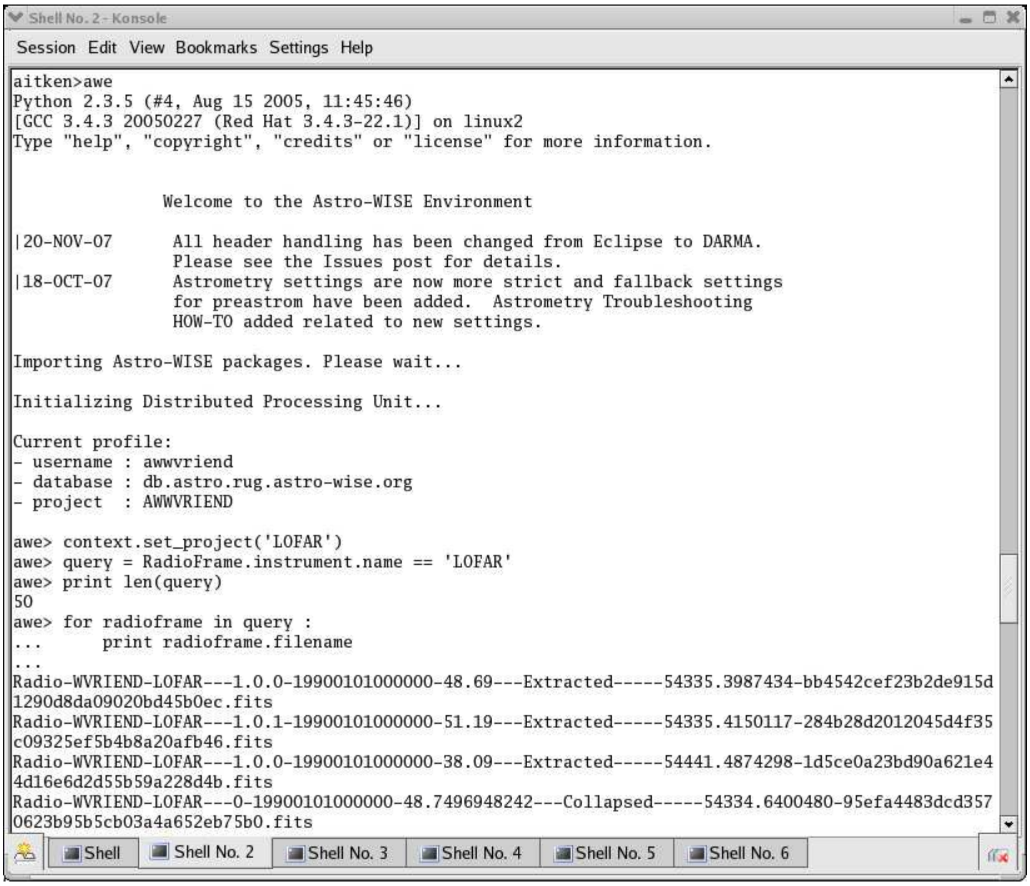

Fig. 3. Command line interface of Astro-Wise - awe prompt.

keep a history of data processing and reprocess data entities from the very raw images. All links and relationships between raw data, processing parameters, reduced images, and catalogs are kept in the system. This feature allows not only to reprocess data from the start but to avoid repeating reprocessing. The user shares data within a project and benefits from the participating in joint data reprocessing within his group.

2.1. Data storage

The data storage in Astro-WISE has two layers: metadata storage and data storage itself. Each data entity (image, catalog, radio cube) has a descriptional part of the data (metadata) and the data. Metadata is always stored in the database (RDBMS Oracle), meanwhile data is placed as FITS files on a special dataservers - the data storage host with http interface for file retrieval.

All dataservers are independent with an ability to communicate in the search for the required data. Currently, Astro-WISE has a capacity of 364 TB data storage which can be extended to PBs.

Currently, national datacenters in the Netherlands (Groningen), Germany (Garching), France (Paris) and Italy (Napoli) and satellites at Leiden (NL), Nijmegen (NL) and Bonn (D) support the Astro-WISE e-Infrastructure deploying dataservers or providing interface to Astro-WISE.

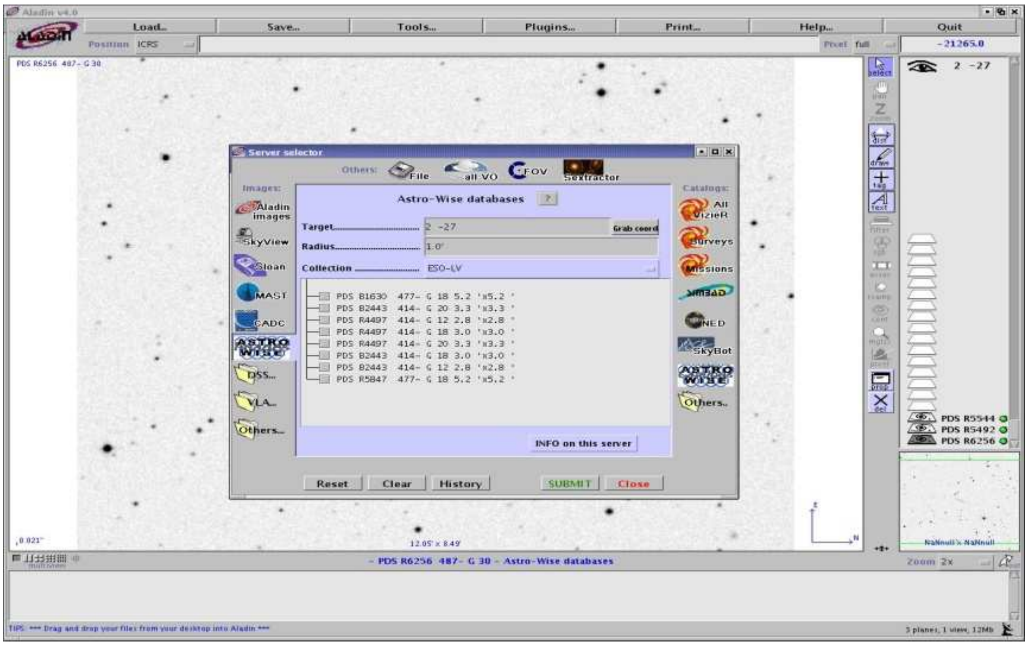

Fig. 4. Astro-Wise data in Virtual Observatory with the use of Aladin.

Astro-WISE architectural design (Fig. 6) combines computational nodes, data storage nodes (dataservers) and user interface to Astro-WISE. The picture on Figure 6 shows the installation of Astro-WISE in Groningen, the installation on other sites can only include a part of this full deployment: for example, computational node and user interface or user interface and data storage. This makes Astro-WISE much more flexible to requests of the research group which is going to use Astro-WISE. All databases installed at different locations are mirrored and synchronized, i.e. changes made at Groningen or Garching immediantly appear in the database at any other location.

As we can see, Astro-WISE has a grid structure with computing elements, data storage elements and user interface to the system. Recently an Astro-WISE interface to EGEE computing element was developed so that EGEE CEs can be incorporated into Astro-WISE structure.

2.2. Multipurpose system

The key feature of Astro-WISE is not the ability to search for the data and retrieve files but to supply user with the whole chain of data processing in one single system. The data processing from the raw images up to publishing of the resulting catalog into the European Virtual Observatory can be done within one system with the same interface.

The core of the system exploits three principles:

1. inheritance. Using object oriented programming (Python) all Astro-Wise objects inherit key properties for database access, like persistency of attributes;
2. relationship between objects. The linking (associations or references) between objects instances in the database is maintained totally, for literally each bit of information it can be traced which bits of information were used to obtain it;
3. consistency. Each step, and the inputs which it used, are kept within the system. The database is constantly growing

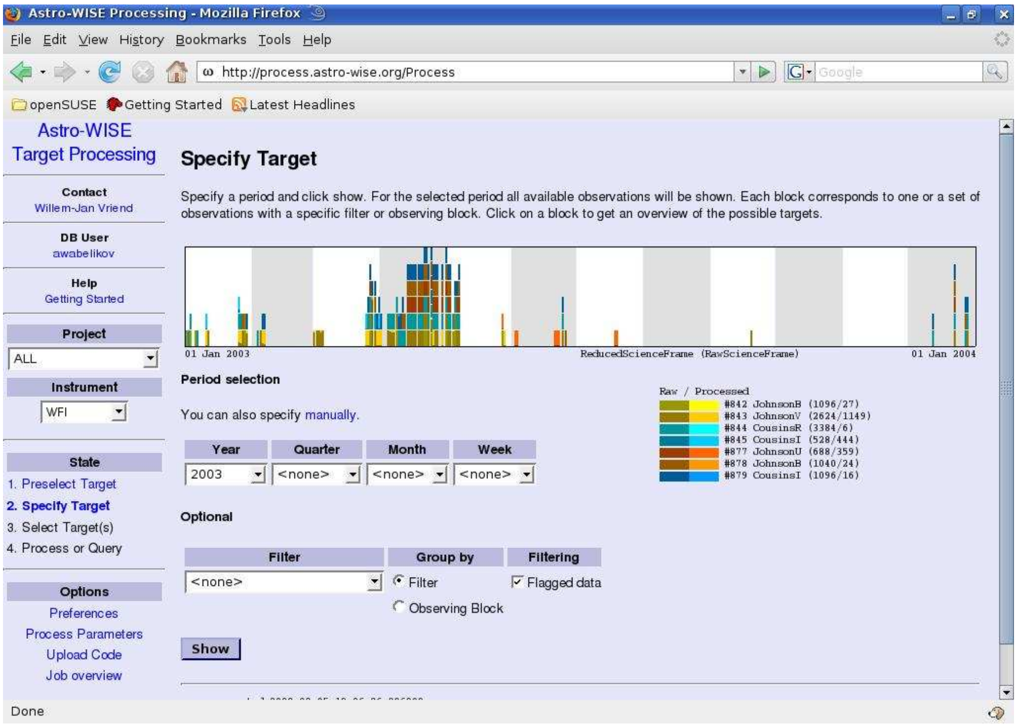

Fig. 5. Target processor of Astro-Wise.

by adding information that was not constructed before, or an improved version of existing information.

The user can not only insert and process his data but share the same dataset with other users creating a user group (project) inside Astro-WISE. The user is not bound to run the default data processing chain but can implement his own data processing. All Astro-WISE components are distributed over nodes enabling research groups in any part of the world to collaborate on shared projects. Knowledge added by one group is immediately accessible by others.

Flexible set of interfaces and detached storage of metadata and data makes it possible to adopt Astro-WISE for a lot of tasks as we saw above. Astro-WISE uses relational database but provides an ability to build an object-oriented data storage.

3. Adaptation of Astro-Wise for Lofar

The main challenge of Lofar is a data storage capacity which will exceed 20 PB for the first 5 years of observations. Because of a number of research groups participating in Lofar (each with their own resources) these data must be stored in different geographical locations but available for retrieval and reprocessing for any Lofar user.

3.1. Lofar

Lofar is a Low Frequency Array, the key radioastronomical project which targets building of a huge radiointerferometer. At the first stage of development LOFAR will consist of a compact core area (approx. 2 km in diameter, 32 stations) and 45 remote stations. Each station will be equipped 100 high band antennas (Fig. 7) and 100 low band antennas. Each

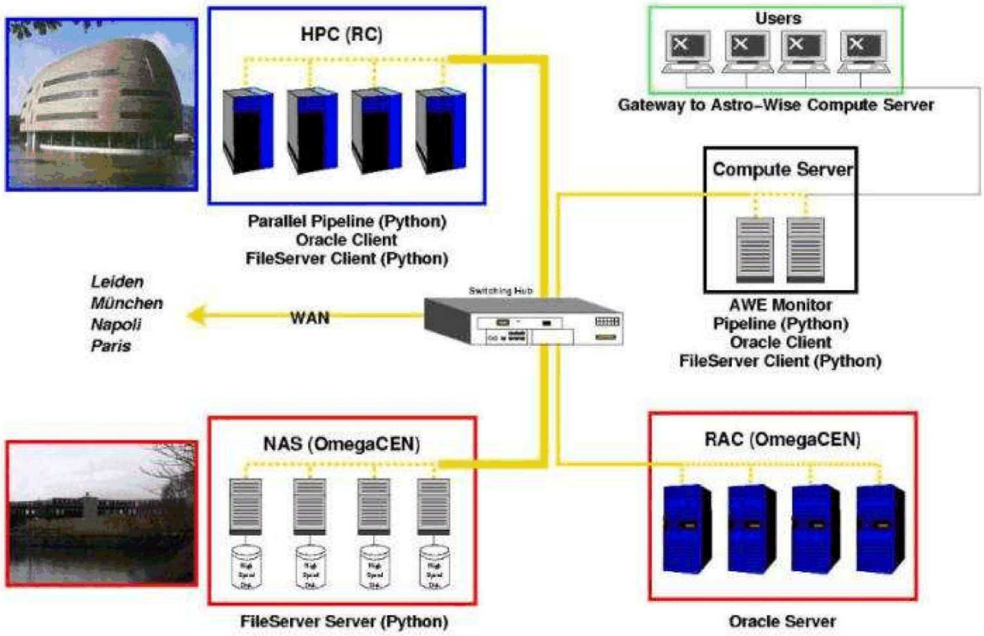

Fig. 6. Astro-WISE architectural design.

Fig. 7. Lofar high band antenna.

station connected with the Central Processing Unit by WAN with 10 Gbps net.

The principle of work of LOFAR is an aperture synthesis performed at the Central Processing Unit based on signals from each station. The Central Processing Unit is using IBM supercomputer, the BlueGene/P system and PC clusters.

Thanks to its unprecedented sensitivity, two Lofar projects offer alternative radio approaches to explore the end of the early stage of the Universe. Lofar surveys will detect extremely distant galaxies with strong radio emission, the radio galaxies, using an empirical correlation between the radio spectral steepness and distance. Only Lofar has the sensitivity and field of view to efficiently pick out radio galaxies at redshifts $z > 6$. This first time inventory of extremely high redshift radio galaxies will constrain the formation of supermassive black holes (radio emission is powered by those black holes), obtain detailed view of the interstellar medium, the fuel of star formation, which shows up as absorption of the radio emission, and identify proto-clusters, as distant radio galaxies have been shown to pinpoint those.

The LOFAR Epoch of Reionisation (EoR) project will detect directly the extremely subtle radio signal (the redshifted 21cm line emission) from gas around the EoR. Research in the last few years suggests that there may have been extended, or even multiple phases of reionisation, the start possibly being around $z \approx 10-20$ and ending at $z \approx 6$. Using LOFAR, the redshift range from $z = 11.4 - 6$ can be probed for the 21cm line emission. It is with LOFAR that one can for the first time determine the redshift range in which the bulk of the neutral hydrogen became ionized and if it happened in a single transition phase or through multiple phases of reionisation.

Lofar information system (Lo-WISE) will deal with the data delievered by the Central Processing Unit and stored in Long-Term Archive. We will review the composition of the data processing and design for the Lofar informational system based on Astro-WISE.

3.2. Lofar data processing

According to Gunst & Bentum (2007), the LOFAR architecture can be described as consisting out of the following components:

1 Stations. A station selects the sky signals of interest for a particular observation out of the total sky. This process results in one or multiple beams onto the sky. Additionally the station is able to store (raw) antenna data.
2 Wide Area Network (WAN). The WAN is responsible for the transparent transport of all the beam data (the beam signals on the sky) from the stations to the central processor.
3 CEntral processor (CEP). The central processor is responsible for the processing and combination of the beam data from all stations in such a way that user data is generated as was specified by the user.
4 Scheduling, Administration and Specification (SAS). Given the specification, the main responsibility of SAS is to schedule and configure the system in the right mode.
5 Monitoring And Control (MAC). The main responsibility of MAC is to control the system (in real-time) based upon the actual configuration of that moment. Additionally, MAC facilitates the (real-time) monitoring of the present state of the system.
6 System Health Management (SHM). SHM is identified as an autonomous block to predict and act on failures of the hardware before it actually fails. Ideally it should even pinpoint which system component is the cause of a failure. The reason for considering this block separate from MAC is because of the scale of the system and the percentage of time the system should be effectively operational.

The main responsibility for CEP is to correlate the station data and deliver a data product which can be further processed by the user.

The Central Processing Facility is divided in three sections: an on-line section for processing of real-time data streams from the stations, a storage section collecting the processed data streams and making the resulting datasets available to the third section: the off-line processing. MAC is responsible to control CEP and allocates resources to particular observations, which can run in parallel with each other (see Fig. 9).

The Blue Gene/P (BG/P) racks transpose the station data and correlate the stations with each other and consequently reduce the data. Parallel to the BG/L resources, a general purpose cluster is available for auxiliary processsing and in particular for real-time analysis, tuning and/or model creation tasks. The results of those tasks are typically used as control data for the processing applications running on the BG/L platform. The resulting data streams from the on-line processing sections are collected in the temporary storage subsystem.

The BG/P contains the connections to the stations (through the WAN). The data sent by the stations are sent in logical packages, each containing a time-frequency window of a single voltage beam. Each I/O node of BG/P will receive data from the stations and run a data handling application that will buffer the input data and synchronize its output stream with the other I/O nodes based on the timestamps contained in the data.

Large amounts of processing power and internal interconnection bandwidth are provided through the BG/P supercomputer; a peak processing power of 34 TFlops is available for the processing tasks. This processing power in combination with the IO capabilities of BG/P allows for a correlator capable to handle 2926 baselines for the full 32 MHz bandwidth together with the channel filter.

The storage system provides disk space for the collection of data streams and storage of complete observation datasets for off-line processing. This storage is intended for temporary usage (typically 5 days) until the final data products are generated and archived or the raw data i itself is exported or archived. Access to data in the storage system is through storage clients that have access to the metadata and file locations.

Finally, a general purpose Linux cluster is used for the off-line processing. The off-line processing section offers general-purpose processing power and high bandwidth interconnections to the off-line processing applications. The largest part of this cluster is a "normal" Linux cluster computer optimized on cost per Flop.

At the end of the pipeline the data will be ported to the Long Term Archive.

Fig. 8. Lofar station locations on the first phase of deployment, from Lofar Project Site.

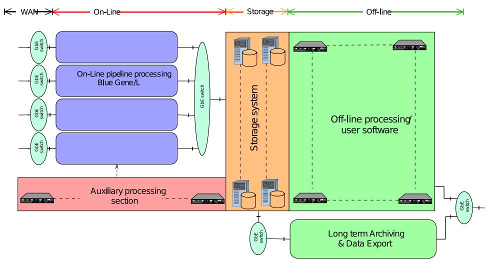

Fig. 9. The structure of Central Processing Unit of Lofar.

3.3. Tiers for Lofar

The scale of the data storage and computing for Lofar reaches the level of the Large Hardron Collider experiment (LHC), and as in the case of LHC/EGEE (Stewart et al., 2007) we adopted a tier architecture for the data storage and data processing grid (see Fig. 10). In the case of Lofar we have both similarities with EGEE architecture ("one-way" data transfer from Tier-0 to Tier-1) and differences (no data will be stored on Tier-2 level, apart from the end-user private usage).

Tier 0 is the Central Processing Unit which receive information, provides necessary computations, and deliver the data to the Long Term Archive - Tier 1. Tier 2 is an end-point for the user access to the data and computational facilities.

Tier 1 consists of a number of data storage nodes and computing nodes deployed in Groningen (Donald Smith Computing Center, CIT) or at any location where the research group participating in the project can provide a facility for data storage. Node of Tier 1 can be an Astro-WISE dataserver, distributed filesystem node (GPFS, Lustre) or EGEE node. The metadata for each data entity will contain a logical address (URL, for example) to the data file which will be converted by the corresponding server to the physical address of the file.

The user can operate from the node of Tier 2 to browse metadata database (which is stored on Tier 1 node), to start computations on computing elements of Tier 1, to retrieve, update or insert data on Tier 1.

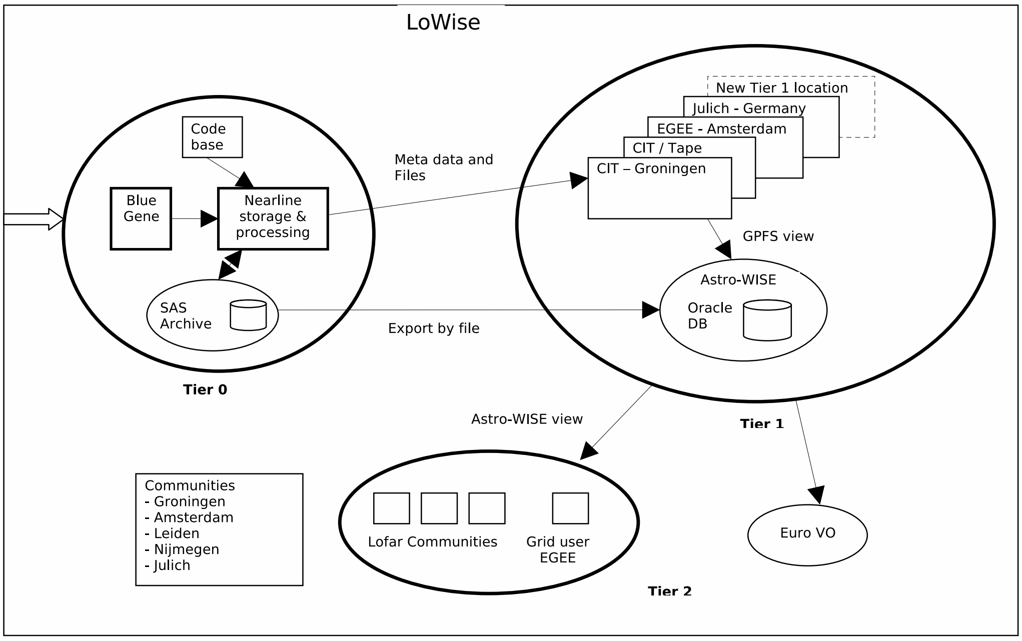

Fig. 10. The architectural design of Lofar Information System Lo-WISE.

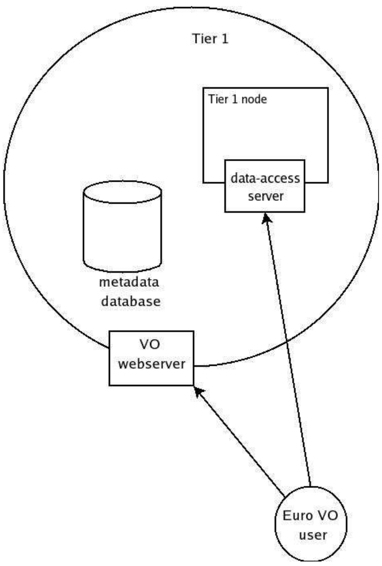

Fig. 11. The design of the access of VO user to Lofar data.

3.4. VO access

Astro-WISE provides an access to the Virtual Observatory data and an ability to publish data to the Virtual Observatory. The user of Astro-WISE can push the data to the Virtual Observatory simply changing privileges of the data entity. The same functionality we plan to exploit in Lo-WISE.

Figure 11 shows the access of the VO user to Lofar data. The user of Lo-WISE can mark a data entity as "published" in the Virtual Observatory. The VO user will able to see this entity in the metadata database of Lo-WISE and retrieve it through the specially developed web interface. We guarantee full compliance with VO standards (VOTable, ConeSearch, SIAP).

4. Conclusions

The development of Lo-WISE will create an unique information system that will able to handle data stored in different environment (distributed filesystems, EGEE data storage nodes, Astro-WISE nodes) and provide a fast an easy access to data and data processing.

The system is based on an existing Astro-WISE information system and will inherit main features and abilities of Astro-WISE. The data storage capacity of the new system will exceed 20 PB. All these data will available on-line for retrieval and reprocessing.

Astro-WISE has already announced building of the "step-in model" of the LOFAR Long Term Archive, the test ground and prototype of Lo-WISE. It will contain 200 Tbytes of storage and database servers and will be delivered before the end of year 2008.

References

Gunst A.W. & Bentum M.J., 2007 Proc. ICSPC 2007, 600

Lofar project Site, `www.lofar.org`

Stewart G.A., Cameron D., Cowan G.A. & McCance G., 2007, 5th Australasian Symposium on Grid Computing and e-Research (AusGrid 2007), Ballarat

Valentijn E.A. et al., 2007, ADASS XIV ASP Conf. Ser. 376, 491

Valentijn E.A. & Verdoes Kleijn, G., 2007, HiA 14, 607