IFERP
connecting engineers... developing research

ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 2, Issue 12, December 2017

Modernized Radio Astronomy using Data mining procedures.

[1] X.Della, [2] V.Subha
[1] Associate Professor, [2] Assistant Professor
[1] Department of Commerce with Computer Applications,Holy Cross Home Science College, Thoothukudi.
[2] Department of Computer Science, Holy Cross Home Science College, Thoothukudi.

Abstract:-- **With the advent of the new class of radio telescopes such as LOw Frequency ARray (LOFAR) and the Square Kilometer Array (SKA) the necessity of using data mining procedures to handle the huge amount of data collected during their observations became essential. Software design having complete shift of methods is used to collect information in the radio frequency domain. Digital processing of signal is collected by simple Omni-directional antennas, which is connected to the software. This technique enables the collecting elements to be spread over an area of hundreds and thousands of square kilometres, in order to map the radio sky with high angular resolution. The large area network provides the infrastructure for sensor networks with applications in other fields. With the data transport requirements of Tbits/s and processing power needs of TFlops/s this new arrays represent a real challenge for computing and existing technology. This raised the interests and stimulated the discussion between scientists and engineers. Here we will provide an overview of the current radio astronomical facilities and those planned for the future to highlight their requirements and their performances.**

Keyword:-- **Frequency, LOFAR, SKA, Radio Signal, Telescopes, Antennas.**

I. INTRODUCTION

The discovery of the radio emission arising from the Galactic plane made by Karl Jansky in 1932 opened a new window on our Universe, tremendously affecting our understanding of the physical processes occurring therein. Since then radio astronomy evolved quickly, and data reduction and analysis techniques, adopted in the modern radio astronomy, as well as new facilities and telescopes are constantly evolving. In addition, it is worth highlighting that repercussions and consequences of based on the discoveries and the scientific advances contribute both directly and indirectly to improve our lives. The development of radar techniques to mobile phone positioning, medical imaging techniques are a couple of clear examples. The same trend of discoveries it is expected once the entire next generation of radio telescopes will be fully operative.

II. GREEN ICT: FROM CORRELATORS TO CLOUDS

Green ICT can defined as embodying "design, manufacturing, utilization, disposal of computers, servers, and associated subsystems—such as monitors, printers, massive storage devices, and networking and communications systems — efficiently and effectively and with minimal or no impact on the environment" Hence, Green ICT paved the way towards more efficient intense computing systems. In fact, the biggest computing challenge within radio astronomy lies within the architecture of the correlator of big synthesis radio telescopes and the second tier processing and storage infrastructures. The correlator processes the data streams arising from the large number of antenna elements of say, with N>1000 antennas. The optimum architecture is planned to minimize power consumption as much as possible by following several approaches: minimizing I/O (storage media, and network interconnects) and memory operations, implying preference for a matrix structure over a pipeline structure and avoiding the use of memory banks and choose among the lowest power computing chip technology. For instance, the ALMA correlator selected for its core design the StratixII 90nm technology based on considerations on power dissipation and logic resources while much lower power technologies are available now. The SKA, under the Central Signal Processor Element Consortium, is currently developing design concepts for design for N>2000 and over 1 GHz frequency bandwidth, based on Application-specific integrated circuits (ASICs) fabricated in a 20nm CMOS process, still better than 20nm for FPGAS with low power considerations. Excluding antenna data pre-processing, the SKA correlator is estimated to consume less than 100 kW. After data is integrated by the correlator and further processed to create calibrated data, it must be stored in a permanent media, such as the case of massive Storage Area Networks (SANs), relying in rotational technologies such as hard disks. ALMA can output several TeraBytes of data per project that must be stored, and the future SKA infrastructure is expected to produce close to an Exabyte/day of raw information, prior to further processing and data reduction. All these data must be made available in large facilities for further reduction by researchers (eg, using

CASA). Due to the amount of information, and the costs of transmitting data through long distance optical links, it is paramount the use of computation facilities located in close proximity to the source of information, but also close to researchers, in order to reduce latency and cost of the post-analysis process. The typical approach is to create computational behemoths capable of handling the entire operation of the instruments, storage, and frequently further processing of the data produced. However, especially during the first years of operation, large infrastructures are operating with frequent interruptions caused by detection of erroneous or unexpected behavior, or when operations require further tuning. Even after entering into its normal operational status, instruments are, among other factors, affected by maintenance downtime, and also by weather conditions limiting observations. As an example, according to the ALMA cycle 0 report, over the course of 9 months (total of ~6500 hours), the instrument was allocated for 2724 hours of observation time, and this resulted in 38% (1034 hours) of successful observation. This results in a considerable efficiency loss, considering all the processing infrastructure that must be available, independent of the observation status. Although we believe the initial processing must be done close to the location of the sensors, we also believe that processing should be shared or co-located as much as possible to other already existing infrastructures, exploiting time multiplexing as a way of increasing power efficiency. Moreover, further offline reduction methods can be improved as they currently typically use dedicated hardware and facilities, which are only used after a successful observation is obtained, further increasing the total carbon footprint of science.

III. SCIENCE INFRASTRUCTURES : TOWARDS THE DATA DELUGE

A. The Low Frequency Array (LOFAR) The LOFAR Telescope, officially launched in 2010, is a novel phased-array radio interferometer containing more than 10,000 small antennas distributed across the Netherlands, the UK, Germany, France and Sweden. The core of the array, consisting of some 40 stations, is located in the North-eastern part of the Netherlands; an additional eight stations are distributed over the participating countries The LOFAR telescope facilities are jointly operated by the International LOFAR Telescope (ILT) foundation. For training and other purposes in a living lab context four so-called remote stations in the Netherlands have been identified for potential enhancements to full solar-powered operation. Note that in the case of radio astronomy, the issue of radio frequency interference is a key lay-out and design issue. That being so, the present LOFAR operation is designed around an intense real-time raw data stream of tens of Terabits/s reducing to roughly 150 Gbits/s after beamforming. This data stream is sent to the central processor (correlator), requiring partially dedicated fibre networks for long-range data transport. After correlation, typical imaging observations can easily produce visibility data at rates exceeding 35 TByte/h. After processing and analyzing, the data volumes of science products are reduced significantly, leading to an expected growth of 5 pByte per year for the Long-Term Archive (LTA).

B. The Square Kilometre Array The Square Kilometre Array (SKA) is an international multipurpose next generation radio interferometer, an Information and Communication Technology machine with thousands of antennas linked together to provide a collecting area of one square kilometre. It further involves more than 67 organizations in 20 countries, and counts with world-leading ICT industrial partners. The SKA will be built in the Southern Hemisphere in high solar irradiated zones (mainly in South Africa, with distant stations in the SKA African Partners - Botswana, Ghana, Kenya, Zambia, Madagascar, Mauritius, Mozambique, Namibia - and Australia/New Zealand). It can be best described a central core of ~200 km diameter, with 3 spiral arms of cables connecting nodes of antennas spreading over sparse territories in several countries up to 3000km distances, all in high solar irradiance latitudes. Creating a power grid covering several countries with a diameter of 3000km is impractical. It is also impractical to rely on fuel sources such as diesel, due to the remote location of most antennas. Solar Power supply is therefore an option to the power generation mix of the SKA antennas, and to contribute towards a zero carbon footprint during its lifetime, and reduced OPEX. Table 1 shows the estimated power needs of the SKA for the two installation sites. Current electronic technology projections point towards an expected target average power usage of approximately 100 MW [5]-[7], when combining all systems. Since the SKA will continuously scan the sky, it will not present strong power peaks and power fluctuations, keeping a much smoother but demanding consumption profile. Energy generation at a continental scale for this facility, with different load profiles at different locations, means that modular power generators are needed, presenting an ideal scenario for development of innovative solutions with its own degree of customization and grid connectivity. SKA is planned in two construction Phases, with deployment of different sensor technologies. In particular, performance of digital sensors [1],[6] may be driven by the electronics power consumption, as power consumption may likely cap sensor performance with a direct impact on system sensitivity. To cut consumption, innovative forms of passive cooling, plus room temperature operation,

will be considered for the work of the Low Noise Amplifiers of the antenna sensors. Therefore, to extract the maximum scientific potential and maintain costs at appropriate levels it is essential to couple the power cycles with electronic power requirements (power and cooling), in a hot, dry site, with temperatures closer to 50ºC in open field for viable operation. Together with several million so-called aperture array antennas alike employed in (e.g. LOFAR), the core with receptors along networks of three arms stretching out to hundreds of kilometers, constitute a data intense and high performance compute exa-scale level ICT based infrastructure. This should be done without degrading receiver system temperature to achieve expected sensitivities, detailed in the SKA Key Science Projects. The SKA will have different operational modes, each requiring a different computational load and power, and a different distribution of this power over the subsystems. Therefore, not only the power consumption of the SKA as a whole will vary over time, also the spatial distribution of the power consumption over the SKA systems will vary. Other aspects that may influence the distribution of power consumption are error-induced re-routing of data-streams, and future upgrades and system extensions. On the supply-side there will be additional constraints such as the interrupted availability of solar irradiance and limited availability of energy storage capacity. This clearly implies that a Smart
Grid approach, both at system level and subsystem level, is needed in order to reduce the carbon footprint of the telescope. Overall, the main characteristic concerning the SKA power system can summarized as:

• Many Antennas nodes are far away from civilization centers and power grid in climates with high thermal amplitudes.

• Exquisite control of Radio Frequency Interference and EMI from Power systems is needed, since RFI would impair the radio telescope sensitivity.

• Different Power requirements over large distances: the SKA Core will require approximately 50MW; the High power Computing around 80 stations of about 100kW over the spiral arms.

• Continuous operation (meaning 24/7 availability) for sky surveying points out that some serious storage capabilities are required, and power supply for night operations must be carefully considered.

• Power stability: control of current peaks, for operation, cooling, computing and telescope management and monitoring.

• Scalability: the power infrastructure should scale from SKA Phase 1 to the later, more extended, and more power demanding Phase2.

• Data output: ~1-10 Exabyte per day for the correlator to process; 1Petabyte of data per day for further offline reduction and analysis.

IV. CONCLUSION

The reduction of radio astronomy large infrastructures carbon footprint requires the development of innovative strategies combining power efficiency gains in electronics and intelligent data center management along the trend of Green ICT: towards Greener Processing and Storage coupled with availability of a Greener power mix. Solar Energy is being investigated as one of the power supply options to mega science projects in remote, highly irradiated locations. This will require SmartGrid management of the available power mix at system and subsystem level. Among several power option investigations, already some demo projects like the BIOSTIRLING-4SKA are studying the viability of new provision technologies like Stirling Dish-engine-based pilot plant for a 24/7 operation of radio astronomical prototype technologies, in Portugal. These developments will set a new paradigm in sustainability in large scale radio astronomy infrastructures.

V. REFERENCES

[1] Thompson, A.R., Moran, J.M. & Swenson Jr, G.W., 2001, Interferometry and Synthesis in Radio Astronomy, 2nd Edition (Wiley). ISBN: 978-0-471-25492-8

[2] Journal of Electrical and Electronics Engineering, Australia, Special Issue, 1992, Vol. 12, No. 2

[3] van Haarlem, M.P. et al. 2013, LOFAR: The LOw-Frequency Array, Astronomy and Astrophysics, 556, 2

[4] Johnston, S. et al. 2008, Science with ASKAP. The Australian Square-Kilometre Array Pathfinder, Experimental Astronomy, Vol. 22, Issue 3, 151

[5] Tingay, S.J. et al. 2013, The Murchison Widefield Array: The Square Kilometre Array Precursor at Low Radio Frequencies, Publications of the Astronomical Society of Australia, 30, 7