Astronomical Data Analysis Software and Systems XV **O10.4**
ASP Conference Series, Vol. XXX, 2005
C. Gabriel, C. Arviset, D. Ponz and E. Solano, eds.

Transforming the way VLBI is done

Huib Jan van Langevelde, Arpad Szomoru, Harro Verkouter, Mark Kettenis, Bauke Kramer, Friso Olnon, James Anderson, Cormac Reynolds, Mike Garrett

Joint Institute for VLBI in Europe, Postbus 2, 7990 AA, Dwingeloo the Netherlands

Abstract. The way VLBI is done in the European VLBI Network is subject to revolutionary changes, initiated by new equipment, often originating in the computer industry. Using high speed connectivity, eVLBI will improve the quality and scientific capacity. On the output side data volumes are also increasing, in order to image large areas of the sky. The large data volumes require new calibration tools and an environment to do distributed computing.

1. Introduction

Advances in computer and communication technology are changing the way Very Long Baseline Interferometry (VLBI) is done. Improvements in computing speed, storage capacity and communication bandwidth enable new scientific opportunities in VLBI and reshape the operations.

The power of VLBI is the ability to obtain unsurpassed spatial resolution by combining the digitized signals from several antennas, which can be several thousands of kilometers apart. The sensitivity of the observations determines how bright sources must be in order to be observable with VLBI. This sensitivity depends on the size of the participating telescopes, but also on the bandwidth over which the incoming signal can be sampled. Thus, the sensitivity depends on the bit-rate that can be recorded.

2. eVLBI

Traditionally VLBI has applied longitudinal tapes for recording typical rates of up to 256 Mb/s. Recently the EVN has switched to the so called Mk5 system (Whitney 2004), which has allowed operations at 1 Gb/s. In addition it has considerably improved the recording quality and the efficiency. The obvious next step is to eliminate the transport of magnetic media, by sending the bitstreams over Internet connections. In fact, we are aiming for a system that enables real-time VLBI, in which the data is not buffered on magnetic media anywhere, but streams into the correlator directly.

This eVLBI offers some important advantages over current practice. The first is economic; by taking out the transport costs, as well as the precious recording media a substantial reduction of operational costs is possible. We

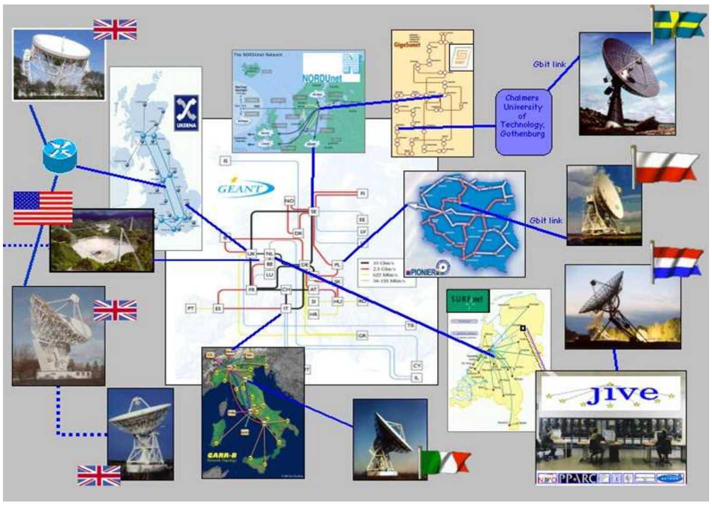

Figure 1. The various components in the EVN eVLBI proof of concept project.

think this is the way forward for increasing the recording bandwidth in the future. Moreover, eVLBI will reduce the latency between observation and user data production. This is not only important for the astronomer end user, but it also shortens the feed-back loop on quality control, spotting instrumental problems at an early stage. Finally, eVLBI will allow more interactive observing strategies, for instance to react more quickly on transient events.

In the European VLBI Network (EVN) we have conducted a proof of concept project in which up to 6 telescopes participate regularly in observing tests and even scientific observations (Fig. 1). These telescopes in the Netherlands, Sweden, Poland and the UK have fiber links and hardware for 1 Gb/s connectivity. Sometimes, even the Arecibo telescope (in Puerto Rico) participates, albeit at a reduced datarate. In order to perform these observations, dedicated fiber connections have been established in collaboration with SURF from the Amsterdam Internet exchange to the EVN correlator at the Joint Institute for VLBI (JIVE) in Dwingeloo. For this purpose a real-time mode has been implemented for the correlator.

Besides the enhancements at the correlator, a lot of progress has been made understanding the various bottlenecks associated with eVLBI. The European backbone (GEANT) has in principle sufficient capacity to support the signals. Occasionally the Gb/s connections will support up to 512 Mb/s data-rates, but often only reduced rates are possible. Understanding this requires understanding of all the intervening transport equipment (e.g. routers) and the performance of the sending and receiving hardware. The transport protocols require various parameters to be tweaked in order to reach optimal throughput.

Along with similar efforts in Japan and the USA, considerable progress has been made in the last two years. In the EVN we have reached a state where observing with 5 stations at 256 Mb/s can be realized regularly. With such capacity it is possible to image astronomical sources and in fact two science observations have taken place (e.g. Paragi 2005).

Supported by the EC, we will start in 2006 to make eVLBI an operational facility. An important aspect of this will be the physical connection of the telescopes to high-bandwidth networks; traditionally these were built at remote sites to minimize civilian interference. Transforming the operational procedures around the EVN and the correlator further is part of this project.

In fact, the project will also look at eVLBI beyond 16 telescopes at 1 Gb/s. This implies upgrading the data acquisition and upgrading the correlator capacity. The next generation correlator should also take advantage of the progress in off-the-shelf computing. However, a correlator that handles up to 10 Gb/s data-streams from 32 telescopes is still a major challenge for modern cluster super computers. It also is a challenge in data routing. In order to address this we are considering the implementation of the next generation correlator on Grid computing. In this scheme the data from each telescope will be split and sent to some compute node in Europe in order to be correlated against the data from another telescope.

There are advantages for doing correlation on standard floating point processors rather than special purpose 2-bit hardware. The correlation on such "software correlators" can be more precise, yielding a better signal to noise. It can also be more flexible, arriving at arbitrary resolution for parts of the spectrum. This has recently been shown by JIVE scientists tracking with VLBI the descent of the Huygens probe on Titan.

3. Data rates and user product

Modern computing components are transforming the input side of VLBI, but they are also enabling new scientific user products at the output side. Traditional VLBI images a very small (arc-second) patch around a source of interest. Most VLBI recordings are actually sensitive over the many arc-minute beams of the contributing telescopes. But imaging sources over a large area of sky requires fine sampling of the output visibilities in time and frequency. The effects of time and beam smearing scale with baseline length and are particularly limiting for VLBI. For this reason the EVN correlator has been upgraded with high speed hardware. each correlator rack now has a single-board computer that transfers high data rates to an offline cluster with parallel disk systems.

Recently an effective data-rate of 48 MB/s has been accomplished, marking an important milestone for the final operational system (Fig 2). This implies that the data volumes from a standard 8 hour observation may go up to reach the TByte regime. In fact, a few users have already asked for datasets of 300 - 800 Gb.

Most users still receive there datasets on DAT tapes, but this will not be feasible for these large datasets. We have started distributing data directly from the EVN data archive. On the archive server all the user data from the EVN are permanently kept on-line. As the user has proprietary rights the datasets are password protected for the first year. The archive has quite a bit of additional functionality. It also serves as a tool for users to assess the data quality of their experiment, and the initial calibration and even preliminary images are available from there. Moreover, it is also a database where station performance can be

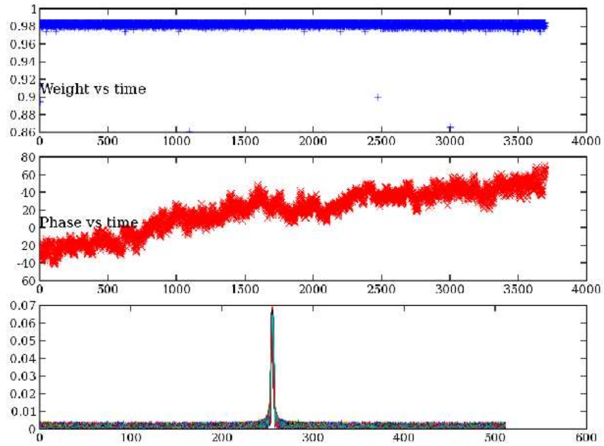

Figure 2. High speed data read-out of the EVN correlator, the data on the Effelsberg-Medicina baseline is sampled at 1/16s.

monitored. Although VLBI is not obviously suited for VO access, an effort is made to make the data available with similar methods.

The high resolution entries in the archive call for new and different user products. For example a user may be interested to get data for a target in the field. We are contemplating an interactive service in which the data can be selected and averaged before downloaded by the user. Besides selection on target position also simpler operations should be possible, such as making a selection in frequency or baseline.

This latter activity is part of the RadioNet effort called ALBUS (for Advanced Long Baseline User Software). This is a collaboration between JIVE, ASTRON Dwingeloo, MPIfR in Bonn and Jodrell Bank Observatory. It addresses various aspects of radio interferometry data processing. A number of calibration methods are to be made available more transparently, such as ionospheric and tropospheric calibration. Other parts focus on distributed computing, which is necessary in order to deal with the enormous datasets that will be produced. In order to accommodate these algorithms and make them distributable to the user domain, an Python interface to classic AIPS has been developed. This software has been called "ParselTongue" and is described by Kettenis et al., (2005).

References

Kettenis, M., van Langevelde, H.J., Cotton W,D., Reynolds, C., this volume, [P.4]

Paragi, Z., Garrett, M.A., Paczynski, B., Kouveliotou, C., Szomoru, A., Reynolds, C., Parsley, S.M., Ghosh, T., 2005, MSAIt, in press, astro-ph/0505468

Whitney, A., 2004, in proc of the 7th EVN symposium, eds R. Bachiller, F. Colomer, J.-F., Desmurs, P. de Vicente, Observatorio Astronomico Nacional, p 251