Mem. S.A.It. Suppl. Vol. 13, 79

Memorie della
Supplementi

Radio astronomy data reduction at the Institute of Radio Astronomy

J.S. Morgan

INAF – Istituto di Radioastronomia, Via P. Gobetti, 101 40129 Bologna, Italy.
e-mail: jmorgan@ira.inaf.it

Abstract. Correlation is the fundamental step in data reduction for all radio interferometers. Traditionally this computationally expensive process has been carried out by purpose-built hardware, however, so-called software correlators running standard computer clusters are becoming more popular. We report on the use of the DiFX software correlator at the Istituto di Radioastronomia. The use of this correlator along with existing antennas and other infrastructure provides the institute with all the resources required to perform VLBI. We discuss the experience we have had with this correlator and present some initial results.

Key words. Instrumentation: interferometers – Methods: data analysis – Techniques: interferometric – Radio continuum: general – Radio lines: general

1. Introduction

1.1. The correlation process

Correlation is the process of multiplying and accumulating (time-averaging) the waveforms from two or more antennas in an interferometer array (Thompson et al. 1986, 3.2). For typical interferometers this involves processing very large volumes of data. However, the computation is embarrassingly parallel.

It should be noted, however, that in practice correlation involves more than this simple mathematical process: The correlator requires further information in addition to the digitised waveform. Other parameters, such as clock delays are also required. Moreover, in order to process the data, the precise positions and velocities of the stations with respect to the source must be calculated for the epoch of the experiment.

Finally the data must be converted into a format readable by standard data analysis software.

Thus for our purposes correlation is the process whereby we transform a digitised broadband data (stored in a Mark 5 format) into a uv dataset (in standard UV FITS format).

1.2. Very long baseline interferometry in Italy

Italy participates regularly in several VLBI programs. Two Italian dishes, Medicina and Noto, regularly participate in the European VLBI network. In addition, along with another dish at Matera, they observe regularly for the International VLBI Service for Geodesy and Astrometry

Send offprint requests to: J.S. Morgan

The Sardinia Radio Telescope which is currently under construction will be a very valuable instrument for VLBI particularly at shorter wavelengths.

1.2.1. e-VLBI

For Antennas which are linked to high-speed fibre networks it is possible to send the data from the antenna to the correlator in real time. Medicina has already participated in European and global VLBI observations using a 1 Gbps data link to the Joint Institute for VLBI in Europe (JIVE) in the Netherlands. It is hoped that the other antennas in Italy can soon be added to this high speed network. A high speed connection to the Sardinia Radio Telescope is planned once it becomes operational.

While these instruments will regularly participate in international VLBI, the four antennas would, in principal, form an effective interferometer on their own.

2. Correlators

2.1. Hardware correlators

Traditionally the correlation process is carried out by purpose-built hardware dedicated to a particular interferometer. Examples include the VLA correlator (Socorro, NM, USA), and the Merlin correlator (Jodrell Bank, UK). This approach is also being used for many instruments which are currently being built or upgraded. Examples include ALMA, e-MERLIN and the EVLA.

Examples of VLBI correlators include the VLBA Correlator (Socorro, NM, USA), the JIVE correlator (Dwingeloo, The Netherlands), and the VLBI correlator at the Max-Planck-Institut für Radioastronomie (Bonn, Germany).

2.2. Software correlators

For LOFAR(Falcke et al. 2007), a slightly different approach has been taken. The telescope uses an IBM Blue Gene/L computer, for which a purpose-built software correlator has been written. While certainly non-standard hardware, the IBM Blue Gene/L is an all purpose super computer which could, in principal, be used for other tasks.

2.2.1. DiFX

Another approach is to build a software correlator which runs on standard computer hardware. We are using DiFX which was developed at the Centre for Astrophysics and Supercomputing, Swinburne University of Technology (Deller et al. 2007). It uses MPI and the intel IPP libraries to allow any standard x86 cluster to be used as a correlator.

2.2.2. Some advantages of software correlation

Software correlators generally allow greater flexibility in choosing the parameters of the correlator output such as the number of channels and integration time. The volume of data which can be processed is limited only by storage requirements and, of course, processing time. Provided the processing power is sufficient for real-time correlation, the correlator can be used for e-VLBI. It is also feasible to modify the source code if a particular feature is required. As a result software correlators are often used for unusual experiments such as pulsar imaging and space-based VLBI.

Since most astronomy institutes have access to a high performance cluster, software correlation allows researchers to take control of the correlation process themselves. Advantages to this approach include the ability to correlate more than once based on the results of an initial correlation. The cluster can also be used for further data reduction once the correlation is finished.

Another great advantage is the ability to upgrade the correlator steadily simply by adding new machines. This is in stark contrast to hardware correlators, many of which continue to be in service after several decades.

For interferometers with large numbers of baselines and/or very large bandwidths, the running costs of software correlators become prohibitive. For VLBI however, the bandwidth is limited by data transport considerations and

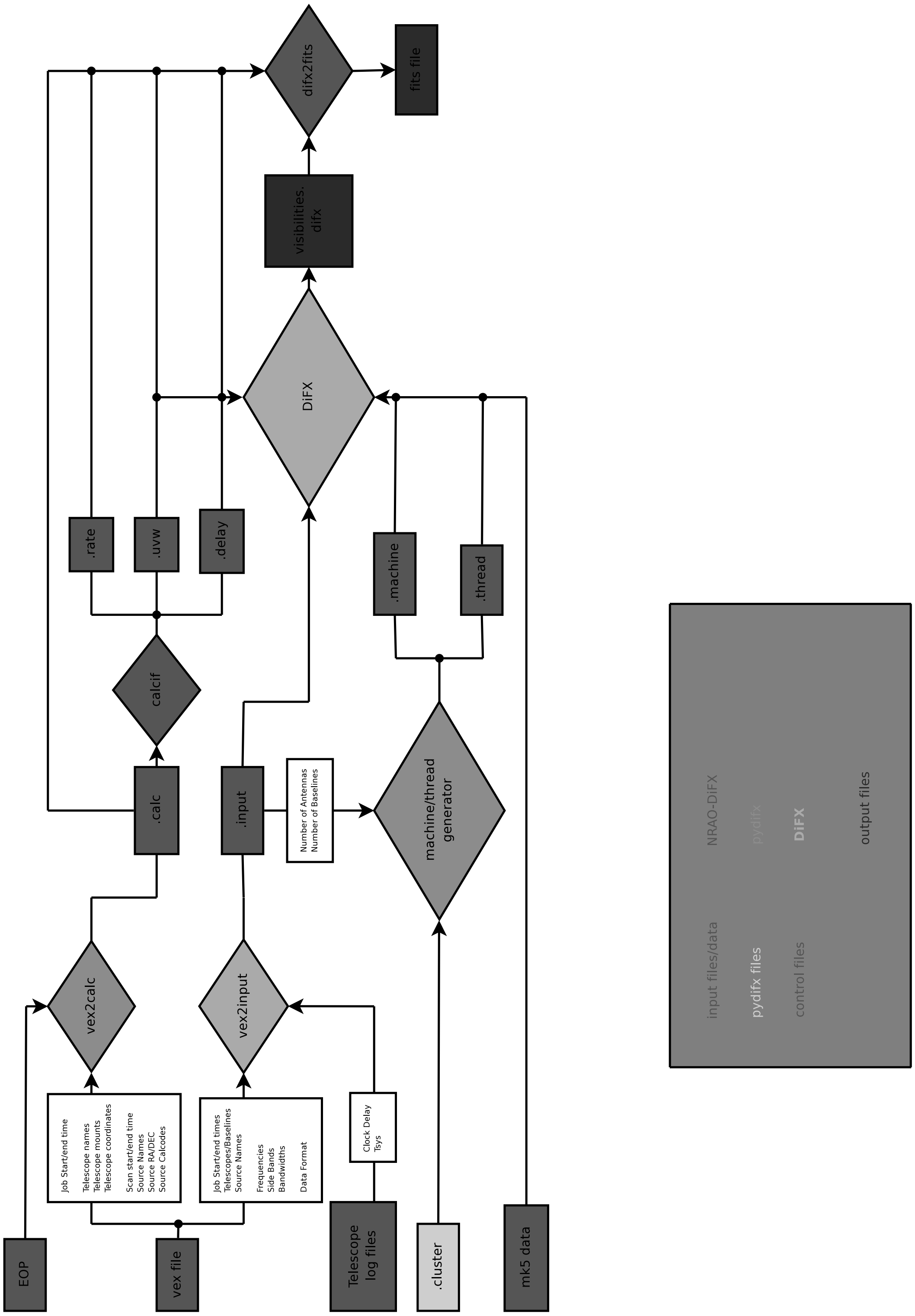

Fig. 1. Correlator pipeline

Table 1. Cluster used for the correlation

Machine	Number of Cores	Type of Processor	Clock Speed
wn01	2 (2 × single core)	Intel 4	3 GHz
wn02	2 (2 × single core)	Intel 4	3 GHz
wn03	2 (2 × single core)	Intel 4	3 GHz
wn04	2 (2 × single core)	Intel 4	3 GHz
wn05	4 (2 × dual core)	AMD Opteron 270	2 GHz
wn06	4 (2 × dual core)	AMD Opteron 270	2 GHz
wn07	4 (2 × dual core)	AMD Opteron 270	2 GHz
wn08	4 (2 × dual core)	AMD Opteron 270	2 GHz

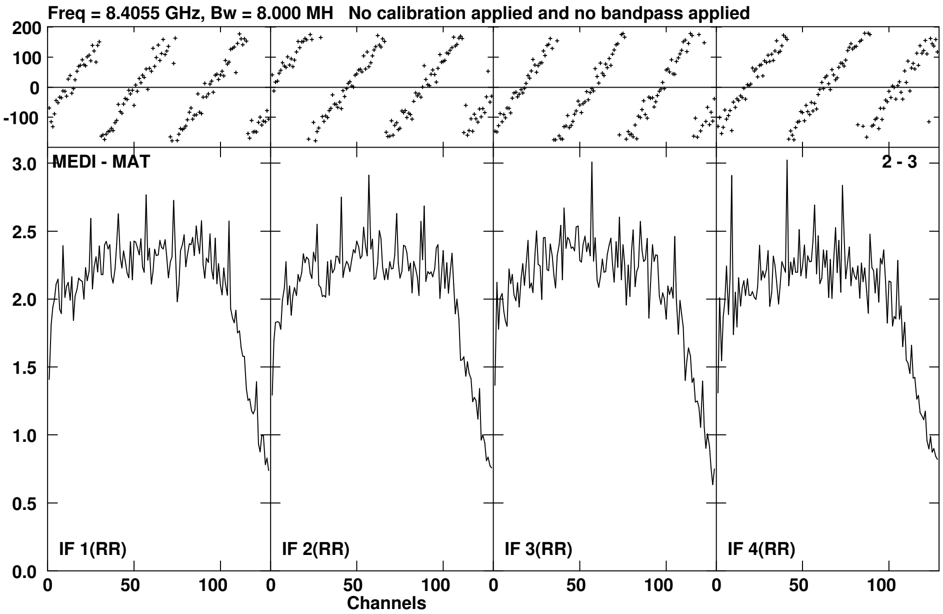

Fig. 2. A plot showing the amplitude (arbitrary scale) and phase (degrees) for a single baseline across all channels and IFs integrated over 1 minute.

VLBI networks typically consist of less than 10 stations making software correlation an attractive solution.

3. DiFX at the Institute of Radioastronomy

3.1. Our cluster

We have installed the DiFX correlator on a heterogeneous cluster consisting of 8 machines connected with a 1 Gbit switch (Table 1). This 24 CPU cluster represents a minimum amount of computing power to correlate sensible amounts of data.

3.2. Managing the correlation

While the correlation process is very straightforward, a comprehensive suite of software is required in order to process the input and log files, and produce output in the desired format (see figure 1). We have chosen to build upon the NRAO-DiFX pipeline (Brisken 2008). This is a collection of software tools being developed at the NRAO which is designed to give identical output to the VLBA correlator. Additional tools are required in order to make this pipeline compatible with the vex files commonly used for controlling European VLBI experiments. The details of this system will shortly be available in a technical report.

4. A VLBI experiment to test the correlator

4.1. Experiment parameters

In order to test the capabilities of this software correlator we scheduled a ad-hoc 4 station VLBI experiment. While the parameters of the experiment where chosen to be as undemanding as possible, they are nevertheless sufficient to produce scientific results. The Effelsburg 100m dish, along with 3 other dishes (Medicina, Matera and Wettzel) observed for 3 hours with 4×8 MHz bandwidth producing approximately 100GB of data per station. This produced results similar to those we would expect from a 4-station Italian VLBI network.

4.2. Data handling

500 GB of data is not a particularly large volume of data to store, however this is still a significant volume of data to transfer over a network. We were able to use the European e-VLBI network to transfer some of the data, others were sent via courier on a physical medium.

4.3. Correlator performance

In spite of the modest specifications of our cluster, we found its performance to be perfectly acceptable. For this experiment the cluster is able to correlate 6 minutes of data in around 8 minutes.

4.4. Results

The results of a correlation are shown in figure 2. Technical problems with the Effelsburg have prevented us from analysing the data further. Nevertheless we have demonstrated that the institute now has the ability to correlate VLBI data in-house.

5. Further work

Working with uncorrelated data gives us very good experience in handling the large volumes of data which will become commonplace once the next generation of inteferometers is completed. It is also a useful platform for exploring the use of high performance clusters for radio astronomical data reduction. This experience will be very valuable as the institute prepares for its role as an ALMA regional centre.

Acknowledgements. Many thanks to S. Tingay, M. Nanni and F. Mantovani. This research was supported by the EU Framework 6 Marie Curie Early Stage Training programme under contract number MEST-CT-2005-19669 “ESTRELA”.

References

Brisken, W. 2008, A Guide to Software Correlation Using NRAO-DiFX Version 1.0, Tech. rep., NRAO

Deller, A. T., Tingay, S. J., Bailes, M., & West, C. 2007, PASP, 119, 318

Falcke, H. D., van Haarlem, M. P., de Bruyn, A. G., et al. 2007, Highlights of Astronomy, 14, 386

Thompson, A. R., Moran, J. M., & Swenson, G. W. 1986, Interferometry and synthesis in radio astronomy (New York, Wiley-Interscience, 1986, 554 p.)