The future of long-baseline interferometry

Simon Garrington, Anita Richards, Anthony Holloway

MERLIN/VLBI National Facility
University of Manchester
Jodrell Bank Observatory
SK11 9DL, UK
Email: `stg@jb.man.ac.uk`

ABSTRACT

The fibre connection of long-baseline interferometer arrays, such as MERLIN and the European VLBI Network will improve their sensitivities by up to an order of magnitude, opening up new areas of scientific discovery. However, the volumes of data produced will often be too large for many users to handle at their desktops. We describe the pipeline approach used to establish the MERLIN archive of processed images and the prospects for remote processing of data on clusters, including preliminary results from the COBRA cluster at Jodrell Bank Observatory.

INTRODUCTION

Long baseline radio interferometer arrays and VLBI networks (MERLIN, EVN, VLBA) have provided radio images with exquisite resolution (1-150 mas), but their sensitivities have been limited by bandwidth, their fields-of-view have been restricted by their output data rates, and their use has often been hampered by the lengthy data reduction required.

Optical fibre technology now offers a straightforward way to increase the rate at which data can be transported to the correlator. At observing frequencies of 1-30 GHz, the observing bandwidth can be increased to several GHz, compared to the current limitations of several tens of MHz.

e-MERLIN: Opportunities and Challenges

From its conception almost 30 years ago, MERLIN was designed to provide sub-arcsec images at centimetre wavelengths. The radio-linked interferometer, pioneered at Jodrell Bank, enabled the development of an array spanning 217-km with a resolution of 50 mas at 5 GHz, but limited the bandwidth to 15 MHz per polarization. The e-MERLIN project, now funded, will exploit existing optical fibre networks to increase the data rate from 128 Mb/s to 30 Gb/s from each telescope (see R. McCool et al, these proceedings). In conjunction with improved receivers (based on those developed by JBO for the Planck satellite) and the resurfaced Lovell Telescope e-MERLIN will have 40 times the sensitivity of the current MERLIN array.
A standard 12-hr observation should reach a sensitivity of 1-2 uJy/beam, better than the most heroic efforts with deep integrations made to date, and long integrations with e-MERLIN will explore populations of radio sources well below our current detection limits.

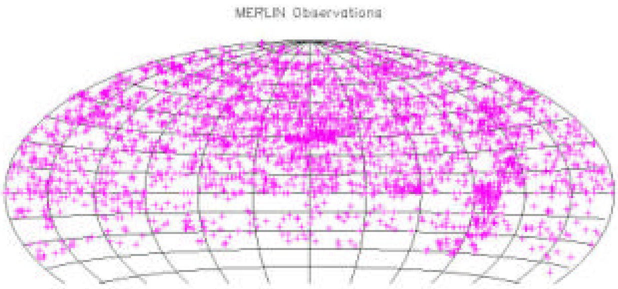
All observations with e-MERLIN will be made in wide-field mode, ie with integration times of <2s and up to 4096 frequency channels across the 2 GHz observing band. This will allow the whole primary beam of the individual telescopes (8 arcmin diameter at 5 GHz) to be imaged at the full resolution of 50 mas, but the output data rate of the correlator will be 0.5 TB/day even for these standard continuum observations.

The correlator output will be handled by an array of data acquisition computers, each responsible for a number of sub-bands. A separate processing cluster will be used for the initial processing of data using a 'pipeline' approach to provide users with an initial image.

In order to achieve the full sensitivity allowed by the bandwidth, it will be necessary to subtract the effects of faint 'background sources' within the field of view. This will require 'coarse imaging' of the whole field followed by detailed imaging of the individual sources and then subtraction of their contribution to the visibility data. The wide bandwidth will also provide almost complete aperture coverage for e-MERLIN via Multi-Frequency Synthesis, but special deconvolution techniques will be required to handle the effects of variations in source properties across the observing band (eg Conway et al 1990)

PIPELINES AND ARCHIVES

Processing data from long baseline arrays includes a number of steps which are unfamiliar to astronomers outside the field and these present a significant barrier to the use of long baseline interferometers. Although many expert users prefer to get their hands dirty and work from scratch, several groups have developed 'pipeline scripts' to automate the process. For surveys, such as FIRST, NVSS and CLASS, this has been a necessity, and this approach has worked well with these VLA data. For longer baseline, or more heterogeneous arrays , such as MERLIN or EVN, developing robust pipelines is more challenging, especially since these sparse arrays can have calibration information of a more variable quality. Pipelines are now regularly used for both MERLIN and EVN user projects.



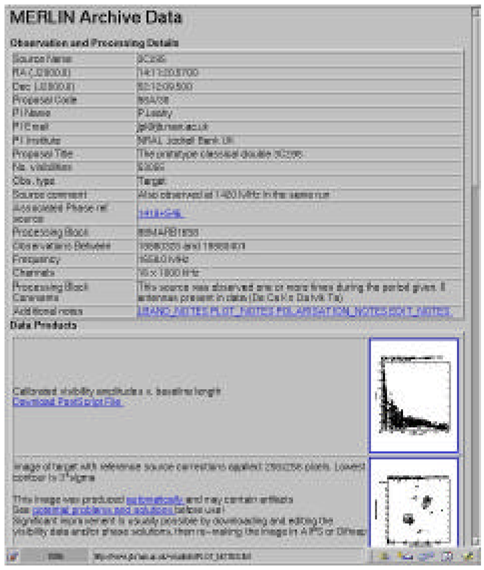


Fig.1: Sky plot of MERLIN observations; about 20% of these have been processed in the archive so far. Example output from the MERLIN archive.

While all long baseline arrays hold archives of their raw correlator output, the processing of the data has proved a disincentive to retrieving the data from earlier projects. For MERLIN, it was decided to produce a uniform archive of calibrated and fully processed data, including reference images, for all continuum observations. This archive can be searched on the web and processed data and fits images can be downloaded (see www.merlin.ac.uk/archive)

C.Reynolds has also developed a pipeline for EVN user experiments (C.Reynolds, these proceedings, see also www.evlbi.org/user_expts/user_expts.html)

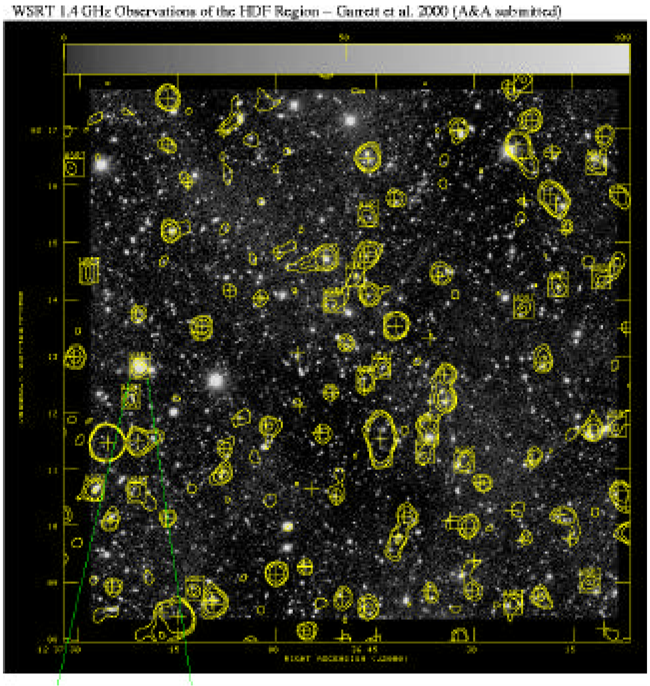
So far, these pipelines have been implemented in AIPS, but aips++ offers an ideal environment for developing scripts where individual tasks can be efficiently farmed out within a cluster (eg Willis 2000).

REMOTE PROCESSING

While some users may be satisfied with images produced by pipelines, there will always be cases where access to the visibility data is essential. In most cases, expert users will be able to improve upon the images produced by pipelines either by more careful editing of the visibility data or by careful control of the imaging and deconvolution.

For wide-field imaging it is much more efficient to extract small images covering regions of interest by applying a phase shift to the visibility data followed by a relatively small FFT, than to attempt to image the whole field and make an image cut-out. Fig. 2 shows an example of this technique to select sub-field of the Hubble Deep Field at 24 mas resolution (Garrett et al 2001, see also these proceedings).

For large data sets it may be impractical for many users to receive the data on tape or disk to load and then process it. Since the data stores may themselves be sets of disks, it may be simpler to access the data remotely and do the processing with CPU attached to the archive.



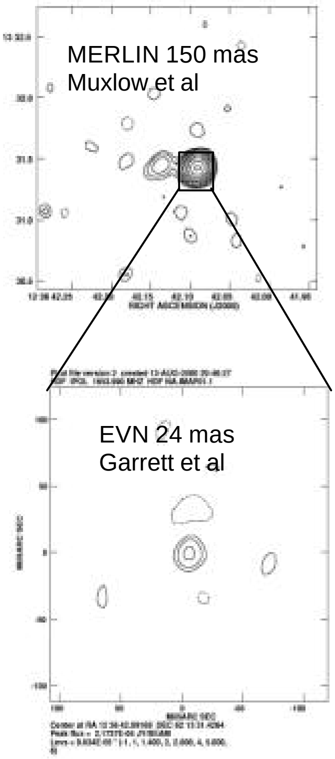


Fig. 2 The Hubble Deep Field at increasing resolution. The large image shows the 10' square field centred on the HDF with contours of the WSRT image at 15" resolution (Garrett et al 2000). The inset shows details of a z=4.4 galaxy at 300 mas with MERLIN+VLA (Muxlow et al 2002) and at 24 mas resolution with the EVN (Garrett et al 2001).

With e-MERLIN or e-VLBI, the majority of observations will provide background fields which are deeper than current long integrations. These data will provide a vast resource on the nature of faint radio sources if they can be accessed efficiently.

We are currently investigating the use of COBRA at Jodrell Bank Observatory as a 'visibility data server' to provide thus type of access to large MERLIN and EVN data sets. COBRA is a 'super-cluster' of 182 1.13 GHz CPUs connected with a 132 MB/s SCI interconnect. We are using the aips++ implementation of pimager (Roberts et al 2000) to investigate the scaleability of this extraction process on large CPU clusters.

ACKNOWLEDGEMENTS

This work is supported by the UK AstroGrid project (www.astrogrid.org) and the EU funded Astrophysical Virtual Observatory (www.eso.org/projects/avo). We would like to thank Wes Young (NRAO) for his help with pimager.

REFERENCES

Conway, J.E, Cornwell,T.J, Wilkinson, P.N. 1990, MNRAS 246, 490
Garrett, M. et al 2000, A&A 361, 41
Garrett, M. et al 2001, A&A 366, 5
Muxlow, T. et al 2002, in prep.
Roberts, D. et al 2000, in ADASS IX, eds Manset, N. et al San Francisco: ASP
Willis, A.G, 2000, Jodrell Bank SKA Workshop (www.jb.man.ac.uk/ska/workshop/Willis1.pdf)