Parallelization of wide-field imaging in AIPS++

K. Golap[(1)], A. Kemball[(2)], T. Cornwell[(3)], W. Young[(4)]

[(1)] ***National Radio Astronomy Observatory,***
1003 Lopezville Road, Socorro, NM 87801 USA.
e-mail: kgolap@nrao.edu
[(2)] ***As (1) above but e-mail: akemball@nrao.edu***
[(3)] ***As (1) above but e-mail: tcornwel@nrao.edu***
[(4)] ***As (1) above but e-mail: wyoung@nrao.edu***

ABSTRACT

Large synthesis arrays in radio astronomy, operating at low frequencies, require wide-field imaging using a full 3-D Fourier transform rather than the conventional 2-D approximation. The wide-field imaging problem results from a geometric effect: image synthesis arrays can have different projected geometrical shape from different points in the sky being imaged. This problem presents unique computational and algorithmic challenges.

Scalar CPU performance, as governed by Moore's law, is insufficient for current scientific needs in wide-field imaging. New telescopes under construction plan to provide more interferometer elements and many-fold increases in available bandwidths and data rates. Parallelization is the only solution in this case.

Overview of the wide-field imaging problem

Imaging using synthesis arrays involves inverting a 3-D integral (given below) to obtain the brightness distribution on the sky, $I(l,m)$, from a measured set of visibilities, V(u,v,w), in the *uv* plane. In most practical cases, the non-coplanar term w can be neglected and the inversion is a direct 2-D Fourier transform.

$$V(u,v,w) = \int I(l,m) \exp j2\pi(ul + vm + wn) \frac{dldm}{\sqrt{1-l^2-m^2}}$$

The wide-field problem occurs when imaging large fields of view with relatively long baselines and non-coplanar arrays

However, for wide-field imaging, if the 'w' term is not taken into account there is usually a substantial loss of dynamic range and it is also impossible to faithfully image regions far from the field center.

The wide-field imaging algorithm used in AIPS++

Several algorithms exist to solve the full 3-D problem listed above (Cornwell & Perley 1992). In AIPS++ a multi-faceted transform approach has been chosen for its efficiency. This covers the image plane by a series of facets, in each of which a 2-D transform holds.

In this formalism we can decompose the visibilities into a summation of re-phased faceted visibilities:

$$V(u,v,w) = \sum_k V_k(u,v) \frac{\exp j2\pi(ul_k + vm_k + w\sqrt{1-l_k^2-m_k^2})}{\sqrt{1-l_k^2-m_k^2}}$$

where :

$$V_k(u,v) = \int I_k(l-l_k, m-m_k) \exp j2\pi(u(l-l_k) + v(m-m_k)) dldm$$

The iterative multi-stage algorithm implemented in AIPS++ proceeds as follows:

- Calculate residual images for all facets(using 2-D transforms)
- Partially deconvolve individual facets and update the image model for each facet

- Reconcile different facets by subtracting the model visibility for all facet models from the visibility data.
- Recalculate residual images and repeat. In the process of making residual images, a *uv* plane coordinate system is chosen so that the final image from each of the facets is projected on a common tangent plane (Sault et al. 1996).

The parallelization effort and progress

The parallelization model we are using is that of a master processor which distributes data to, and collects back from, worker processors. We are using the Message Passing Interface (MPI version 1), standards and libraries to pass information between processors. The design of the parallel algorithms and usage of MPI ensures that the parallel version of 'imager' is portable and will run on either a cluster or on shared memory multi-processor machines.

It should be noted that as the visibility data can be very large we have chosen that the worker processors each access the data sets independently without going through the the master.

The first level of parallelization was aimed at parallelizing the nearly embarrassingly parallel sections of the wide-field algorithms. There are three distinct sections which we have identified in the wide-field algorithms which fall under this category and which have been parallelized:

- The formation of the point spread function (PSF). The PSF for each facet is needed in deconvolution. These can be estimated totally independently of each other, requiring only the *uv* coverage seen from each facet.
- The model visibility estimation from the source model components. As the visibilities from different sources (or different facets) are additive, they can be estimated independently for each facet model and cumulatively added into the final model visibility. This has parallel I/O implications.
- The residual image estimation. In calculating the residual image the residual visibility re-projection for the different facets can be estimated independently.

An example

The image below, Fig. 1, comes from a dataset of ≈ 600000 visibilities from a VLA observation at 74 MHz in the B and C configurations (R. Perley et al.). Reducing such a dataset using the AIPS++wide-field algorithm with 225 facets takes close to 20 days to process on a desktop workstation (SGI Octane). A similar observation in the A-array of the VLA would require some 10 times more computer resources to process. Along with other overheads, such as improved deconvolution algorithms for larger baselines, we are facing computation of 200 to 300 days on a typical desktop. This problem strongly justifies the need for parallelization of this algorithm. The problem will be more pronounced with future arrays such as the E-VLA.

The image below comes from reducing the dataset using the parallelized version of AIPS++imager for the widefield case. It took slightly more than 6 hours using 32 processors on a origin 2000 SGI.

Ongoing and future work

The parallel 3-D imaging approach is close to full operational use. Areas of ongoing work include:

- Parallelize interferometric-mosaic imaging.
- Migration loosely coupled clusters and fully measure the speed up of each algorithm computed.
- Further work in Parallel I/O using MPI-2
- Investigation of statement level parallelization using OPEN-MP.

REFERENCES

[1] Cornwell, T. J. and Perley, R. A. Astron. & Astrophys, 261, 353-364, 1992.

[2] Sault, R. Staveley-Smith, L. and Brouw, W. N. Astron. & Astrophys. Suppl., 120, 375-384, 1996.

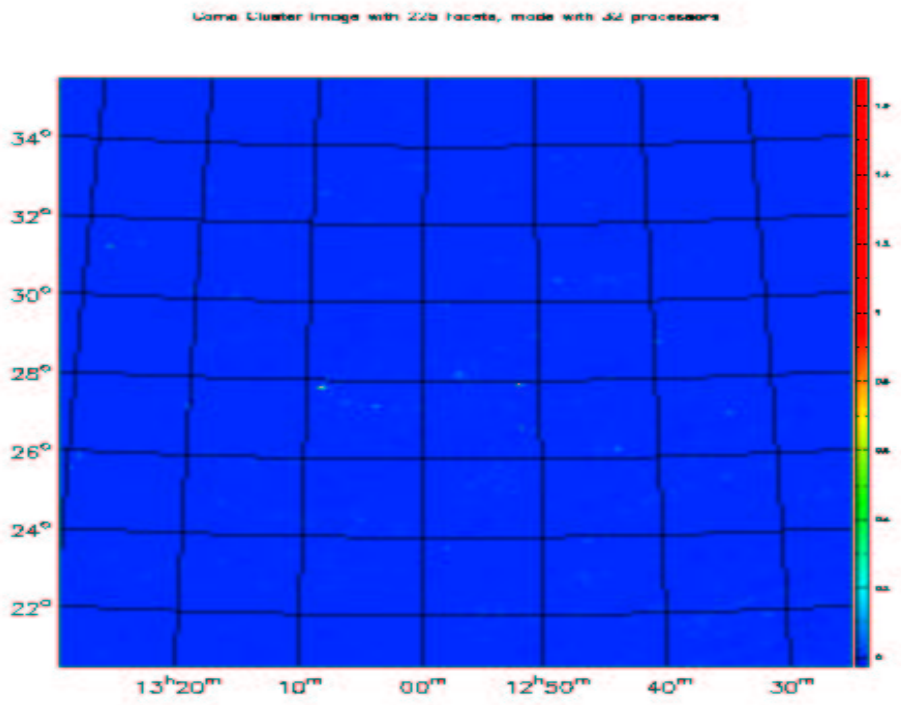

Figure 1: Image of Coma cluster at 74 MHz reduced in AIPS++, with 32 processors, using 225 imaging facets