Astronomical Data Analysis Software and Systems XXV
ASP Conference Series, Vol. 512
Nuria P. F. Lorente, Keith Shortridge, and Randall Wayth, eds.
© 2017 Astronomical Society of the Pacific

Early Science Pipelines for ASKAP

Matthew Whiting, Maxim Voronkov, and Daniel Mitchell
on behalf of the ASKAP team

CSIRO Astronomy & Space Science, Sydney, NSW, Australia;
`Matthew.Whiting@csiro.au`

Abstract. The Australian Square Kilometre Array Pathfinder (ASKAP) is an innovative wide-field, high-data-rate radio-synthesis telescope, that requires high-performance processing pipelines for calibration and imaging. These pipelines are being demonstrated with BETA, the 6-dish prototype Boolardy Engineering & Test Array.

1. The Demands of ASKAP

The Australian Square Kilometre Array Pathfinder (ASKAP, DeBoer et al. 2009) is a radio-synthesis telescope located at the Murchison Radio-astronomy Observatory (MRO) in Western Australia, featuring thirty-six 12m dishes equipped with Phased-Array Feed receivers[1]. These provide a wide, 30 square degree field-of-view by forming up to 36 separate dual-polarisation beams at once. This results in a high data rate: 70 TB of correlated visibilities in an 8 hour observation, requiring high-performance calibration & imaging pipelines.

The ASKAPsoft package has been developed specifically for meeting this challenging problem. Particular design decisions in the imaging algorithms, and the framework under which they run, have been made to meet processing requirements on memory and run-time, and scientific requirements such as dynamic range.

The ASKAPsoft software will run in pipelines on the Cray XC30 'Galaxy' supercomputer at the Pawsey Supercomputing Centre in Perth, Western Australia. Since ASKAP will be primarily a survey telescope, these pipelines must run in near-real-time, commencing once an observation is completed and completing in time to process the next. All calibration and imaging will be done in these pipelines, followed by the formation of astronomical catalogues, and the resulting data products (calibrated visibilities, images, cubes, and catalogues) will be stored in the CSIRO ASKAP Science Data Archive (CASDA, Chapman 2016).

2. Software pipelines for BETA

The Boolardy Engineering & Test Array (BETA, Hotan et al. 2014) is the prototype ASKAP array, consisting of 6 dishes equipped with Mk I Phased Array Feeds, capable

[1] At the 2015 CSIRO Awards Ceremony, held at the same time as this ADASS meeting, the ASKAP team was awarded the Chairman's Medal, CSIRO's highest award, for the development of the PAF receivers.

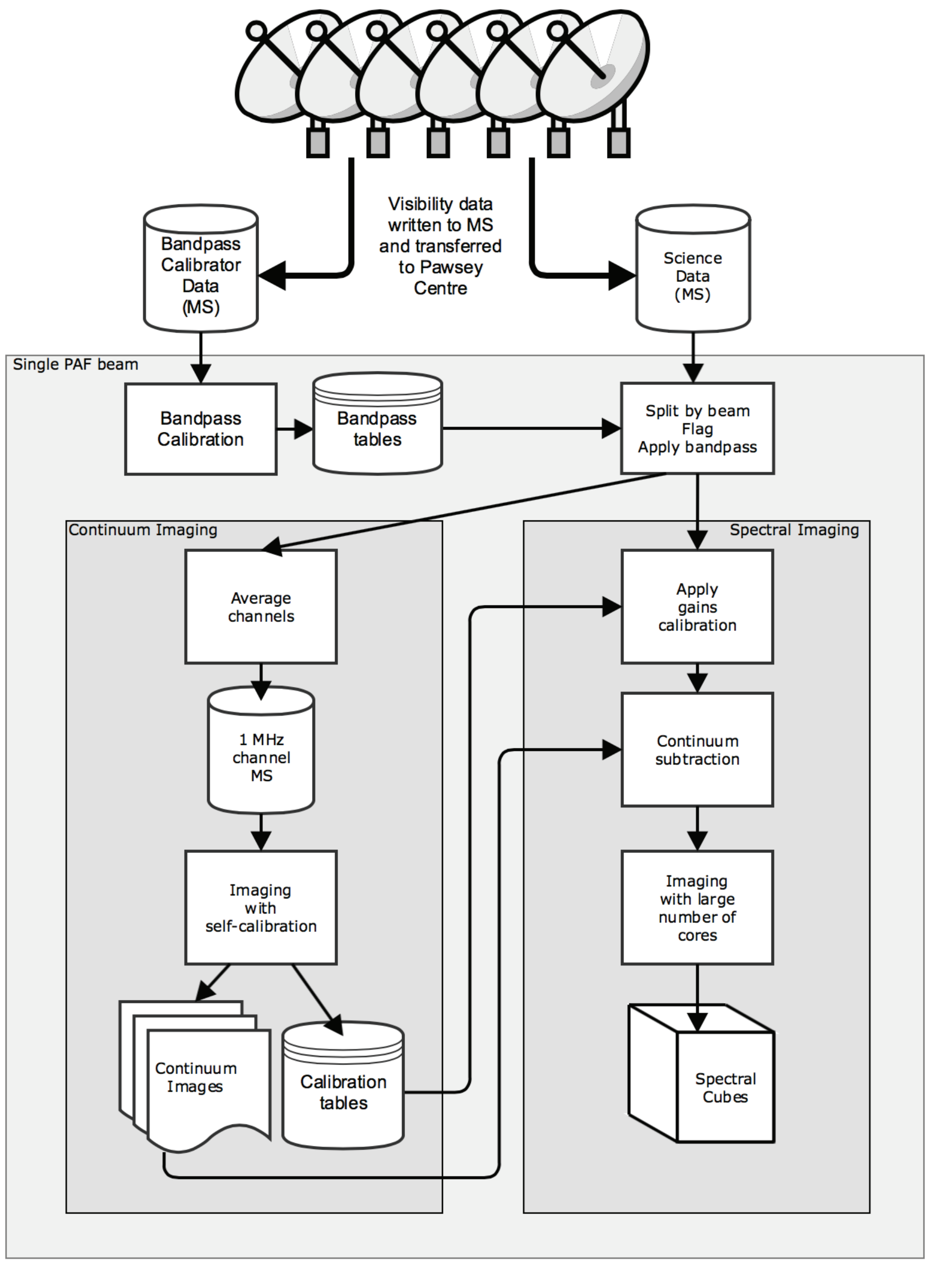

Figure 1. A portion of the BETA processing pipeline, showing the steps required for a single PAF beam. The telescope writes visibility data to a measurement set (MS) for offline processing at the Pawsey Centre. The bandpass is calibrated against an observation of PKS B1934−638, then a continuum and/or spectral-line pipeline is run to produce image data products. Other PAF beam images can be mosaicked together to form a full-field image. Source-finding can then run on the final image to produce catalogues.

of generating 9 dual-polarisation beams. BETA has been used for science observations (see, for instance, Serra et al. (2015) and Allison et al. (2015)), although its main purpose is to gain experience with the PAF receivers and to commission the telescope as a whole, including the processing software. The BETA datasets are small enough to not require real-time processing, and to allow processing by other, more established (if not high-performance) software packages. This allows us to validate the ASKAPsoft algorithms by comparison with well-understood workflows created by experienced astronomers.

We have developed a pipeline that runs on the 'Galaxy' supercomputer to allow processing of typical BETA observations by the ASKAP commissioning team. The pipeline structure reflects the typical observing strategy used for BETA: dedicated calibrator observations (where the southern flux-density standard PKS B1934−638 is observed in each beam for bandpass and flux calibration), followed by a 'target' or 'science' field observation. Separate PAF beams are calibrated and imaged independently, prior to mosaicking to form the final image. The key elements of the pipeline when processing a single PAF beam are shown in Fig. 1, while Fig. 2 shows the resulting continuum image from a single 9-beam BETA observation.

The pipeline software has been designed to hide much of the complexity of the ASKAPsoft tools from the user. The individual tools have a multitude of possible input parameters for configuration, but the pipeline software provides default values for many of these parameters that are appropriate for typical BETA observations, so that a typical pipeline run can be completely specified by a relatively small number of parameters.

The way the jobs are run on the supercomputer is also handled by the pipeline – when run, the pipeline launches a series of jobs that are queued for execution with appropriate dependencies. This removes the need for the user to have a detailed knowledge about how best to structure the large, distributed processing jobs. For those users that are interested, however, the pipeline provides information on the progress and performance of the various jobs (looking at metrics such as the time taken and the memory used).

3. From BETA to ASKAP

The next-generation MkII PAFs, that show improved sensitivity across the ASKAP band (Chippendale et al. 2015), are now being deployed on ASKAP antennas, and the first 12 such antennas will be used in 2016 to conduct the ASKAP Early Science program. The data for this program will be processed by ASKAPsoft pipelines that will build on those currently used for BETA, implementing improvements identified in commissioning and providing greater monitoring and control for the processing scientists. To prepare for such observations, we use of simulated 12-antenna observations, allowing testing of the imaging at a larger scale than is possible with BETA data.

As the ASKAP array grows, the increased sensitivity will mean a real-time calibration pipeline becomes more feasible, and the progressively larger data rates will eventually require a move to online (rather than offline, user-driven) processing.

Acknowledgments. We acknowledge the Wajarri Yamatji people as the traditional owners of the Observatory site. We would like to thank the ASKAP Commissioning & Early Science (ACES) team for their wonderful efforts in commissioning BETA, as well as our colleagues in the wider ASKAP computing team, past & present. This

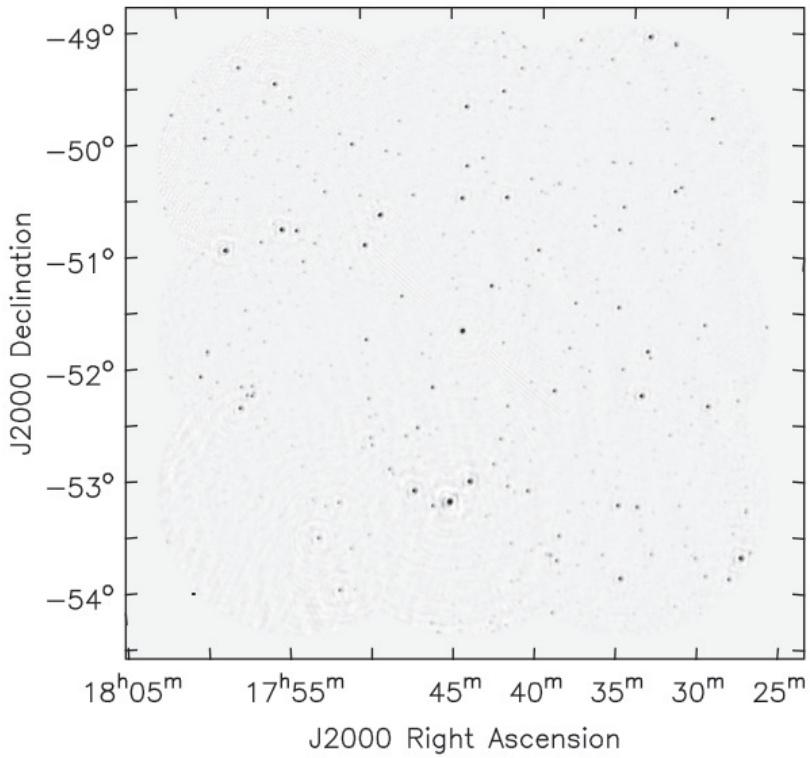

Figure 2. An example continuum image, as processed by the ASKAPsoft pipeline. This is the field surrounding PKS B1740−517, at a frequency of 863 MHz with a 304 MHz bandwidth: a total of 511 components above 5 mJy are detected. The entire field of ∼ 27.4 square degrees is imaged at once, using 9 PAF beams.

work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia.

References

Allison, J. R., et al. 2015, MNRAS, 453, 1249
Chapman, J. M. o. 2016, in ADASS XXV, edited by N. P. F. Lorente, K. Shortridge, & R. Wayth (San Francisco: ASP), vol. 512 of ASP Conf. Ser., 73
Chippendale, A. P., et al. 2015, in International Conference on Electromagnetics in Advanced Applications, 541
DeBoer, D., et al. 2009, Proceedings of the IEEE, 97, 1507
Hotan, A. W., et al. 2014, PASA, 31, e041
Serra, P., et al. 2015, MNRAS, 452, 2680