The relentless pace of Moore’s law has made it feasible and useful to port many scientific applications from dedicated hardware to commodity equipment. Radio interferometry employing the aperture-synthesis technique\(^1\) enables high-resolution imaging of the radio sky, despite the comparatively long wavelengths of radio emission. In this method, the coherence of the incident wavefront between pairs of radio antennas (baselines) is estimated by correlating a sampled representation of the electrical field at each antenna. The resulting ‘visibility’ data is subsequently inverted (using a Fourier transformation) to yield a distorted representation of the sky-brightness distribution. Deconvolution methods are then applied to obtain the best possible representation of the true sky-brightness distribution.\(^2\) Progressively higher resolution can be obtained by using more widely separated antennas. For instance, very-long-baseline interferometry (VLBI) uses antennas on different continents, and even in space, to reach baselines of thousands of kilometers and resolutions better than one milliarcsecond.

The ‘correlator’ used to calculate the baseline visibilities is a digital system in all modern interferometers, whose computation load is dominated by vector arithmetic: complex multiplication, addition, and Fourier transformations. The underlying algorithm is parallel, because the correlation operation can be divided into time and/or frequency. The desire to process wide bandwidths for high sensitivity has generally led to correlator implementations on custom high-speed digital signal-processing hardware, either application-specific integrated circuits (ASICs) or, more recently, field-programmable gate arrays (FPGAs). This approach is not without drawbacks: significant engineering effort is expended on each new design, and upgrades after design completion are typically costly or impractical. However, bandwidth increases for VLBI arrays have been slowing down because of the cost of data storage. Figure 1 shows the historical progression of maximum recording rates for VLBI.

Figure 1. Historical maximum recording rates for very-long-baseline interferometry (VLBI), with Moore’s law shown for comparison.

The relatively slow growth of computational requirements for VLBI correlation, coupled with the fact that VLBI arrays are typically small (5–20 antennas), has meant that custom hardware is no longer required for this small-\(N\) (i.e., low number of antennas), moderate-\(B\) (i.e., moderately wide bandwidth) region of parameter space. An alternative approach, ‘software’ correlation, uses commodity computing hardware and algorithms written in high-level programming languages such as C++. This allows correlators to be built quickly, since an intimate knowledge of the underlying hardware is not required. Because of the ease of programming, it also allows more features and flexibility than is typically feasible for a hardware correlator. The consumer trend toward multicore machines has been particularly useful for correlators, by virtue of the ease of parallelization.

However, potentially the most significant advantage for software correlators is their portability between generations of
commodity hardware and the commercial availability of optimized arithmetic libraries. The short test and development time needed for implementation of new software correlators allows commodity-based approaches to delay equipment purchases, thus taking advantage of ongoing improvements in commodity computing performance. This is a luxury not available to systems requiring significant (re-)development between architectures, such as graphics-processing units (GPUs) or FPGAs. Such systems, despite being largely developed in high-level ‘software,’ are thus not software correlators in the sense defined here.

The most widely used software correlator at present is the distributed-FX (DiFX, where FX represents the Fourier-transformation and multiplication ordered pathway) approach, written in C++. It uses a user-specified vector-arithmetic library (usually Intel® Performance Primitives) for the vast majority of computations and the message-passing interface for interprocess communications. The performance details and science results discussed here are derived from implementations of this correlator in US and Australian VLBI arrays, both of which use Linux as their operating system. In addition to DiFX, software correlation has been widely used in Japanese VLBI arrays and a distributed software correlator is also under development for the European VLBI network.

Performance statistics for DiFX at the Very Long Baseline Array (VLBA) are shown in Figure 2. The VLBA consists of ten antennas, and Figure 2 shows the ratio of correlate time to observe time for varying numbers of antennas, showing the VLBA’s current maximum bit rate (128MHz bandwidth). We ran these tests on a prototype cluster containing five rack-mounted nodes (with a height of one rack unit), each with four quad-core computer processing-units connected using dual gigabit Ethernet (total cost ~$30,000). Not shown here is the scaling with aggregate antenna bandwidth, which is almost perfectly linear. The addition of one or two nodes will allow this prototype system to replace the existing 18-year-old ASIC-based correlator, which occupies five large equipment racks. In addition to the obvious savings in space and power consumption, the DiFX correlator facilitates many advantages in flexibility over the existing ASIC correlator, such as arbitrarily fine frequency and time resolution for the output visibilities, thus allowing wide-field imaging.

The relentless pace of Moore’s law has made it feasible and useful to port many scientific applications from dedicated hardware to commodity equipment. Software correlators are just one such example, albeit particularly well suited because of the inherently parallel nature of the algorithm. In general, they allow more flexibility and more rapid development than traditional ASIC- or FPGA-based approaches, and they offer standardization between different VLBI arrays. Extrapolating into the future, software correlation will become cost effective for larger and larger arrays, and it may be adopted by substantial numbers of non-VLBI arrays.

Author Information

Adam Deller and Walter Brisken
National Radio Astronomy Observatory
Socorro, NM
http://www.aoc.nrao.edu/~adeller/

Adam Deller is a Jansky fellow. He studies radio sources at very high angular resolution using the VLBA. During his PhD at Swinburne University (Melbourne, Australia), he developed the DiFX software correlator for use with the Australian Long Baseline Array, which is now being adopted by the VLBA.

Figure 2. Benchmarks of the distributed-FX (DiFX) software correlator at the Very Long Baseline Array. The aggregate bandwidth was 128MHz for each antenna (data rate: 512Mbps/antenna). The correlation used a prototype cluster with five nodes, i.e., a total of 80 computer processing-units. The shaded area represents unreachable parameter space due to the limitations of the gigabit Ethernet interconnect from the data sources.
Walter Brisken is an associate scientist. His scientific interests focus on long-baseline observations of pulsars. He has been heavily involved in an ongoing sensitivity upgrade to the VLBA, which includes migration to the DiFX software correlator and wide-bandwidth data-acquisition systems.

References