

SCALABLE, EFFICIENT ASICS FOR THE SQUARE KILOMETRE ARRAY: FROM A/D CONVERSION TO CENTRAL CORRELATION

Martin L. Schmatz* Rik Jongerius† Gero Dittmann* Andreea Anghel*

Ton Engbersen* Jan van Lunteren* Peter Buchmann*

*IBM Research – Zurich, Switzerland †IBM Research, The Netherlands

mrt@zurich.ibm.com r.jongerius@nl.ibm.com {ged,aan,apj,jvl,pbu}@zurich.ibm.com

ABSTRACT

The Square Kilometre Array (SKA) is a future radio telescope, currently being designed by the worldwide radio-astronomy community. During the first of two construction phases, more than 250,000 antennas will be deployed, clustered in aperture-array stations. The antennas will generate 2.5 Pb/s of data, which needs to be processed in real time.

For the processing stages from A/D conversion to central correlation, we propose an ASIC solution using only three chip architectures. The architecture is scalable—additional chips support additional antennas or beams—and versatile—it can relocate its receiver band within a range of a few MHz up to 4 GHz. This flexibility makes it applicable to both SKA phases 1 and 2. The proposed chips implement an antenna and station processor for 289 antennas with a power consumption on the order of 600 W and a correlator, including corner turn, for 911 stations on the order of 90 kW.

Index Terms—Square Kilometre Array, Radio Astronomy, ASIC Design, Beamforming, Correlation

1. INTRODUCTION

SKA1-Low is an aperture-array instrument, built in phase 1 of the Square Kilometre Array (SKA), receiving signals in the band between 50 and 350 MHz. More than 250,000 antennas will be grouped in 911 stations and spread out over an area 200 km in diameter.

Signals received by the instrument are digitized and processed in the digital domain. Data is transformed into sky images or into time series for pulsar science. Figure 1 shows an overview of a potential processing pipeline for imaging modes.

After digitization, an optional beamformer creates coarse beams from a subset of the station's antennas. Coarse beams are channelized to increase the frequency resolution of the signal. An initial calibration step is performed before the final beams are formed. Furthermore, each pair of antennas is correlated such that calibration parameters can be determined for all subsets of antennas. Both beamforming steps delay and sum the antenna signals in order to change the directionality of the aperture array.

In the correlator, signal bands are aligned in time to correct for the geometric delay between stations. The aligned signals are correlated per pair of stations and integrated. The resulting *visibilities* are sent to the imager where sky images are generated.

Current telescopes implement these processing steps using various platforms. Aperture-array station processing is commonly implemented on FPGAs, whereas the central correlator algorithms are

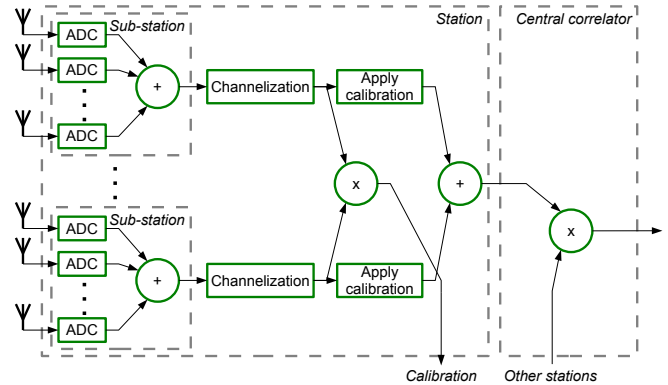


Fig. 1. Overview of the digital processing pipeline for aperture-array instruments up to and including correlation.

executed on a variety of platforms ranging from custom ASICs to general-purpose architectures such as GPUs or the IBM Blue Gene/P supercomputer [1, 2].

On the scale of the SKA, the energy consumption of compute platforms becomes a major issue. Anghel et al. [3] show that ASIC solutions can be up to four times more power-efficient than FPGA solutions. Therefore, in contrast to programmable platforms used for existing aperture-array telescopes, we propose an implementation for the SKA station and central processing using ASICs to reduce energy consumption.

The SKA precursor MeerKAT will use FPGAs for central correlation [4]. The use of ASICs, however, is not excluded where reconfigurability is not needed. We envision using our ASICs also where some reconfigurability is required due to, for example, changing specifications. By accounting for scalability the compute system can be adapted to the new environment without redesigning the ASICs.

D’Addario analyzes the power consumption of different architectures for correlator ASICs [5]. In contrast to some of these approaches, our architecture requires more memory, which we have available in the form of embedded DRAM. However, our design has the advantage that samples do not need to be re-ordered (corner-turned) between station processing and correlation, which would require one-to-all communication.

We focus here on implementing channelization, beamforming, and correlation functionality. We propose an ASIC solution to implement these aspects of the chain.

2. ASIC ARCHITECTURE

We propose an ASIC solution in 14 nm technology, which is expected to be available when SKA phase 1 construction starts in 2016,

This work is conducted in the context of the joint ASTRON and IBM DOME project and is funded by the Netherlands Organisation for Scientific Research (NWO), the Dutch Ministry of EL&I, and the Province of Drenthe.

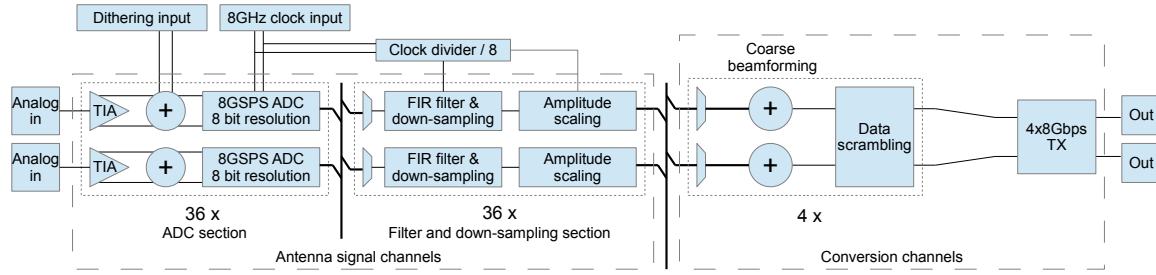


Fig. 2. Block diagram of the A/D converter ASIC.

for station processing and correlation in order to reduce the power consumption of the telescope. The ASICs are primarily designed for, but not restricted to, SKA1-Low.

The basic functionality, as shown in Figure 1, is implemented in three different ASIC designs. The first design integrates very small, low-power A/D converters (ADCs) [6] and coarse beamforming functionality, whereas the second design combines both channelization filters and station beamforming. The third design implements station and central correlation.

2.1. ADC chip

The ADC chip shown in Figure 2 consists of 36 dual-polarized antenna channels and four dual-polarized conversion channels.

The antenna channels are implemented in two sections, an ADC section and a filter and down-sampling section. The ADC section contains 36 identical blocks, each containing two amplifiers and two ADCs clocked at 8 GS/s, one for each polarization. The ADC resolution is 8 bits per sample, resulting in 128 Gb/s per block.

An FIR filter suppresses noise signals outside the bandwidth of interest. The passband is designed to be 400 MHz wide, the stopband starts at 600 MHz. For a 70-dB suppression, an FIR filter of order 119 is sufficient. By feeding the FIR filter with eight consecutive samples, the signal is down-sampled to 1 GS/s.

Next to stopband suppression, the FIR filter generates a fine time delay in order to align the incoming antenna signals for the coarse beamformer in each conversion channel. By using a 128-coefficient FIR filter with 128 different sets of coefficients, we can achieve a fine-grained time delay with approximately 1 ps resolution. Together with a 32-element delay chain, a time delay of up to 32 ns can be achieved.

Each conversion channel can optionally beamform (delay and sum) multiple antenna channels, selectable by a bus. A typical use case would be for the four conversion channels to beamform groups of nine antennas each. If not all antenna channels are used, e.g., if coarse beamforming is disabled, the unused channels are powered down.

The resulting output data rate of the chip, based on 16-bit beamformed samples, is 128 Gb/s for four conversion channels, for which 16 8-Gb/s I/O macros are placed on the ASIC. Furthermore, the chip contains logic for data scrambling, dithering, and control, which is beyond the scope of this paper.

2.2. Beamforming chip

Data generated by the ADC chips is processed further by the beamforming chips. One ASIC, shown in Figure 3, can handle up to four dual-polarized channels from the ADC chips and generates up to

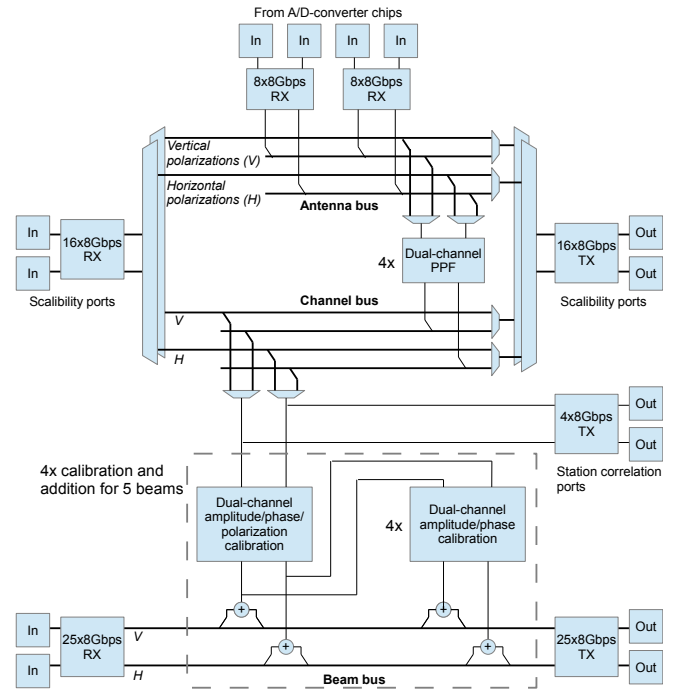


Fig. 3. Overview of the beamforming chip.

five calibrated station beams. Several scalability ports are available for scaling to higher numbers of antennas or beams.

By far the most demanding function of the beamforming chip is the channelizer, which is capable of generating channels with 1-kHz resolution. The channelizer is implemented by a polyphase filter (PPF) bank—FIR filters followed by a fast Fourier transform (FFT). We implement a 2^{20} -point FFT generating 2^{19} complex frequency channels from DC to 500 MHz.

The FIR filters are programmable with 4 to 16 filter taps each. In order to relax timing constraints, the PPF banks are clocked at 125 MHz. Given the sample rate of 1 GS/s, eight parallel FIR filters are implemented.

Out of the 2^{19} complex frequency channels, 327,680 channels are selected for beamforming, covering a bandwidth of 312.5 MHz. These channels can be located anywhere in the signal band between a few MHz and 400 MHz. In order to process signals between 400 MHz and the Nyquist frequency of 4 GHz of the 8-GS/s A/D converters, we use image rejection techniques. The ADC and beamforming chips are designed such that two channels combined can implement a Hartley image rejection mixer.

Table 1. Required memory for beamformer ASIC.

| Memory (per channel) | Sample size | Size |
|-------------------------------|------------------|--------|
| FIR taps (shared by channels) | 2 Byte | 32 MB |
| FIR samples | 2 Byte | 60 MB |
| FFT samples (double-buffered) | 8 Byte (complex) | 16 MB |
| FFT twiddle factors | 8 Byte (complex) | 4 MB |
| Calibration parameters | 8 Byte (complex) | 4 MB |
| Total (4 channels) | | 368 MB |

After channelization, antenna data is calibrated, delayed, and beamformed. In order to point the beam at the sky, antenna data is delayed depending on the beam direction. The calibration block applies a phase rotation to the antenna signals for inter-sample delays and compensates for gain differences between antenna channels. Coarse delay is performed using a time delay buffer.

The calibrated samples for all five beams are added to partial station beams received from other beamforming chips. Partially formed station beams are encoded using 64 bits for complex numbers, resulting in a data rate of 40 Gb/s per beam or five I/O macros. The resolution of the final beams transmitted to the correlator is reduced to 8 bit.

Station calibration parameters are calculated based on the correlation of all antenna signals. The station correlator does not need to correlate all antenna and frequency channels in real time. Time samples from all frequency channels are sent to the correlator in a time-interleaved fashion.

The channelizer requires data rates of more than one terabyte per second to memory. Reaching this memory bandwidth using external DRAM would be infeasible given the required number of memory lanes and power consumption, whereas embedded DRAM (eDRAM) technology is capable of delivering this bandwidth on chip.

Table 1 lists the amount of embedded DRAM required. The total memory size results in a large chip with plenty of chip edge where a large number of I/O macros can be placed. We make use of this by providing plenty of bandwidth in the scalability ports and in beam generation. Although not all of those ports will be used in SKA phase 1, they enable a gradual extension towards phase 2.

2.3. Correlator chip

Correlation is performed in two locations in the system: station beams are correlated for imaging, and individual antennas in a station are correlated for station calibration. The same correlator ASIC is used in both cases.

For correlation, 8-bit samples from each frequency channel are correlated with samples from the same frequency channel from a different antenna or station. Figure 4 shows the basic correlator unit cell that is instantiated many times on the chip. One pair of stations forms four correlations, which have to be integrated, employing four complex multipliers and four complex adders. Intermediate values are stored per frequency channel in eDRAM.

By connecting multiple chips as is shown in Figure 4, all correlations can be calculated. Note that there is no need for a corner turn: each station only connects to one input of the correlator.

In order to minimize chip-to-chip interconnect we place as many correlator unit cells on a single chip as possible. The size of the embedded DRAM and the number of I/O macros limits this. Only $1/5^{\text{th}}$, or 62.5 MHz, of the signal bandwidth is correlated per chip in order to reduce eDRAM size and I/O macros. To process the full signal

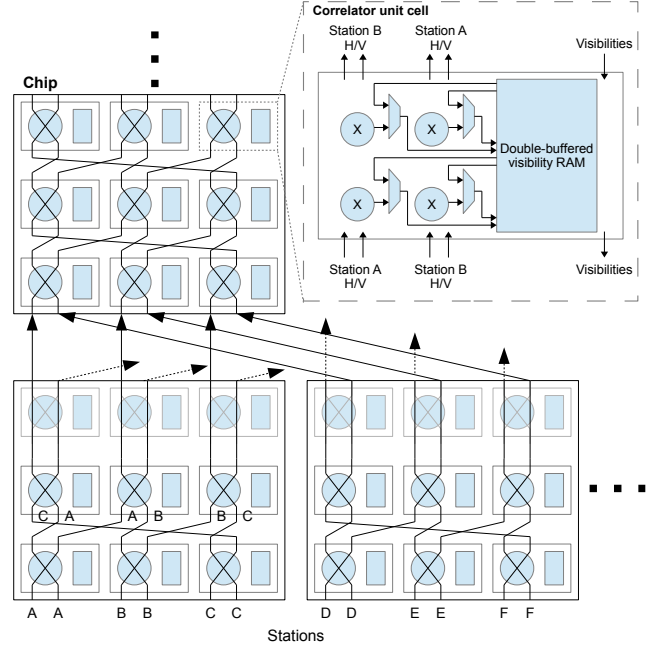


Fig. 4. A correlator chip contains several correlator unit cells, nine in this example. Multiple correlator chips are connected in a rectangular structure to calculate the full correlation matrix.

bandwidth, five correlator *frequency planes* are needed. As there are five macros per beam available on the beamformer ASIC, data can be routed to five different planes without additional hardware.

With five frequency planes, each correlator chip can contain 64 unit cells (nine are shown in Figure 4). Each unit cell contains 4 MB of double-buffered eDRAM to store the 8-byte complex visibilities. We use double-buffered RAM such that visibilities are integrated while the previous set is read out for further processing.

3. APPLICATION TO SKA1-LOW

The low-frequency aperture array stations in SKA phase 1 will consist of 911 stations with 289 antennas each [7]. A total of 73 A/D-converter chips are needed to convert the analog signals to the digital domain, assuming that no coarse beamforming is used. As each station produces one beam, each A/D-converter chip connects to one beamforming chip. Station correlation requires 703 chips to correlate all 289 antennas, only one frequency plane is used and the full bandwidth is processed in a time-interleaved fashion. A total of 849 chips are needed per station.

Correlation of all stations requires 6,612 chips per frequency plane. Five frequency planes require 33,060 chips for real-time correlation of all stations.

To estimate chip area and power consumption we assume that both are dominated by the A/D converters, the embedded DRAM, and multiply-accumulators for digital processing. Current state-of-the-art technology for these elements is scaled to 14-nm technology based on the ITRS road map [8]. Power is assumed to scale by a factor of 0.84 per technology node, area by a factor of 0.74.

A low-power and area-efficient A/D converter has been developed by Kull et al. [6]. The current 8.8-GS/s A/D converter in 32-nm technology uses a chip area of 0.025 mm² and consumes 49 mW. Scaling these numbers to 14 nm results in a chip area of 0.014 mm²

Table 2. Power and area estimates for MAC unit cells in 14 nm used by the ASICs. A correlator unit cell performs 16 MAC operations.

| MAC use | Clock | Power | Area |
|----------------------------|-------|---------|------------------------|
| ADC chip 8 b FIR filters | 1 ns | 0.42 mW | 792 μm^2 |
| BF chip 16 b FIR filters | 8 ns | 0.12 mW | 1,332 μm^2 |
| BF chip 32 b FFT butterfly | 8 ns | 1 mW | 11,211 μm^2 |
| BF chip 32 b calibration | 4 ns | 1.76 mW | 11,211 μm^2 |
| Correlator chip unit cell | 16 ns | 0.67 mW | 12,668 μm^2 |

and a power consumption of 34.6 mW.

Power and area consumption of embedded DRAM technology is estimated using CACTI [9]. A model is created for each DRAM element needed and the number of memory elements and access width is optimized to reduce power consumption and area. The model includes energy consumption per access, where we assume read and write energy to be equal, as well as leakage and refresh power.

Power and area estimation of the multiply-accumulate (MAC) unit cells is based on estimates from synthesis tools using standard threshold voltage (SVT) transistors in 22 nm. Scaled synthesis results are summarized in Table 2.

The ADC ASIC implements 72 antenna signal channels, each containing one 8 GS/s ADC. The ADCs require 1 mm² of chip area and consume up to 2.5 W.

The 128-tap FIR filter is clocked at 1 GHz and calculates a new output sample every cycle. As each tap in an FIR filter can be implemented using one MAC, 128 MAC unit cells are needed, which results in a compute rate of 128 GMAC/s. All 72 FIR filters add up to an area of 7.3 mm² and consume 3.9 W.

On the beamforming chip, eight 16-tap FIR filters implement one out of four polyphase filter banks. These filters operate on 125 MHz and compute 16 MACs each. The 128 MAC unit cells occupy 170,496 μm^2 and consume 15.4 mW.

One complex-to-complex radix-4 FFT is used per dual-polarized antenna channel to calculate two real-to-complex FFTs. Clocked at 125 MHz, 264 multiplications and 484 additions are needed per clock period. As we expect that energy and area usage are dominated by the multiply units, we assume that 264 MAC unit cells are implemented. Per dual-polarized antenna channel this results in an area of 3.0 mm² and a power consumption of 264 mW.

Calibration and beamforming require four complex MACs, or 16 real MACs, clocked at 250 MHz per dual-polarized antenna channel. The MAC units occupy 179,376 μm^2 and consume 28.1 mW.

The largest contributor to power consumption and die area of the beamformer ASIC is the 368 MB of integrated eDRAM. The aggregate memory bandwidth is 1.2 TB/s. Optimizing the different memory elements for power consumption and area results in 6 W power consumption including accesses, leakage, and refresh. For eDRAM, a total area of 191.2 mm² per chip is required.

As the MAC unit cells for all four antenna channels occupy 13.4 mm² and consume 1.3 W, the complete beamforming chip consumes 7.3 W and has an area of 204.5 mm², corresponding to 14.4 mm along each side.

Each correlator chip contains 64 unit cells, consuming 42.9 mW of power and occupying 0.8 mm² of chip area. A total of 256 MB of eDRAM consumes a further 2.7 W and 144.3 mm².

Correlation of antenna signals, used for calibration, does not have a real-time constraint and can be updated over the course of minutes. Given a 1 % active period of the station correlator and a 1 s integration time, the correlation tables are updated every eight minutes. The correlator is switched off for the remaining time.

Table 3. Power and area estimates based on embedded DRAM and MAC units for the three ASIC designs.

| ASIC | Single die | | SKA1-Low Power | |
|---------------|------------|-----------------------|----------------|------------|
| | Power | Area | Station | Instrument |
| ADC | 6.4 W | 8.3 mm ² | 51.9 W | 47.3 kW |
| Beamformer | 7.3 W | 204.5 mm ² | 530 W | 482.8 kW |
| Corr. station | 2.74 W | 145.1 mm ² | <20 W | <18 kW |
| Corr. central | | | - | 90 kW |

The total power consumption and area can be found in Table 3. Note that the power numbers for a single ADC chip assume using all 72 A/D converters and FIR filters in the chip, whereas the aggregate numbers for SKA1-Low are based on using only 8 A/D converters and FIR filters per chip. A complete system for SKA1-Low consumes 548 kW for station processing in all 911 stations and a 90 kW for the central correlator.

4. APPLICATION TO OTHER SKA INSTRUMENTS

The ASICs are not only designed to be applied to SKA1-Low. The capability to sample up to 4 GHz with the A/D converters, image rejection techniques, and the scalability ports make them amenable to other SKA1 instruments as well.

The SKA1-Survey instrument in Australia will consist of 96 dishes. They are fitted with phased-array feeds (PAFs), multiple antenna elements placed in the focal plane of the dish, and require processing similar to SKA1-Low. The instrument targets signals in three different 500 MHz bands between 350 MHz and 4 GHz. Thanks to the 8-GS/s A/D converters and the image-rejection mixing functionality, the presented design can be applied without modifications.

The phase-1 instruments will be extended in phase 2. It has yet to be defined how this will be achieved. Potential paths for expansion incorporate adding more dishes or antennas, or extending the SKA1-Low receiving bands to higher frequencies.

The scalability of our design makes it applicable to both SKA phase 1 and phase 2, or a gradual transition between both phases. Adding more antennas, beams, or stations is a matter of enabling more functionality or adding additional chips. For example, the number of beams of the SKA1-Low instrument can be extended by enabling the additional four available beams in the chips. Supporting more than five beams is simply a matter of adding more of the same chips: no new ASICs need to be designed.

5. CONCLUSIONS

We presented three ASIC designs for the Square Kilometre Array. The ASICs implement A/D conversion, filtering, channelization, and correlation and can be used, for example, for aperture array digital signal processing from A/D converter up to and including the correlator. The scalable and versatile design allows adaptation to changing requirements, other SKA phase 1 instruments, or phase 2 instruments without the need to redesign the ASICs.

The ASIC designs rely on low-power 8-GS/s A/D converters and embedded DRAM technology for a low-power design and to eliminate the need for high-bandwidth external DRAM. A complete system for SKA1-Low, the low-frequency aperture array for the SKA phase 1, consumes 548 kW for station processing in all 911 stations and an additional 90 kW for the central correlator.

6. REFERENCES

- [1] M. de Vos, A.W. Gunst, and R. Nijboer, “The LOFAR telescope: System architecture and signal processing,” *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1431–1437, 2009.
- [2] S. J. Tingay, R. Goeke, J. D. Bowman, et al., “The Murchison Widefield Array: The Square Kilometre Array precursor at low radio frequencies,” *Publications of the Astronomical Society of Australia*, vol. 30, January 2013.
- [3] A. Anghel, R. Jongerius, G. Dittmann, J. Weiss, and R. P. Luijten, “Holistic power analysis of implementation alternatives for a very large scale synthesis array with phased array stations,” in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014.
- [4] J.L. Jonas, “MeerKAT—the South African array with composite dishes and wide-band single pixel feeds,” *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1522–1530, 2009.
- [5] L.R. D’Addario, “Low-power architectures for large radio astronomy correlators,” in *General Assembly and Scientific Symposium, 2011 XXXth URSI*, 2011.
- [6] L. Kull, T. Toifl, M. Schmatz, et al., “A 35 mW 8 b 8.8 GS/s SAR ADC with low-power capacitive reference buffers in 32 nm digital SOI CMOS,” in *2013 Symposium on VLSI Circuits (VLSIC)*, 2013, pp. C260–C261.
- [7] P. E. Dewdney, W. Turner, R. Millenaar, et al., “SKA1 system baseline design,” Tech. Rep. SKA-TEL-SKO-DD-001, SKA, March 2013.
- [8] International Roadmap Committee, “International Technology Roadmap for Semiconductors,” <http://www.itrs.net/Links/2012ITRS/Home2012.htm>, 2012.
- [9] HP Laboratories, “CACTI,” <http://www.hpl.hp.com/research/cacti/>.