Melvyn Wright, ICCS workshop, Berkeley, Jan 2011

MOTIVATION

• Astronomers primarily interested in astronomy.
  – Data reduction preoccupies radio astronomy specialists.
  – Telescopes should be easily used by non specialists.

• Old paradigm: custom DSP + off-line calibration & imaging.
  – Mismatch between on-line and off-line processing bandwidths
  – Off-line can handle only a few percent of data rates from DSP.
  – Data processing problems for large-N arrays.
- PROBLEMS

• Data reduction bottleneck. Large time before analysis.

• User expertise with aperture synthesis data.

• RFI, bad data, flagging & editing.

• Background & confusing source subtraction.

• High data rate and Power management.

• Lost science opportunities without real time feedback.
• Heterogeneous DSP: FPGA, GPU & clusters for flexible programming.

• Image large FOV with high frequency and time resolution.

• Simultaneously image: science targets, calibrators & confusing sources.

• Variable sources and RFI handled in real time.

• Real time feedback into imaging and deconvolution.
Simultaneous images of multiple regions.

Image fidelity $\sim 10^4$

Bandwidth $\sim 4 - 40$ GHz.

Spectral channels per band $\sim 10^5$.

Sample and integration interval $\sim 1$ ms - 10s.

Image size $\sim 10^6$ pixels.
• Calibration and editing in close to real time using a sky model.

• Calibration parameters fed back into imagers and beam formers.

• Subtract sky model from $uv$ data before imaging.

• Observations update and improve the a-priori model.

• Sky model is final calibrated image when observations are complete.
• High dynamic range need to subtract strong sources from the data.

• Antenna station beam pattern is time variable.

• Deconvolving large FOV is very expensive off-line

• Sky model used to calibrate $uv$ data & subtract confusing sources.
• Multiple delay & phase centers for targets over a wide FOV.

• High performance DSP handle high data rates in parallel.

• RFI mitigation with high time and frequency resolution.

• Calibration and feedback into beam formers in real time.

• Image & deconvolve in close to real time.
Figure 2: Data flow from telescopes to images.
Hierarchical beam formation & correlators image large FOV.

Phased array beams can be formed anywhere in the sky.

Correlators image multiple regions in a large FOV.

Beam formers channel collecting area into expensive back ends.

Correlator and Beam formers are shared DSP resources.
ADVANTAGES OF USING SWITCHES

- COTS rather than custom backplanes for large $N_{ant}$.
- Packets can be routed to multiple asynchronous DSP engines.
- Include metadata needed to calibrate & image multiple regions.
- Flexible routing to beam formers & correlators.
- DSP can be upgraded & reprogrammed with minimum interruption.
- EXAMPLES -

• ATA: 4 32 ant x 4 pol x 100 MHz 1000 channel correlators (FPGA)
• ATA: 3 42 ant x 2 pol x 100 MHz beam formers (CASPER FPGA)
• VLBI: 8 ant x 1 pol x 1 GHz beam former (CASPER FPGA)
• PAPER: 32 ant x 4 pol x 100 MHz correlator (CASPER FPGA)
• CARMA: 23 ant x 1 pol x 4 GHz correlator (custom FPGA)
• ALMA: 50 ant x 4 polys x 8 GHz correlator (custom ASIC)
DATA RATE MANAGEMENT

- Input bandwidth $N_{ant} \times B \times N_{pol}$.
  - digitize & packetize using COTS hardware and protocols.

- Data bandwidth $N_{ant} \times 2B \times N_{pol} \times N_{bits}$,
  $$4 \times 10^{12} (N_{ant}/1000)(B/\text{GHz})(N_{pol}/2) \text{ bytes/s}$$
  using 8-bit digitization.
• Spectral resolution, RFI rejection, multifrequency synthesis.

• Science & RFI require large $N_{chan}$ — favors FX architecture.
  
  – polyphase filter isolates frequency channels. e.g. RFI.

• Data bandwidth $N_{ant} \times 2B/N_{chan} \times N_{chan} \times N_{pol} \times N_{bits}$.
  
  – reduce bit width after RFI rejection

• Parallel processing reduces data rate by $N_{chan}$
- CORRELATORS

- Cross correlation of the sampled wavefront between pairs of antennas.

- Data bandwidth from correlator for full FOV:

\[ N_{ant}(N_{ant} + 1)/2 \times N_{pol} \times N_{chan} \times N_{bits} \times 2 \times sdot \times D_{max}/\lambda \]

\[ \sim 10^8(N_{ant}/1000)^2(N_{pol}/4)(D_{max}/km)(\lambda/cm)^{-1} \text{ bytes/s} \]

using 2 × 16 bits per complex channel.
• Average to reduce data rate and FOV at each phase center.

• Data bandwidth for antenna primary beam is:

\[
N_{\text{ant}}(N_{\text{ant}} - 1)/2 \times N_{\text{pol}} \times N_{\text{chan}} \times N_{\text{bits}} \times 2 \ sdot \times D_{\text{max}}/D_{\text{ant}} \\
\sim 10^5(N_{\text{ant}}/1000)^2(N_{\text{pol}}/4)(D_{\text{max}}/\text{km})(D_{\text{ant}}/12\text{m})^{-1} \ \text{bytes/s.}
\]

using 2 × 16 bits per complex channel.
MIRIAD Timings for 1573x1573x10 channel Imaging
(simulated ALMA data with 60 antennas in 4 km configuration)

1535MHz Athlon XP 1800+ w/256KB cache, 512MB RAM. 10Feb2002.

Calibration – solid line:
8812 uvpoints/sec (1.1 Mbytes/sec)

Gridding – dashed line:
11400 uvpoints/sec (1.3 Mbytes/sec)

FFT: 1573x1573 x 10 channels in 10 sec.
- CALIBRATION

- Station Beam, Gain, Bandpass, & Polarization.
  - primary beam is complex product of station voltage patterns.
  - time variable PB, pointing & atmospheric fluctuations.

- Non isoplanicity – calibrate the data for each phase center.

- Image science targets, calibrators, & confusing sources.
  - identify confusing sources from a-priori images.
RFI MITIGATION

• Subtract RFI before averaging in time and frequency.
  – Gaussian filtering: MAD, Spectral Kurtosis

• Characterize RFI as a function of time, frequency and polarization.
  – SNR improved by pointing some antennas at RFI sources.
  – may need fast sampled $uv$ data (e.g. satellites)

• Null formation by controlling station beam.
Figure 4: ATA data before RFI rejection

Pre-flagged Data

Amplitude

Frequency

Time
Figure 5: Post-flagged Data
Nulling WAAS on Galaxy 15

1 part in $10^8$ of transmission

**On-Axis:**

**Off-Axis:**
10 dB down

**Nulled:**
28 dB down

Figure 6: Null Formation on interfering signal.
• Image parallel $uv$ data streams with a-priori model subtracted.
  
  – difference images update the model & improve calibration.

• Deconvolve by subtracting model & sidelobes of confusing sources.
  
  – stop when model image is consistent with $uv$ data streams.

• Transient source are inconsistent with the model. Save this $uv$ data.
  
  – make and keep $\chi^2$ image to identify transient & RFI sources.
• Imaging is a dynamic process.
  – look at convergence of sky model and $\chi^2$ image.

• Phase centers can be moved for science goals, or better calibration.
  – new sources are discovered in the imaging process.
  – isoplanatic patches may vary during observations
  – calibration across the FoV.
• High bandwidth archive is an integral part of real-time system.
  – save $uv$ data streams with metadata.
  – $uv$ data can be replayed though imaging system.

• Better sky model used to improve calibration.

• Save transient source data for further analysis.
• Delayed calibration & analysis limit the science which can be done.

• Real Time Imaging reduces burden of data reduction on users.
  – expertise which many users do not have.

• Bandwidth Management at every stage of data processing.

• Make best use of both telescope and human resources.