ALFA Pulsar Consortium Workshop: Summary & Action Items

Shami Chatterjee

Arecibo, 1–2 Nov 2002

Minor update, 12 Dec 2002

Welcome & Introduction

To Do list (Chris Salter)

1. Determine details of Consortium membership.
2. Decide consortium constitution and structure.
   Who will implement?
   Representative(s) to deal with other ALFA consortia.
3. Consortium proposal to NAIC: route and timescale.
4. Identify an NAIC “Point of Contact”.
5. Technical and other requirements: (begin to) tell NAIC
6. Consider details of student participation: how?
7. Funding ...?

Hardware: Front End

(German Cortés)

- 7 stepped TE11 mode horns;
  Aperture = 25 cm; center spacing = 26 cm.

- $768 \times 658''$ on the sky.
  Central: $230 \times 201''$; Avg pixel: $231 \times 204''$

- 1225—1525 MHz;
  Gain 11 K/Jy (center), 8.5–9 K/Jy (other).
  Variation of +1/–2 dB over band.

- T_{sys}: 35 K (1.2 GHz) to 25 K (1.5 GHz).
  T_{scatt} = 4 K; $T_{receiver}$ = 8 K;

- Linear polarizations: Cross pol –28 dB for central pixel;
  –25 dB to –21 dB for peripheral pixels.

Hardware: Front End (cont.)

Delivery: 2004 April.
Commissioning: $\sim$ 6 months.
Surveys could begin 2004 Nov.

Receivers: Which stay up, which removed? L-wide stays...?

Consider an 8th beam for RFI mitigation?

Hardware: IF/LO

(Lisa Wray / Eddie Castro)

- 1225–1525 MHz downconverted to 150–450 MHz.
- 14 IF systems for 7 feeds (only 1 LO)
- WAPP backends: what is the maximum number of backends?

Note: Strong RFI from GPS: 24 MHz at 1175 MHz
Is this cut off by the waveguide?
No, waveguide cutoff is 1125 MHz

Hardware: Backends

4 possibilities considered:

- Enhanced WAPPs
- ATA filter board (“F”-card) based design
- FPGA-based modular digital backends
- Good old Filterbanks

Backends: Enhanced WAPP

(Hagen / Sisk)

Basic design builds on WAPPs:

- 2 correlator boards (1 IF each) = 2 $\times$ 100 MHz; 32 bit lags.
- Alt: Digital filter board: direct sampling.
- Either one feeds a new lag demultiplexer card: 16 or 32 bit; 45 Mb/s

Develop 7 such backends: *hopefully* by ALFA deadline.

3 $\times$ 100 MHz boards? Data rates and manpower issues.

Can a new correlator board be dropped in? NOT easily.

Backends: ATA-based Filter Board

(Backer)

Polyphase filterbank (PFB): FFTs with weights.
$\Rightarrow$ Window function is better behaved than sin(x)/x.

ATA “F-card” based on Xylinx Virtex II board
($800 each; some for free?)

What is needed?

- Input A/D bytes, 100 MHz, double sideband.
- Channels in FFT/PFB: determine with FPGAs.
- Detection & decimation of power vector.
- Power data samples I/O to host CPU (50 MB/s).
- Voltage I/O $\rightarrow$ PC array.

Flexible PC array: real time search, RFI mitigation, timing?

Backends: Modular Reconfigurable Digital Backends

(Ellingson / Hampson)

FPGA based spectrometer: reconfigurable, reuse modules, test as you go, incremental upgrades.

Currently: 80 MHz bandwidth $\rightarrow$ ADC $\rightarrow$ ... $\rightarrow$ 16+16 bits.

Asynchronous Pulse Blanker for RFI mitigation:

- Running estimate of μ, σ.
- If (sample $> \mu + \beta \times \sigma$) then blank.
- Blanking *both before and after* trigger.
- Need more Blanking Time Registers (only 1 so far).

Note, also need frequency domain blanker (after time blanking: impulsive signals cause power leakage in FFTs).

Contact hampson.8@osu.edu for documentation.

Backends: Good Old Filterbanks

(Dunc Lorimer)

Challenge: can you beat this?

Pros:

- Cheap, no R&D, proven track record (Parkes).
- 1 bit: optimal storage; RFI-robust.

Cons:

- It’s soooooo 20th century!
- 20% sensitivity compromise vs. many bit sampling.
- Inflexible: may not work for other piggyback surveys.

Backends: Good Old Filterbanks (cont.)

Circuit diagram:
`http://www.jb.man.ac.uk/~tsi/projects/arecibo`

To commission: 70 weeks; $110,000 + $84,000 overhead.
(Alternative scenarios available.)

Other Issues:

- 600 μs smearing for DM=200, 1.4 GHz, 1 MHz channels.
- High pass filter: suppresses long period pulsars.
  e.g. PMB: longest period = 6.7 s (at 3.35 s harmonic).

Backends: Summary

Goal should be most efficient use of telescope: *observe as much sky as possible, as deeply as possible*.

$\rightarrow$ Flexibility is very important.

$\rightarrow$ Proven track record and well-defined costs also important.

To do: Detailed cost and time estimates.

To do: Other mystery designs?

Surveys: Lessons from Parkes

(Ingrid Stairs)

Parkes had: 13 beams, average T_{sys} = 21 K;
14′ beams spaced by 2 $\times$ FWHM;
Filterbank: 96 channels, 288 MHz.

Main survey ($260^\circ \leq l \leq 50^\circ, |b| < 5^\circ, T_{int} = 2100\ s$)
$\rightarrow$ 600+ pulsars, 9 binaries, 4 MSPs; to reprocess.

Flanking survey ($260^\circ \leq l \leq 50^\circ, 5 < |b| < 15^\circ, T_{int} = 256\ s$)
$\rightarrow$ 69 pulsars, 8 recycled + binaries.

Collaboration: $\sim$dozen people; centralized decision making.
All team on discovery papers; not all on follow ups.
Group votes on external collaborations.

Surveys: Lesons from Parkes (cont.)

- Expect some reprocessing after things are running.
- Do not underestimate time for confirmation, follow up.
- Do not underestimate manpower, processing load, RFI. (Parkes: Self interference was major – NAIC role?)
- Clear organization, roles for young scientists / students. (see astro-ph/0105351)
- Processing allocation?
- Parkes data are “public” but not practically available. Archive? Role for the NVO?

Pilot Survey: 1 square degree

(M. McLaughlin)

Pilot 1 square degree search: 793 beams.

- L-wide recvr + WAPP: 1475 MHz, 100 MHz, 430 s.
- 256 chans, 102.4 μs, 2^{22} point FFT.
- 3 level, 16 bit $\rightarrow$ 2 GB files.

- 0.05 mJy @ 0 DM; Observed / Expected S/N = 0.98.

- 10–20 candidates per pointing; expect 3–4 pulsars.

Should surveys go to $|b| = 5^{\circ}$, or less?
Based on Parkes, maybe not. But too early to say.

Should τ_{int} vary based on $T_{sys} = T_{sys}(l, b)$? Maybe not.

Strawperson ALFA Surveys

(Jim Cordes)

Boundary conditions:

- Future: Full Galactic Census (LOFAR, SKA).
- New NAIC mode of ops.
- RFI mitigation – general tools.

Ideas:

- Gal Plane $|b| < 5^\circ$, 300 s, 2–3000 hrs.
- Intermediate $|b| < 15 - 25^\circ$?
  (400 MHz better above10°.)
- Deep targeted surveys (MB for RFI).
- Extragalactic: M33.
- Pulsar / transients piggyback on high $|b|$ HI survey.

Strawperson ALFA Surveys

- Need 2 or more passes on sky? (Nulling, ISS)
- Search vs Confirmation obs: 2 searches?

To do: Survey simulations

To do: RFI study

To do: Merge search and simulation code?

To do: Telescope survey modes & interlacing

RFI: The Bad News

(Ramesh Bhat, Murray Lewis)

Pulsar & Spectral Line perspectives: rather not have RFI.

WAPP: 1125–1225 and 1425–1525 MHz used: okay.
ALFA: 1225–1525 MHz: lower 200 MHz is awful!

- San Juan Airport radar: 1330, 1350 MHz – bad.
- GPS L5:1175 MHz; L3:1380 MHz: 600+ hrs/yr.
- Internal RFI is critical as well.

$\sim 10\%$ data loss; $\sim 5\%$ with radar blanking.

RFI: The Even Worse News

- RFI that saturates IF/LO cannot be mitigated!
- Integrated plots are very misleading:
  Between spikes, band can be clean.
  (e.g. Arecibo radar band)
  “Clean” band can have broad band low level RFI.
  (even worse!)
- SETI: Avoid stacking RFI excision algorithms!

Pilot studies:
Dual site (AO,GBT); Dual beam (Gregorian, CH).

Piggyback Surveys: Galactic Science with GALFA

(Carl Heiles)

HI maps: supernova remnants, supershells, loops, etc.
Galactic fountain, bipolar outflows from stars.

Continuum: Polarization $\sqrt{Q^2 + U^2}$; Cold neutral medium.

Need / want:

- Spectral line mode: 6.25 MHz, 8192 chan, T_{int} ~10 min.
- Faraday rotn obs: need full Stokes.

Will need lots of time: no reason not to use same pointings.
Can coexist, IF observing pattern is set for baseline subtraction (observations with the same part of the dish).

Piggyback Surveys: Galactic HI

(Tom Bania)

Example: BU AO Survey (1983–1990: long time!)
21 cm self absorption (HI absorbing against background HI).

Distances are fundamental: Galactic rotation ambiguity.
Combine optical/IR extinction: match ^{13}CO survey features.

Need / want:

- 5 MHz BW, ± 500 km/s $\Delta V_{LSR}, 7 \times 10^4$ channels.
- Coexist with pulsar surveys? Yes.

Consortium of consortia: best possible case for survey?

Piggyback Surveys: Extragalactic HI

(Riccardo Giovanelli)

Faint end of HI mass function (HIM).
Local density dependence of HIM.
Map luminous and dark matter in local ($z < 0.1$) Universe.
Nearby galaxy groups, Local Group, HVCs.

Need / want: several choices of parameters.

- $D \propto T_{int}^{1/4}$; Volume $V \propto \Omega D^3$
  $\Rightarrow V \propto \Omega t^{3/4}$: survey large angles quickly.
- Fast, all-sky survey (5 s, 50 MHz, 25 kHz) $\rightarrow$ 1100 hrs.
- Virgo: 120 MHz, 25 kHz, 60 s $\rightarrow$ 800 hrs.
- Zone of Avoidance: 120 MHz, 25 kHz, 300 s, $|b| < 10^\circ$
  $\rightarrow$ 10,000 hrs.
- All have synergies with other surveys.

Piggyback Surveys: SETI

(Dan Werthimer)

Searches have improved by huge amounts
(*but telescopes keep collapsing mysteriously...*)

Various time and frequency scales (e.g. SERENDIP IV on CH: 100 MHz, 16.8 million channels) – over 10^{18} fruitless tries.

Also search SETI@home data:
e.g. for μs pulses at DM +100 to –100.

Need / want:

- Piggyback SERENDIP V:
  300 MHz, 2 pols, 7 beams, 5 billion chans.
- Lot of parameter space to cover.
- Look at all 4 Stokes?

Software: An End-to-end Approach

(Dunc Lorimer)

Lots of people have their own search code.
e.g. `http://www.jb.man.ac.uk/~drl/seek`

A standard pulsar search package? ($\equiv$ Tempo?)

- Scheduling
- Book keeping (cooperate with telescope control)
- Storage format: homebrew binary? FITS?
- RFI excision: time, frequency, cross correlation.
- Search: mostly automated?
- Digest: are they real?
- Candidate dbase: where and how?
- Confirmation: future & past data.

Software: Things to Consider

- Balance rewriting & recycling (e.g. `netlib.org`).

- Well-written and well-documented code.

Sanity checking and Quality control? Code check-in?

CVS (Parkes: enforce locking, or else!) vs. software by committee?

- A modular approach with UNIX pipes?

What modules? What is in realtime?

- Databases & scripting languages vs flat text files.

Example: `http://boinc.berkeley.edu`

- Portability of code: multi-platform compiling.
- Pipeline vs. cross-beam RFI analysis.
- Generality: maybe even X-ray searches?

To do: Come up with list / description of available code.

Computing Resources

(Vicky Kaspi)

What's out there?

- McGill: "The Borg"
  52 dual Athlon 1.4 GHz nodes, lots of RAM, disk.
  $200K – power, wiring upgrade, air conditioning.
- Jodrell: "COBRA"
  91 dual PIII – lower power; RAM, disk.
  Very fast switch: 150 MB/s node-to-node.
  $500K including switch.

(We wouldn't need fast switch.)

Computing Resources: Processing Needs

- MAM: 1.4 GHz Athlon + 1 GB RAM: typical beam $\sim$24 hrs.

At that rate, 300,000 CPU days – 8 years on Borg.

- Acceleration searches: huge CPU requirements.

- Save dedispersed time series... bulk of CPU time.

- Moore's Law; problem is inherently parallel.

Speed vs power consumption.

Computing Resources: Processing Needs

- 300,000 beams: $\sim$800 TB+ overhead $\Rightarrow$ 1 Petabyte.

- What data should be archived?
RFI cleaned, dedispersed time series / raw data / both?

- Who archives? Consortium? Observatory?

- Technology: tapes not looking good; disk arrays?
New tech on the horizon? Mass-market technology?

- Plug into HEASRC, NVO?
Most science interest from broadest use.

Computing Resources: Arecibo Hardware & Network

(Arun)

- 28 Gigabit ethernet port switch.
- OC-3 lines; links to Internet, Internet 2.
- Near future: 155 Mbit/sec link off the island.
  Mainland: 2 Gb/s backbone.
- A shared storage network:
  Fed by all backends, serves all access requests?
  Each consortium member contributes storage...?
- CPU: distributed computing?
  Large computers: shielding?

Organizing a Group Effort: Who does what?

Based on preliminary guidelines (Salter).

- Working groups – open membership.
- NAIC tasks (Well defined?)

  - Front end (horns, IF/LO) - April 2004.
  - Back ends: 7 WAPPs - done by late 2003.
  - Telescope monitor & control software.
  - Software for NAIC backends.
  - Calibration, testing, maintenance & repair.
  - Data archiving.

Organizing a Group Effort: Who does what? (cont.)

- NAIC point of contact
- Consortium meetings: hosted by NAIC
- Consortium should return:
  - Constitution, membership rules.
  - Science goals; scope.
  - Data acquisition, time estimate, processing, software.
  - Funding needs, PI, volunteers for pre-commissioning, etc.
  - Ongoing: progress reports, data & data products release.
  - Ongoing: software release.

Organizing a Group Effort: Who does what? (cont.)

Misc. thoughts (Desh):

- Optimal use of telescope will require alt backends.
- Depth of survey vs survey time.
- Parallel / multiple search processing strategies?
- Documentation ...?
- Other data products: TOAs? Dyn spec?

Organizing a Group Effort: Consortium

Organization chart: people on subcommittees – by function.

What about decision making?

Coordinating committee = leaders of Subcommittees & Secretary & NAIC member?

To do: Populate subcommittees.
Volunteers? (Preliminary and NOT exclusive!)

Consortium Committees

- Survey: *Cordes*, Weisberg, Nice, McLaughlin, Ransom, Kaspi, Camilo, Arzoumanian, etc.

- Data Acquisition: *Freire*, Stairs, Jason, Ramachandran, Bhat, Backer, Jin Lin Han, etc.

- Processing: *Lorimer*, Cordes, McLaughlin, Ransom, Kaspi, Nice, Bhat, Ramachandran, Stairs, Han, etc.

- Data Management: *Ransom*, Kaspi, Stappers, Lorimer, Stairs, Chatterjee, Cordes, Hessels, Xilouris, etc.

- Follow up: *Lommen* (Westerbork), Kramer (JBO, Bonn), Gupta (GMRT), Backer (ATA), Stairs (GBT etc), Freire (Arecibo), Gaensler, Stappers, Thorsett (GLAST), Cognard (Nancay), Han, Desh (Ooty) etc etc.

- Coordinating committee: Group chairs + NAIC (Desh/Freire)

- Acting Chair: *Cordes*

- To do: Subcommittee Temporary Chairs: write description of the committee job, solicit members, elect chair.

Consortium Report & Discussion

Not quite a proposal, but a list of goals and requirements.

Todo: Report by 6 December 2002, 10 pages.

- Strawperson survey requirements: 1024 channels, 300 MHz...
- Computers & Procesing: Berkeley, Jodrell, UBC, Astron, McGill, NRL, Cornell, UPR.
- Archiving: NAIC?
- Time fraction to ALFA vs other projects? We think...?
- What is the scope of the consortium vs other efforts? (e.g. Globular clusters; GLAST tie-ins.)

Consortium Report: cont.

- NAIC requirements.
(e.g. RFI mitigation, archiving, hardware, spectrometers)

- Cost and Timelines.

- Instrument Maintenance: consortium vs NAIC.

- Processing power on site: how much extra do we need?

- Observing concept: remote by consortium members?
Need continued Internet-2 access.

- Other stuff: protocols, proprietary periods (18 months, but open to proposals?), analysis, publication policy – internal proposals?

- Student projects? New members?

Funding: Canadian Foundation for Innovation

(Vicky)

CFI: `http://www.innovation.ca`
Will fund up to 40% of project cost for projects $> \$150k$.

Interest in high performance computing, databases.
Possible role for offline processing, archiving, etc.

Deadlines: Notification of intent Dec 18;
Proposals due May 2003 from institutions;
Decision Feb 2004.

Funding: NSF et al.

(Jim Cordes)

Preliminary proposals (Cordes, Camilo – devel work): Okay.

Ask for more? For what purposes?

- Hardware? Spectrometers? 9–15 months lead time.
  7 WAPPs = 7 $\times$ \$30K less what's built.
  Filterbanks = \$110K + overhead.
  "F-card" = $\sim$\$100K...?
- Grants from NAIC: NAIC should be aggressive.
- Explore instrumentation grants at Jodrell.
- Consider Co-PIs for each country funding.

Future Meetings

Coordinate with other groups. Milestones:

- Extragalactic meeting: March 2003.
- GALFA meeting $\sim$May 2003?
- IAU Sydney: July 2003.
- 1 year: Have some software ready?

Jan 2004 – signal path should be ready (with fake front ends).

Action Items

1) Subcommittee temporary chairs: describe committee job, call for and organize membership, elect a chair.

2) Report by 6 Dec 2002: Cordes, Nice, Chatterjee, Stairs, Ransom, McLaughlin, Freire.

3) Explore various funding options: each country.

This document is available online:

`http://www.astro.cornell.edu/~shami/share/ALFA-nov1.pdf`