

ECSV Discussion

24 April 2012, at 10am in room 317.

Attendees:

Andreas Brunthaler, Bryan Butler, Claire Chandler, Vivek Dhawan, Eric Greisen, Amanda Kepley, Casey Law, Josh Marvil, Drew Medlin, Heidi Medlin, Emmanuel Momjian, Steve Myers, Kristina Nyland, Frazer Owen, Jon Romney, Nirupam Roy, Michael Rupen, Bob Sault, Lorant Sjouwerman, Philip Schmidt, Deb Shepherd, Gustaaf van Moorsel,

Minutes:

Correlator and general system health (Michael)

- Stokes V problem and frequency setup issues have been resolved. Remaining issues are at such low levels that it is below the noise ($\sim 10^{-3}$) errors. This means we can make Zeeman observations: the first Zeeman science observations were last night (Emmanuel saw the line).
- In one high red-shift science program, the user claimed that the observed JVA frequencies did not agree with PdBI – We tested this and there is no current problem in the correlator. Gustaaf is looking at the data to see if there was an error in post processing.
- Vivek has uncovered band pass variations that change on time scales of minutes. Levels are at a few percent in amplitude and a degree or so in phase. All bands except X-band appear to have this problem. The problem is intermittent on a particular antenna but it is always there in a given dataset where we are observing strong calibrators. This did not exist in a previous configuration and it is being actively pursued.
- We had another baseline board failure yesterday (antenna 25, one sub-band). We are having about one BIB failure per week and this is worrisome. Problem is being investigated.
- We are continuing to do Mk5C testing and have written multiple streams without losing data in Vdiff format. Adam has been able to correlate the 2 streams that we have taken. We are going to try 4 streams and will do a VLBA correlation test later this week. This is good progress.
- We took 10ms dump science data for Casey's RRAT (rotating radio transient) program. We achieved ~ 100 MB/s data rate using 3 sub-band pairs. We found that the data was being collected so fast that it started collecting in a buffer because the write speed was not as fast. But the executor didn't know this and would cut off the program when the time was up without allowing time for the remaining buffered data to be written to disk. Padding the SB with dummy scans at the end allowed the final scans to be written to disk and the data looks excellent. Martin has been looking into the problem of miss-matched data collection versus writing speeds – adding more CBE nodes might help the throughput. This kind of data remains solidly in the domain of Resident Shared Risk.

- We completed a Ku-band holography test that ran without a hitch. We will be getting 2 6hr runs later this week. Michiel Brentjens wrote new holography reduction code and this will be verified with this new data.

Software status (Bryan)

- The exposure calculator had some changes made based on the technical review discussion we had last week. More changes are in the works.
- New requirements for the PST have been communicated to the software firm "Open Sky", but the written "work-order" hasn't been given to them yet (our contact at Open Sky went on vacation).

CASA status (Steve/Jeff)

- There is a new casapy-stable 3.4 build 19407 that fixes many of the issues that came up (such as the multi-threading slowdown – multi-threading is turned off by default in this release). It is undergoing testing. If you want to make use of the multi-threading you can run casapy-test which has this enabled (but beware of slowdowns for normal datasets). Also in the new stable are new region capabilities in (image) viewer.
- There will be another new stable later this week after some testing.
- The move of 3.4 to a new Release is still a bit uncertain. We hope to release the software to our users in early to mid-May.

General capabilities being offered and PST requirements (Claire)

- We now have an agreed-upon set of general capabilities that will be delivered to our users. Claire has a presentation summarizing these capabilities from a user perspective (see her presentation for details at: https://safe.nrao.edu/wiki/pub/EVLA/ECSVMeetings/Claire-JVLA_capabilities_rev2.pptx). Some notes from the presentation and discussion are:
- We will offer the following general capabilities to our users for the 9jul12 proposal deadline (for observations starting in Jan 2013):
 - 8-bit samplers with up to 2 GHz band-pass, up to 2x16 sub-band pairs (16 per baseband), up to 16,384 channels (baseline board stacking only; no recirculation).
 - Note: 16k channels are the most you can get without recirculation.
 - 3-bit samplers with up to 8 GHz band-pass and 128 MHz sub-bands that are contiguous in frequency within a baseband.
 - Phased VLA output for VLBI with 8-bit samplers. 1 or 2x128 MHz sub-bands, independently phased.
 - We need to discuss what will be shared-risk and what will be offered as a general capability. Walter has suggested that we have narrower resolutions and this may be considered shared-risk.
 - Up to 3 independent sub-arrays using the 8-bit samplers with 128 MHz contiguous sub-bands only. Each sub-array will be controlled

with a separate scheduling block. The SBs will be submitted independently to the observing queue and then started all at the same time.

- We will have data rate restrictions:
 - 20 MB/s max – This will support 3s integrations with the 8 GHz BW system.
 - We will also limit the allowed dump time per array configuration.
 - Blind red-shift survey CO work will need 3 to 5s integration time in D and C configurations. With so many channels, 3s integration times will take us to 20 MB/s.
 - If we implement real-time frequency averaging for continuum observations, this will take the data rates down. We haven't tested this fully yet – once we get this working then Claire proposes that we change the NRAO default for 3-bit continuum to include frequency averaging.
 - D and C configurations will be fine with our current ideas about data restriction. B config in late 2013 will need frequency averaging or we will have to increase the archive size – this needs to be evaluated carefully.
- We will have shared risk programs for, e.g.,: 1 or 2 recirculation setups, Sub-arrays with 3-bit samplers, Moderately fast dump times but not the fastest we can achieve, More than 16 sub-band pairs per baseband and 3-bit sub-band BWs other than 128 MHz (the science case for this is to cover the full frequency BW of C-band)
 - Notes:
 - Shared risk programs will have 1 hour of test time. Test time will be considered test-priority A. But approved shared risk programs will be given observing priority B (just in case we can't get them done). We will request that the TAC limits the amount of hours that can be allocated to shared-risk, and we will give the TAC that limit.
 - Priority details about how shared risk programs will be handled needs more discussion.
- We will maintain the resident shared risk program but the minimum residency requirement will be relaxed and it will scale to 1 month per 20 hrs of observing.
 - RSRO subjects will include: OTF mosaics, phased array for all but VLBI, pulsars, complicated sub-arrays, general recirculation set-ups, real-time transient detection.
 - We haven't discussed the maximum number of hours that will be allocated to shared risk programs in a given configuration. But we will have to discuss this.
- Bryan's SSS group will create a separate java webstart tool that will output correlator configuration information. Our standard NRAO defaults will be in

a drop-down PST menu. User-defined set-ups will be able to be uploaded into the PST.

Pipeline testing (Claire) – see Claire’s presentation at:

<https://safe.nrao.edu/wiki/pub/EVLA/ECSVMeetings/04-ECSVmtg-24apr12-PLtesting.pptx>

- There is now a prototype pipeline that can be run on pre-flagged measurement sets. It runs with CASA 3.3 (no automated flagging). If you have a pre-flagged dataset that you have processed by hand, please test the pipeline and compare your processing results with the PL results. Please respond to Claire about any issues, and whether you would be willing to use this as the start for your data processing.
 - See Claire’s presentation for the full list of feedback she would like from each person testing.
 - Note: Claire sent around an e-mail last week that describes the heuristics of the pipeline. This can be downloaded at: <https://safe.nrao.edu/wiki/pub/EVLA/ECSVMeetings/04-ECSVmtg-24apr12-PLheuristics.pptx>
 - One of the outputs of this testing will be to identify what datasets the PL will work for; understand the boundaries of where the current PL works and then start pushing the boundaries.
 - Claire will find out who has a pre-flagged MS to see if there are enough testers.
 - We will continue to work on developing methods for, e.g., using weak calibrators.
 - We need people to start testing on Friday (not Thursday).
 - The proto-pipeline will also be ported into CASA 3.4 which has rflag automated flagging available. So anyone who has rflag experience, please tell Claire about the heuristics you have developed.
- There will now be pipeline testers meetings on Fridays at 9am.