

SCUBA-2 Data Reduction Pipeline Lab Acceptance

Frossie Economou (frossie@jach.hawaii.edu)

Description of test

The purpose of the acceptance exercise was to allow JAC staff to assess whether the software complies with all the requirements that can be evaluated prior to the start of commissioning. Some criteria can be assessed by inspection. Others require a practical demonstration (henceforth "test"). The test was carried out on JAC systems by on May 5-9th 2008, and consisted of running the software under the current operational platform (64-bit Linux), examining the output and assessing performance. The tests were conducted by Frossie Economou (JAC). Also attending were Andy Gibb (UBC), Ed Chapin (UBC) and Tim Jenness (JAC). At various times, Craig Walther (JAC), Wayne Holland (ATC) and Douglas Scott (UBC) were present.

The input to the test is simulated data files served by a simulated DA provided by the project representing output that would be typical MSB. The data set was specified to have the following properties:

- Observation type
 - a pointing
 - a focus
 - a standard observation
 - a science target of various properties
- Size of field and files
 - four 5x5 deg maps in a diagonal configuration (to test output file sizes do not exceed 2 GB limit)
 - a 25 x 1 deg map oriented at 45 degrees wtr tracking co-ordinates
- Target type:
 - Fixed
 - Moving
- Array configuration
 - 2 arrays at 850 and 2 arrays at 450(for data handling)
 - 1 array at 850 and 1 array at 450 (interim performance test)
- Observation Mode
 - SCAN (all 3 observation types)
 - STARE/DREAM (all 3 observation types)
- Auxiliary files
 - disp.dat
 - stripchart configuration
- Duration
 - Minimum 1 hour's worth of observing total for at least one test

Copies of the input and output files of the test are available by request, please email frossie@jach.hawaii.edu

Systems tested:

1. DR (QL)
2. DR (science)
3. stripchart

Test details

The following test runs formed in order to inform conformance of the acceptance criteria.

1. An MSB-like dataset consisting of pointing, focus, standard, and science target. All data simulated in SCAN mode. Target for pointing, focus and standard was a moving target (Mars), whereas the science data was a field in Orion simulated from IRAS data. The science observations consisted of 5 observations of 8 minutes each, with the total data set estimated to take 50 minutes on-sky. The pipeline was run in *off-line* or *science* mode.
2. As in test 1, but data was simulated to be in STARE mode and run in *quick-look* mode.
3. As it test 1, but pipeline was run in "quick-look" mode
4. As in test 2 but pipeline was run in *science* mode
5. A series of 1x1 degree scan observations offset in a diagonal pattern in order to test behaviour with very large maps

Note that from the point of view of the pipeline, STARE and DREAM observations are identical.

Issues known prior to acceptance

Since the requirements were reviewed, SKYDIP was identified as a necessary observing mode. No simulated data or reduction for SKYDIP observations is available at this point, but it has to be developed as required by the SCUBA-2 commissioning plan.

Performance was only evaluated with 2 arrays simulated. The pipeline is know to fail its performance test with 8 arrays simulated, and work to address this is ongoing and targeted for December 2008, prior to the installation of the full array complement inside the instrument.

Acceptance criteria - Derived from Requirements

The acceptance exercise covered the points listed in the CDR compliance matrix (SC2/SRE/S210/002), excluding items marked as "C" (compliance demonstrated at commissioning) and items marked as "D" (compliance demonstrated by design) unless otherwise noted . Most items considered were listed are "S" (simulation) and "I" (integration and acceptance) in the original compliance matrix. Other changes have been made to account for any relevant changes in the SCUBA-2 project since then, for example the delivery of the instrument with only 2 of the final 8 arrays, and the anticipation of the SCAN being the primary observing mode. All these changes are noted in the

"Modification" Column. Also, no points relating to the FTS and the Polarimeter were be considered at this time.

Requirement	Description	Compliance	Description of test	Pass
G1	The Linux/x86 and Solaris/SPARC operating systems shall be supported for the DR Software	D	Solaris/SPARC compliance waived. Test performed on 32 and 64-bit Linux (no performance criteria to be applied to the 32-bit test)	64 bit:Y 32 bit:N/A
GR3	The Software shall conform to the JAC Software Requirements.	I	See separate section on JAC software requirements below	See below
GR9	There shall be a small number of data reduction Recipes to select from, with no free parameters.	S	By inspection of RECIPES directory.	Y
GR12	There shall be a procedure to correct for calibration drifts.	S	ORAC-DR primitive to calculate, apply and log FCFs	Y but log per arcsec and per beam FCFs in different files
GR15	The data generated by the Pipeline shall conform to a well documented format and content.	I	SC2/SOF/IC210/02	Y but JAC should review IC210/02
GR16	All file sizes shall be kept less than 2GB and shall be able to be manipulated by systems with a 32 bit address space.	S	By inspection of file sizes.	N Y Addendum #1
GR17	Errors shall be propagated in all data reduction methods.	S	Presence of NDF VARIANCE component	Y
MR1	All available observing modes shall be supported by the Pipeline.	I	Tested DREAM/STARE, SCAN, SCAN POINTING and FOCUS	N Focus failed under most circumstances Y Addendum #1
MR2	In Off-line mode the DREAM or STARE data shall be re-reducible from the raw 200 Hz time series rather than just the 1 Hz images (using the same algorithms).	I	Is there a way to calculate the DREAM solution independently of the online system?	Y (not tested but there is a documented application in SMURF to do this)
MR3	There shall be an efficient SCAN reduction method which is fast enough to compute that it does not make the Pipeline fall behind.	S	Performance to be evaluated at a later date, prior to delivery of all 8 arrays.	PUNT
MR4	Any reduced image shall not degrade the resolution by more than 10% compared with a perfect telescope beam.	S	For the simulated STANDARD source, compare FWHM of simulated input image to result reported by pipeline.	Y (no degradation) but weird bug where GAIA reported FHWM in pixels not "
MR5	There shall be no unrecoverable astrometry errors introduced by the DR system which are worse than the resolution at 450 microns.	S	Compare astrometric positions of sources in the simulated and output image. Test on moving and non-moving sources.	Y
MR7	There should be the possibility of running a nearly lossless SCAN reduction algorithm off-line even although that may take much more computing resources.	S	Iterative map-maker can be run with more iterations offline.	Y

Requirement	Description	Compliance	Description of test	Pass
PR2	By default the Pipeline shall be generating just images (which may subsequently be converted to polarization Stokes parameters or an FTS spectral cube).	S	Products are 2-D. FTS or POL not in scope for this acceptance.	Y
PR3	Engineering recipes shall exist for array setup and flat-fielding etc.	I	Removed. Now performed internally by the DA.	N/A
PR4	Specific Recipes shall be written that will be optimised for different situations, e.g. for when there is a bright point source in the map, when large areas are being mapped, or the map is going very deep.	I	While this is marked as verifiable at integration time, the level of simulations available do not allow such custom recipes to be tested. Therefore these will be evaluated at commissioning	PUNT
PR5	The Pipeline shall be able to handle the expected data rate, with up to a 30% margin, so that quality control information is delivered in a timely manner in most circumstances.	S	Pipeline tested to reduce data faster than the simulated acquisition time in both QL and science mode.	Y: 2 arrays N: 8 arrays
PR8	The Pipeline shall generate and store calibration data from each night in order for it to be possible to use the information later.	S	Presence of log.fcf and other logs as appropriate in \$ORAC_DATA_OUT	Y
PR9	The noise shall be estimated robustly and accurately based on detector element variances and integration times.	S	Inspect iterative map maker noise residuals for presence of any structure.	Y though one map did have higher noise than the others, presumed simulator bug
PR10	It shall also be possible to obtain a noise map based on the variances of pixel values among individual images (this is separate from the local spatial variances in the image, which includes confusion and source structure).	S	(KAPPA) display COMP=ERR	Y
PR11	A measure of the weight per pixel (the number of recorded bolometer values contributing to each reconstructed pixel) shall be available.	S	(KAPPA) display .MORE.SMURF.WEIGHTS	Y but not propagated to group correctly
PR12	As well as noise estimates, Recipes shall make basic data quality assessments, including the expected range of the data, so that data can be flagged...	S	Logs of: FCF, FWHM, error beam contribution, NEFD, bolometer noise statistics, sky background	Y: FCF, FWHM, noise, sky N: NEFD, error beam
PR13	All images generated by the Pipeline shall have a full astrometry header.	S	(KAPPA) ndftrace fullframe	Y
PR14	The Pipeline shall be able to process data from the two wavelength channels independently.	S	Run pipeline tests from different xterms.	Y
PR15	The Pipeline shall be able to reduce a combined image at each wavelength (although the DA system deals with each sub-array separately).	S	Run test with 2 arrays per wavelength simulated data	Y

Requirement	Description	Compliance	Description of test	Pass
PR16	Observers will need to know that the observations are proceeding correctly and that the signal-to-noise is improving through the night.	S	Report RMS of the current group (display in stripchart?)	Y
PR17	The Pipeline should be able to correct for pointing drifts using bright sources in fields which are observed multiple times in one night.	S	Primitive in the science recipe or in mapmaker?	N
PR19	The Pipeline should be able to compare reduced image data with source catalogues extracted from a previous night's data.	S	Can be broken down into (a) generate source catalogues (b) store them in a manner accessible by future invocations (c) compare catalogues	N
OR1	The Off-line system shall be able to reproduce all of the steps performed by the Data Acquisition System on the raw data	S	Identical to MR-2	Y
OR2	In Off-line mode the Pipeline shall be able to process all 4 sub-arrays as a single unit using the 200 Hz data.	S		Y
OR3	In Off-line mode 16 hours of data shall be reducible in 24 hours, including calibration, but this can be implemented using multiple computers.	S	Performance to be evaluated at a later date, prior to delivery of all 8 arrays.	PUNT (currently fails)
OR5	The Software shall provide Recipes which are optimized for mosaicing large areas and for carefully combining images taken in the deepest fields.	S	HEALPIX projections	N
OR6	The Software shall be able to deal with at least 20° x 20° maps (where the pixel area distortion in the tangent plane projection exceeds 10%), subject to staying within the 2GB file size limit.	S	Simulated data to cover 20° x 20°	N needs to tile output to 2GB files
OR10	The available Off-line Software should contain a decent, robust source extractor (similar to SEXTRACTOR 9,10 perhaps) which works well in uncrowded fields.	S	Point source extraction primitive (calls SEXTRACTOR?)	N
OR11	The Off-line extraction Software also should report a measure of how extended vs. point-like each source is.	S	Clump extraction primitive (CUPID)	N
OR12	The available Off-line software should make it relatively straightforward to search for variable or moving objects off-line.	S	Needs to write catalogues that can be cross-correlated with a 3 rd party tool.	N
DR1	There shall be a set of near real-time displays available to the observer. These displays may grab partially processed data from the Pipeline data reduction, but they shall not interfere with the running of the main Pipeline.	S	Run the test, inspect displays	Y

Requirement	Description	Compliance	Description of test	Pass
DR2	An early version of the Data Display System shall be ready for use in testing the instrument.	I	QL and stripchart was used at ATC	Y
DR3	A map of the currently taken data shall be displayed, available as fast as possible, but dependent on the observing mode. This will be in Nasmyth coordinates, but with an approximate (RA,Dec) overlay.	S	Run the test, inspect displays. Displays now in RA/Dec, exceeding requirement as written.	Y
DR4	The most recently co-added map shall also be displayed separately. This shall be available in (RA,Dec) coordinates and be calibrated.	S	Run the test, inspect displays	Y
DR5	There shall be a means of viewing the array variations in near real-time during observing.	S	MOVIE display system. Withdrawn in the light of GAIA's capabilities	DROP
DR6	There shall be an Off-line means of viewing the array variations (including the ability to combine data from multiple files).	S	Use WCSMOSAIC to combine subarrays. Use GAIA cube display to view data	Y
DR7	For the DREAM mode, frames shall be updated at the same rate they are being written to disk (i.e. As fast as 1Hz), and with essentially no delay behind real-time.	S	Run the test, inspect displays. [DA simulator may not be able to run at 1Hz]	Y
DR8	The DREAM/STARE Quick Look system shall not fall behind in displaying frames, and should drop the smallest number of frames in order to keep up.	S	Run test. System reports when it drops frames.	Y
DR9	For SCAN mode the image shall be updated every time an individual scan has been completed, with the display appearing before the next scan is complete.	S	Run the test, inspect displays. Data should appear roughly every minute.	Y
DR10	There shall also be a plot of estimated noise in the co-added map. This will be estimated using knowledge of the detector variances and integration times in sub-images.	S	Run test. Inspect stripchart.	Y
DR11	A map of the noise of the detector elements measured over a fixed time interval shall be available (used for assessing which detector elements are misbehaving for example).	S	This would be a product of the pipeline. [Can the iterative mapmaker do this?]	N
DR12	A "strip-chart" shall be used for displaying time varying data generated by the Pipeline	S	Run test. Inspect stripchart	Y
DR13	It shall be possible to display multiple strip-charts in a single user interface.	S	Run test. Inspect stripchart	Y

Requirement	Description	Compliance	Description of test	Pass
DR14	It shall be possible to display multiple data sets on the same strip-chart with shared x and y-axes.	S	Run test. Inspect stripchart	Y
DR15	It shall be possible for the observer to adjust the time period covered by the strip-chart.	S	Run test. Try it.	Y but time period in MJD very unhelpful, please adjust
DR16	The y-axis for the strip-chart shall be configurable to either autoscale or use values supplied by the user.	S	Run test. Try it.	Y
DR17	It shall be possible to pre-configure the strip-chart so that it can automatically locate the type and location of data it is expected to display (e.g. data from log files, parameter systems etc).	S	Examine stripchart configuration file.	Y
DR18	It shall be possible for more than one strip-chart to be displayed. (e.g. one for the observer and one for the telescope operator).	S	Try it.	Y
DR19	The strip-chart shall include a facility for displaying monitor information, including at least the median power across the array, FCF performance, opacity information (e.g. Water Vapour Radiometer, 12 the CSO tau meter ¹¹ and skydips), and beam sizes.	S	Examine stripchart configuration file to determine whether its syntax supports these features.	Y
DR20	Calibrated images of particular significance (e.g. mosaics) shall be displayed on a separate window so that they remain visible for longer periods than the ongoing co-adds.	S	Run test. Establish that group files are displayed in a different window.	Y
DR21	It should be possible to overlay catalogues of objects on the co-added images, or import other images to compare.	S	GAIA does this.	Y
DR22	The display of a large area should allow the possibility of pan and zoom function.	S	GAIA does this.	Y

Acceptance points - Derived from JAC Software Requirements

These requirements were predominantly developed for low-level software. Some are waived or modified given the DR pipeline context.

Requirement	Description	Description of test	Pass
BSR1	The software must be written in a programming language acceptable to JAC operations	Perl, C	Y
BSR2	The source code must be delivered checked in as a module in the JAC CVS source code repository.	In JAC SciCom subversion repository	Y
BSR3	All the code in the system must be able to be fully compiled by checking out from the CVS repository executing a single configuration command to set the installation directory of the system and its dependencies (this can be a file edit) and a single make command to build all the code. The exact procedure must be documented. No code written as part of the project shall generate any compiler warnings with the strictest compiler warning settings.	Standard Starlink build system for SMURF. ORAC-DR part not compiled code.	Y
BSR4	All external software dependencies required for building or running the system must be documented and dependent software supplied, if required, as part of the project deliverables.	No external dependencies outside the Starlink tree.	Y
BSR5	If software is compiled from code to executable form, the executable files should be installed into a different tree from the source code. The system should continue to run after the source code tree has been deleted.	This is the case for Starlink packages.	Y
BSR6	User documentation must be provided describing how to take data in normal operations.	The SCUBA equivalents of this are: SUN/216 (SURF) and SUN/231 (SCUBA ORAC-DR recipes)	N

BSR7	Programmer documentation must include a system design document outlining the software structure, a brief description of the functionality and operation of each software module, plus reference documentation of a brief description of every source code function. The function descriptions must be able to be extracted from the source code using an appropriate program	What is required here is a SMURF equivalent to SSN/72 for SURF.	N
BSR8	All the external interfaces to the software should be explicitly documented and agreed with JAC operations as part of the software development. External interfaces should be as generic as possible, using existing generic interfaces as often as possible.	All external interfaces are through the file system and are described into SC2/SOF/IC210/01	Y
BSR9	The entire system must be able to be run up by entering a no more than two commands after a login to one of the JAC workstations. If two commands are required, the first would be to setup the environment for a particular configuration of the system.	To be demonstrated (test run by typing oracdr_scuba2[s q] and then oracdr).	Y
BSR10	The system may be distributed across machines and operating systems. Software that has real time constraints may use the VxWorks operating system and the 68000 and Power PC architectures.	Waived	N/A
BSR11	All network communication between machines must utilize either the Drama IMP or EPICS Channel Access protocol or another, agreed upon, widely used, protocol.	Only DRAMA used.	Y
BSR12	It must be possible to switch the currently running version of the system by running the system down, executing a single command (for example, changing a soft link), re-initializing the running environment, and running the system up again.	ORAC_DIR and STARLINK_DIR allow relocation	Y
BSR13	No DRAMA task should have the paths to other tasks hard-coded, and they should generally also not even hardwire any names (including their own).	Quick look monitor to be inspected to verify this requirement.	Y

BSR14	All commands that are run from the Unix shell must have -h (prints usage notes) and -v (prints version number) options. They should also have a -sim (simulate) option, if appropriate.	Test oracdr and stripchart -h and -v SMURF follows Starlink convention: type <i>starversion SMURF</i> for version; type <i>smurfhelp</i> for help.	Y
BSR15	All Software documentation should be installed in the JAC documentation system before commissioning has been completed.	All user and programmer documentation to be issued SUN numbers and committed to SVN.	N

Outcome

Following four days of running and re-running the various tests, the status of the SCUBA-2 DR project can be summarised as follows:

- Reasonable progress has been made for this stage of the projection and many requirements are met
- A number of requirements are not met at this point
- A number of bugs were uncovered during the test; these have been communicated to the DR team and need to be fixed prior to on-sky observing at the telescope

In regards to the requirements that are not met, they fall under three categories:

- should be met prior to installation at the telescope: GR16, MR1, OR6 (do not generate large files),
- should be met during or by the end of science commissioning: PR12, DR11, BSR6,
- do not need to be met until the full complement of science arrays is delivered: PR5, PR17, PR19, OR5, OR6 (tile larger area into smaller files), OR10, OR11, OR12, BSR7, BSR15

Action

A “delta” acceptance test will be performed on the week of May 26th to verify that all bugs uncovered during acceptance testing have been fixed, and all requirements needing to be met prior to installation are in place.

Addendum #1

A “delta” acceptance test was performed on Monday June 3rd, 2008. Tests 1, 2, 4 and 5 were repeated using a local code tree for ORAC-DR checked out from the SVN repository. It was established that

- Bugs reported to the team during the main acceptance test appear to be fixed, with the pipeline running well and with no unexpected warnings
- FOCUS observations were correctly handled, fulfilling requirement MR1
- All files, including temporary files, remained under the 2GB file limit, fulfilling requirement GR16
- Final mosaics on large areas were adjusted to show the middle of the field subject to the 2GB file limit. This is an acceptable intermediate step towards full compliance with requirement OR6.

This successfully completed the lab acceptance exercise.