AIPS Memo 77

Summary of DDT Accuracy Results

Ernest Allen and Glen Langston

92 September 3

The current results of the *AIPS* DDTs (Dirty Dozen Tests) are presented. The DDTs consist of a set of inputs and outputs to a few of the most commonly used *AIPS* tasks. The purpose of the DDTs is to determine whether the *AIPS* software output is changing with time. The DDTs also measure the relative performance of the CPUs running *AIPS*. The changes to the *AIPS* task outputs have two types of causes, *1)* improvements and *2)* bugs. The latter should be detected by these tests and the bugs fixed.

Three types of tests have been developed, *1)* continuum imaging (DDTs), *2)* continuum *uv*-data calibration (VLAC), and *3)* spectral line *uv*-data calibration (VLAL). A DDT for VLBI software is needed, but has not yet been implemented.

A new continuum calibration (VLAC) master tape has been released, to accommodate *AIPS* calibration changes.

1 INTRODUCTION

In order to assure the continuity of the *AIPS* tasks results, sets of test procedures have been developed. These tests are important for tracking changes in the software and detection of bugs. The results are presented in three sections, continuum imaging, continuum calibration and spectral line calibration. Current results of these tests are described in turn below.

The Small, Medium and Large DDT results are presented in Appendix A, the VLAC results in Appendix B, and the VLAL results in Appendix C. The forms useful for reporting the DDT results to the *AIPS* group are presented in Appendix D. A listing of *AIPS* messages produced by the VLAC test are shown in Appendix E and the VLAL test messages are shown in Appendix F.

2 Continuum Imaging, DDT

The Continuum imaging DDTs consist of three size tests, Small, Medium and Large. The small test creates 256x256 pixel images, medium 512x512 and large 1024x1024. The method for running these tests has been described in *AIPS* memo 73.

The latest results are presented in the tables in Appendix A. These results compare the current operation of the *AIPS* tasks with the results of these tasks when the last master tape was created. The current master tape was created in October 1989. Three different types of CPUs are listed in the results tables. Table 1 lists the names of the CPUs tested.

Two types of tests are executed, imaging and *uv*-data tests. The imaging test results are expressed as the number of `bits` of agreement between the master and the test images. In modern CPUs, the floating point numbers are represented with 32 `bits`, 8 `bits` of exponent and 24 `bits` of mantissa. In the tables, two `bits` numbers are given for each image compared,

Nrao1	A Convex C1 computer
Baboon	A Sun Sparc2 Workstation
Lemur	An IBM 6000 (Power Station), model 530

Table 1: NRAO CPU's tested

a) the number of `bits` of agreement in the pixel with the maximum difference, and *b)* the average number of `bits` of agreement between the master and test images.

The *uv*-data tests compare differences between the master and test *uv*-data using the *AIPS* task UVDIF. This task looks for three kinds of *uv*-data differences, *1)* differences in flux density of the visibilities, *2)* differences in *u,v and w*, and *3)* and other differences, when there does not appear to be any relation between the master and test visibilities. The third type of difference often occurs when the *uv*-data end up with a different sort order. This happens frequently because several of the visibilities in the master dataset have nearly identical *u* values. The test resorts are reported as list of the three numbers of differences, ie. the 0,0,2 in column 6 of the first DDT results table.

The typical accuracy now expected for the DDT tests is 10 to 16 `bits`. The *AIPS* accuracy is calculated as the number of `bits` that match exactly between the master and test images. The number of `bits` of accuracy is a function of the internal floating point algorithm of the CPU, the improvements of the software and errors introduced in inadvertently. Since most CPUs now conform to IEEE floating point standards, the first of these should have little affect. The second of these confuses the test results and requires periodic creation of new master images. The inadvertent introduction of errors is detected by strangely low numbers of `bits`.

The algorithm for the *AIPS* task MX has improved for large images to the point where the reference image has only 9 `bits` matching when compared to current results.

Note that because the computers were often very heavily loaded during the DDT tests, the relative CPU times are NOT useful in comparing the performance of the different types of computers.

The master DDT tape will probably be updated sometime in the near future. The current Small and Medium DDT master data are available via *anonymous* `ftp`. Below is the README.DDT file in the `baboon ftp` area.

```
----------------------------------------------------------------------
   We have small and meduim DDT files available via ftp.  The small
and medium DDTs have been written to compressed tar files.  In order
to run the DDTs on a remote cpu, the small or medium DDT master
data and images must be transferred.

   Log onto your CPU as user ``aips''.  The Small DDT requires about
7 megabytes and the medium DDT requires 30 megabytes of space in your
FITS area.  The steps for getting the master data and images are
```

listed below: (It is assumed you are using a CPU with the UNIX operating system.)

```
   #  Unix Command.                            Comment
 ---  -------------                   ------------------------------
   0. cd $FITS                        Directory containing images
   1. ftp baboon.cv.nrao.edu          Computer containing images
                                      (telnet # 192.33.115.103)
   2. login anonymous                 Anyone can ftp the files
   3. password= guest name            Identify your self with passwd
   4. > cd pub/aips/DDT               move to directory with DDTs
   5. > binary                        change to binary mode
   6. > get ddt.small.tar             Get small DDT data and images
      > get ddt.medium.tar            or Medium DDT, 4 times bigger
   7. > quit                          Exit ftp
   8. tar -xvf ddt.small.tar          will unpack files DDTS1 - DDTS11
```

It is probably best to get and execute the small DDT first, then if desired, delete the small DDT and get the medium DDT.

If you don't already have the document "AIPS DDT HISTORY", MEMO #73, you will need it for reference in running the test. Change to the pub/aips/memos directory, and get file AIPSMEMO73.TEX. Take a look at the file AAAREADME for info on other documents.

```
                  Loading the DDT master FILES into AIPS
                                                   Comment
1. aips new                               to start up AIPS

                              enter AIPS commands as follows:

2. > task= 'fitld'                         FITLD loads uvdata + images
3. > douvcomp= -1                          Do not compress uvdata
4. > dotable= 1                            Load AIPS tables
5. > ncount =0
6. > outn= ''; outs= o; outcl= ''          Use default out names
7. > outdisk=2                             Set output disk as you wish
8. > input                                 Check FITLD inputs
                                           Use loop to read 11 files
9. > for i= 1 to 11; infile= 'DDTS'!!char(i); go; wait; end

     When all 11 files have finished loading, proceed with running the
     test as instructed in MEMO #73. Repeat the procedure for meduim test
     replacing DDTS with DDTM.
-------------------------------------------------------------------------
```

3 Continuum Calibration, VLAC

The VLA continuum calibration test is designed to check the tables produced by the *AIPS* calibration process for a continuum, two IF observation. There is only one size of calibration test (unlike the DDTs which come in three sizes, small, medium and large).

The VLAC test was created by Bill Cotton in 1990 and the current master data and images were previously updated on December 6, 1990. On 1992 June 16, a new master VLAC tape was created using the 15APR92 release of *AIPS*. This tape is now available for distribution and may be requested with the latest releases of *AIPS*.

The VLA continuum calibration test executes in a manner similar to the DDT tests. The test *AIPS* procedures are created by running the RUNFIL VLACLOAD. The test master images are loaded from tape into *AIPS* using the RUNFIL VLACEXEC. The VLAC test parameters are specified by specifying the adverbs in VLAC (ie. within *AIPS* typing INPUT VLAC). After the adverbs are correctly set, the adverb values must be saved in *AIPS* by typing TPUT VLAC. Then the VLAC test is executed by running the RUNFIL VLACEXEC.

The VLAC tasks tested are listed in table 1.

VLAC Task	Purpose
UVCOP	Copy a selection of *uv*-data
UVFLG	Flag bad *uv*-data
SETJY	Set the flux density in the source table. Defines the observation gain
CALIB	Assuming point source observations, the antenna complex gains are estimated
GETJY	Measure the calibration source flux densities
CLCAL	Interpolate antenna gain measurements into CL table for all observations
PCAL	Calibrate the intrinsic antenna polarization
CLCOR	Correct the CL table for intrinsic antenna polarization
SPLIT	Apply calibration and create single source data sets
HORUS	Fourier Transform the calibrated *uv*-data. Creates Beam I and Q images

Table 2: Continuum Calibration Tasks tested by VLAC

The VLAC results are presented in Appendix B, and the entire *AIPS* messages for the VLAC test are listed in Appendix E. The execution times for the VLAC test are not recorded. The DDTs must be used to check the execution times of the CPUs.

VLAL Task	Purpose
UVCOP	Copy a selection of *uv*-data
AVSPC	Average spectral line channels to create a new *uv*-dataset
UVFLG	Flag erroneous *uv*-data
SETJY	Set the flux density in the source table. Defines the observation gain
CALIB	Assuming point source observations, the antenna complex gains are estimated
GETJY	Measure the calibration source flux densities
CLCAL	Interpolate antenna gain measurements into CL table for all observations
BPASS	Band pass calibration of the spectral line channels
CLCOR	Correct the CL table for intrinsic antenna polarization
SPLIT	Apply calibration and create single source data sets
HORUS	Fourier Transform the calibrated *uv*-data. Creates a cube of images

Table 3: Spectral Line Calibration Tasks tested by VLAL

4 Spectral Line Calibration, VLAL

The Spectral Line Calibration process is tested by VLAL. The test is similar to the VLAC test with the addition of several tasks for the relative calibration the spectral channels (the band pass calibration). The VLA spectral Line calibration (VLAL) tests were created by Doug Wood with modifications and improvements by Phil Diamond and Dave Adler.

The tasks tested by VLAL are listed in the table 2.

The current Spectral Line DDT master tape was created on 1991 November 4, using the 15APR92 version of *AIPS*. The Spectral Line DDT is still being refined and the master data and images will be re-released shortly. The Spectral Line DDT results are presented in Appendix C. A listing of the *AIPS* messages produced by the VLAL test are shown in Appendix F.

5 Results Forms

The DDT results are useful for comparing the IEEE floating point compliance of CPUs running *AIPS* as well as comparing the speed of the CPUs. *AIPS* System managers running the DDTs on new types of CPUs are requested to report their results to the *AIPS* group. In Appendix D, the DDT reports forms are presented.

Test Date	AIPS Version	CPU	Old New Tst	UVSRT CPU Time	UVSRT U-V Diffs	UVMAP CPU Time	UVMAP MAP Peak	UVMAP MAP RMS	UVMAP BEAM Peak	UVMAP BEAM RMS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
92jan15	15APR92	baboon	T	4.9	0,0,2	14.0	14.2	19.5	12.6	18.1
92jan28	15APR92	nrao1	T	11.1	0,0,6	13.8	13.2	19.1	10.2	16.4
92jan09	15APR92	lemur	T	22.9	0,0,4	5.2	14.2	19.5	12.3	18.0

Table 4: Small DDT results for UVSRT and UVMAP

Test Date	AIPS Version	CPU	Old New Tst	APCLN CPU Time	APCLN MAP Peak	APCLN MAP RMS	APCLN CPU Time	APCLN RES Peak	APCLN RES RMS	ASCAL CPU Time	ASCAL U-V Diffs
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
92jan15	15APR92	baboon	T	94.3	14.1	20.7	10.4	16.9	22.2	63.3	0,0,2
92jan28	15APR92	nrao1	T	37.1	18.7	24.0	9.9	16.8	22.2	33.9	0,0,6
92jan09	15APR92	lemur	T	34.0	13.7	16.4	3.0	16.9	22.3	4.0	0,0,4

Table 5: Small DDT results for APCLN and ASCAL

Test Date	AIPS Version	CPU	Old New Tst	MX CPU Time	MX MAP Peak	MX MAP RMS	MX BEAM Peak	MX BEAM RMS	MX CPU Time	MX Real Time	MX CLN Peak	MX CLN RMS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
92jan15	15APR92	baboon	T	15.7	13.9	19.3	14.9	19.7	176.6	270.0	13.8	17.5
92jan28	15APR92	nrao1	T	20.1	12.3	18.8	14.4	19.5	98.4	548.0	14.5	17.3
92jan09	15APR92	lemur	T	6.1	13.8	19.5	14.7	19.8	56.4	82.0	13.6	17.4

Table 6: Small DDT results for MX

Test Date	AIPS Version	CPU	Old New Tst	VTESS CPU Time	VTESS CLN Peak	VTESS CLN RMS	Notes
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
92jan15	15APR92	baboon	T	57.5	19.2	26.8	Sun-Sparc2, light load
92jan28	15APR92	nrao1	T	57.8	18.6	25.9	Convex C-1, heavy load
92jan09	15APR92	lemur	T	15.9	20.3	27.9	IBM/530, light load

Table 7: Small DDT results for VTESS and Notes

Test Date	AIPS Version	CPU	Old New TST	UVSRT		UVMAP				
				CPU TIME	U-V DIFFS	CPU TIME	MAP PEAK	MAP RMS	BEAM PEAK	BEAM RMS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
92aug19	15APR92	baboon	O	4.6	0,0,2	22.9	15.2	18.2	15.2	18.7
92aug19	15APR92	nrao1	O	15.6	0,0,2	27.4	14.7	17.9	14.5	17.9
92aug19	15APR92	lemur	O	2.5	0,0,2	9.3	14.9	18.2	15.1	18.7

Table 8: Medium DDT results for UVSRT and UVMAP

Test Date	AIPS Version	CPU	Old New Tst	APCLN						ASCAL	
				CPU Time	MAP Peak	MAP RMS	CPU Time	RES Peak	RES RMS	CPU Time	U-V Diffs
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
92aug19	15APR92	baboon	O	481.9	11.8	14.7	18.2	15.3	21.2	282.1	0,0,2
92aug19	15APR92	nrao1	O	184.0	11.9	14.6	23.5	15.4	21.2	132.2	0,0,2
92aug19	15APR92	lemur	O	229.1	11.8	14.7	7.5	15.3	21.2	123.6	0,0,2

Table 9: Medium DDT results for APCLN and ASCAL

Test Date	AIPS Version	CPU	Old New Tst	MX								
				CPU Time	MAP Peak	MAP RMS	BEAM Peak	BEAM RMS	CPU Time	Real Time	CLN Peak	CLN RMS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
92aug19	15APR92	baboon	O	23.2	14.9	18.2	15.4	18.7	582.1	656.0	8.1	12.7
92aug19	15APR92	nrao1	O	33.3	14.6	17.9	14.9	17.9	273.0	471.0	8.6	12.4
92aug19	15APR92	lemur	O	10.8	14.8	18.2	15.3	18.7	269.9	309.0	8.6	12.4

Table 10: Medium DDT results for MX

Test Date	AIPS Version	CPU	Old New Tst	VTESS			Notes
				CPU Time	CLN Peak	CLN RMS	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
92aug19	15APR92	baboon	O	66.2	21.1	29.0	Sun-Sparc2, light load
92aug19	15APR92	nrao1	O	90.4	21.5	29.0	Convex C-1 light load
92aug19	15APR92	lemur	O	29.9	21.3	29.3	IBM/530 light load

Table 11: Medium DDT results for VTESS and Notes

Test Date	AIPS Version	CPU	Old New Tst	UVSRT CPU Time	UVSRT U-V Diffs	UVMAP CPU Time	UVMAP MAP Peak	UVMAP MAP RMS	UVMAP BEAM Peak	UVMAP BEAM RMS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
92mar04	15APR92	baboon	T	48.4	8,1,32	194.7	13.1	15.6	14.2	16.2
92feb20	15APR92	nrao1	T	83.6	52,0,148	93.5	12.2	15.6	14.0	16.2
92mar03	15APR92	lemur	T	31.2	18,0,74	44.2	13.1	17.4	15.1	17.7

Table 12: Large DDT results for UVSRT and UVMAP

Test Date	AIPS Version	CPU	Old New Tst	APCLN CPU Time	APCLN MAP Peak	APCLN MAP RMS	APCLN CPU Time	APCLN RES Peak	APCLN RES RMS	ASCAL CPU Time	ASCAL U-V Diffs
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
92mar04	15APR92	baboon	T	2055.7	10.9	16.8	186.7	14.2	20.7	7718.1	8,1,32
92feb20	15APR92	nrao1	T	557.5	13.0	17.3	82.8	14.2	20.7	2620.8	52,0,148
92mar03	15APR92	lemur	T	673.4	10.8	16.8	35.0	14.2	20.6	2454.1	18,0,74

Table 13: Large DDT results for APCLN and ASCAL

Test Date	AIPS Version	CPU	Old New Tst	MX CPU Time	MX MAP Peak	MX MAP RMS	MX BEAM Peak	MX BEAM RMS	MX CPU Time	MX Real Time	MX CLN Peak	MX CLN RMS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
92mar04	15APR92	baboon	T	178.5	13.9	18.1	15.2	19.5	2881.7	3829.0	9.5	15.8
92feb20	15APR92	nrao1	T	101.1	13.8	18.1	14.8	19.5	1048.5	2767.0	9.7	15.9
92mar03	15APR92	lemur	T	43.1	14.0	18.1	15.2	19.6	3722.0	4547.0	9.0	15.3

Table 14: Large DDT results for MX

Test Date	AIPS Version	CPU	Old New Tst	VTESS CPU Time	VTESS CLN Peak	VTESS CLN RMS	Notes
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
92mar04	15APR92	baboon	T	820.5	18.6	27.2	Sun-Sparc2, heavy load
92feb20	15APR92	nrao1	T	368.3	21.2	29.1	Convex C-1 heavy load
92mar03	15APR92	lemur	T	182.6	18.6	27.2	IBM/539 light load

Table 15: Large DDT results for VTESS and Notes

Test Date	AIPS Version	CPU	Old New Tst	CALIB SN Real MAX	CALIB SN Real RMS	CALIB SN Imag MAX	CALIB SN Imag RMS	GETJY SU Iflux MAX	GETJY SU Iflux RMS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
92aug24	15APR92	lemur	O	21.7	27.6	22.0	27.7	99.0	99.0
92aug26	15APR92	nrao1	O	21.4	27.4	21.4	27.6	25.3	26.9

Table 16: VLAC results for CALIB and GETJY

Test Date	AIPS Version	CPU	Old New Tst	CLCAL SU Iflux MAX	CLCAL SU Iflux RMS	CLCAL CL Real MAX	CLCAL CL Real RMS	CLCAL CL Imag MAX	CLCAL CL Imag RMS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
92aug24	15APR92	lemur	O	99.0	99.0	18.9	23.9	19.5	24.2
92aug26	15APR92	nrao1	O	25.3	26.9	21.1	25.6	21.7	26.0

Table 17: VLAC results for CLCAL

Test Date	AIPS Version	CPU	Old New Tst	PCAL AN Polc MAX	PCAL AN Polc RMS	CLCOR SU Iflux MAX	CLCOR SU Iflux RMS	CLCOR CL Real MAX	CLCOR CL Real RMS	CLCOR CL Imag MAX	CLCOR CL Imag RMS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
92aug24	15APR92	lemur	O	20.5	24.9	22.4	23.9	18.9	23.9	18.9	24.0
92aug26	15APR92	nrao1	O	21.2	25.5	19.9	21.5	21.1	25.6	21.2	25.7

Table 18: VLAC results for PCAL and CLCOR

Test Date	AIPS Version	CPU	Old New Tst	SPLIT MAX	SPLIT RMS	HORUS HO Imap MAX	HORUS HO Imap RMS	HORUS HO Ibem MAX	HORUS HO Ibem RMS	HORUS HO Qmap MAX	HORUS HO Qmap RMS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
92aug24	15APR92	lemur	O	15.1	30.4	19.7	23.9	19.4	23.4	18.8	23.1
92aug26	15APR92	nrao1	O	15.2	30.4	16.5	21.5	19.3	23.0	18.0	21.9

Table 19: VLAC results for SPLIT and HORUS

Test Date	AIPS Version	CPU	Old New Tst	AVSPC					SPLIT		
				U-V	CH 0 IF 1		CH 0 IF 2		U-V	Bits	
				Diffs	MAX	RMS	MAX	RMS	Diffs	MAX	RMS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
92aug22	15APR92	nrao1	O	0,0,0	30.2	46.2	30.2	46.2	2526,0,0	7.9	18.9

Table 20: VLAL results for AVSPC and SPLIT

Test Date	AIPS Version	CPU	Old New Tst	GETJY							
				$Real_1$		$Imag_1$		$Real_2$		$Imag_2$	
				MAX	RMS	MAX	RMS	MAX	RMS	MAX	RMS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
92aug22	15APR92	nrao1	O	99.0	99.0	99.0	99.0	99.0	99.0	99.0	99.0

Table 21: VLAL results for GETJY

Test Date	AIPS Version	CPU	Old New Tst	CLCAL							
				$Real_1$		$Imag_1$		$Real_2$		$Imag_2$	
				MAX	RMS	MAX	RMS	MAX	RMS	MAX	RMS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
92aug24	15APR92	nrao1	O	99.0	99.0	99.0	99.0	99.0	99.0	99.0	99.0

Table 22: VLAL results for CLCAL

Test Date	AIPS Version	CPU	Old New Tst	BPASS							
				$Real_1$		$Imag_1$		$Real_2$		$Imag_2$	
				MAX	RMS	MAX	RMS	MAX	RMS	MAX	RMS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
92aug24	15APR92	nrao1	O	14.0	17.8	9.8	13.3	14.3	17.5	9.2	13.4

Table 23: VLAL results for BPASS

Test Date	AIPS Version	CPU	Old New Tst	HORUS							
				HO map		HO beam		HO map		HO beam	
				MAX	RMS	MAX	RMS	MAX	RMS	MAX	RMS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
92aug24	15APR92	nrao1	O	15.0	18.2	99.0	99.0	14.7	17.7	99.0	99.0

Table 24: VLAL results for HORUS