AIPS MEMO NO. 71

A Comparison of DDT results

Patrick P. Murphy

April 8, 1991

Revised by:

National Radio Astronomy Observatory
Edgemont Road, Charlottesville, VA 22903–2475
804–296–0211 (FTS 940–7211), x328; TWX 910–997–0174
aipsmail@nrao.edu

INTRODUCTION

This memo documents the results of running the *AIPS* benchmark suite of programs — the DDT* tests — on the IBM RS/6000 series of workstations and servers, and on the Convex C–1. The systems tested were:

- IBM Model 530: 48 Megabytes, 200–500 Megabytes free disk space
- IBM Model 540: 64 Megabytes, 70–200 Megabytes free disk space
- Convex C-1: 32 Megabytes, 400 Megabytes free disk space, integral vector processor

In addition, an IBM RS/6000 model 320 with 24 Megabytes of memory and about 400 Megabytes of available disk space was tested for a brief period during the course of the tests described herein. Unfortunately, by the time the specific test had stabilized, this system was no longer available. However, it is possible to use some of the early tests to infer the approximate performance of this system.

All IBM systems were running version 3.1 of the *AIX* operating system. The model 320 was a loaner workstation from IBM in Richmond (Thanks to George Latimer), the 530 is the Charlottesville `lemur` workstation, and the 540 is the `polaris` server. The Convex system is `nrao1`.

Close to 30 separate tests were done on these systems, ranging from the small DDT suite, through subsets of the large, to the full-blown large test. Two main versions of *AIPS* were used: `15JAN91` and `15JUL91`, although at one point a `15APR91` version of `MX.EXE` was tried for comparative purposes. The main purpose of the many tests was to eliminate spurious effects from the results. The effects in question ranged from incorrect values of `EDGSKP` and the AP size to competition for CPU cycles from runaway emacs "zombie" processes.

It should be stressed that the version of *AIPS* used for most of these tests (`15JUL91`) was a *frozen snapshot* of the continually changing *AIPS* "TST" version that exists on `nrao1` and the many machines throughout NRAO served by the *AIPS* midnight jobs. This snapshot was taken on about February 20. The intent here of course is to prevent day-to-day changes made by *AIPS* programmers from affecting the results. There is nothing special about either the date or the characteristics of the *AIPS* code resulting from this snapshot. Among the IBM systems, the exact same binaries were used in running the program. These were initially compiled on the RS/6000–320 and moved to the other systems as needed, although the final figures presented here resulted from the identical code rebuilt on the RS/6000–540 and moved to the 530 later. It is assumed that a rebuild of the code on each RS/6000 system would not significantly affect the results.

In the final results, the compilation characteristics of the *AIPS* code was either identical or as near as possible to identical among the different systems tested. The system was compiled without debugging information and with no optimization, except for the `QPSAP` routines which were as fully optimized as the native compilers would allow. The AIX fortran compiler only has one optimization setting, whereas the Convex fortran has several. The value used on the Convex was `-O2`.

* The so-called *dirty dozen tests*.

ACCURACY RESULTS

For the purposes of comparison, the medium DDT was run on all systems with the same size pseudo-array processor (1280 kwords). All the relevant input parameters were the same for each run and the `EDGSKP` parameter was set to 8. The accuracy results were identical for all three IBM tests. The conclusion one can draw from these figures is that there are few significant differences in accuracy between the Convex and the IBM RS/6000 series. If anything, the IBM seems to give better overall results than the Convex, although the difference is slight. The results are presented in Table 1.

Table 1. Accuracy Results, 15JUL91 MEDIUM test on IBM RS/6000 and Convex C–1 systems.

	RS 6000		Convex C–1	
Task	Peak Bits	Rms Bits	Peak Bits	Rms Bits
UVMAP	14.97	18.25	14.63	17.90
UVBEAM	15.15	18.68	14.52	17.89
APCLN	11.78	14.73	11.92	14.61
APRES	15.27	21.26	15.37	21.20
MXMAP	14.85	18.25	14.36	17.91
MXBEAM	15.28	18.68	14.68	17.90
MXCLN	9.99	14.19	10.29	14.18
VTESS	21.31	29.34	23.12	30.85

In the course of the trial runs, it was determined that reducing the AP size to 256 kwords resulted in slightly poorer accuracy results for both `MXMAP` (14.25/18.24) and `MXBEAM` (14.83/18.67), on the RS/6000 systems.

Additional accuracy results were also obtained for all DDT test sizes with the 15JAN91 version of $\mathcal{AIPS}$. This was performed on the model 540 with `EDGSKP` set to 4, 8, or 16 as appropriate, and a 256 kword AP.

Table 2. Accuracy results, 15JAN91 tests, RS6000/540.

Task	Large		Medium		Small	
	Peak	Rms	Peak	Rms	Peak	Rms
UVMAP	12.33	17.29	14.97	18.25	13.49	19.33
UVBEAM	15.12	17.75	15.15	18.68	10.83	16.71
APCLN	10.95	16.79	11.78	14.73	13.74	16.38
APRES	14.11	20.62	15.27	21.26	17.03	22.43
MXMAP	13.88	18.92	14.25	18.24	13.04	18.97
MXBEAM	15.10	19.59	14.83	18.67	14.61	19.72
MXCLN	9.47	15.86	9.99	14.19	13.80	17.53
VTESS	18.71	27.37	21.31	29.34	21.74	28.84

TIMING RESULTS

Table 3 below summarizes the results of the timing tests on the Convex, the IBM rs540 and rs530. The numbers are derived from the printed values reported by $\mathcal{AIPS}$ on termination of the various tasks. The numbers for tasks `UVSRT` and `UVDIF` represent the *sum* of the two times those tasks are run in the DDT suite, and those for `COMB` represent the sum of the eight separate runs of that task within the DDT procedure.

Note that the scratch disk used in all cases was the same as that used for the data — the IBM internal disk.

Table 3. Times for MEDIUM DDT tests, $\mathcal{AIPS}$ 15JUL91.

Task	RS6000/530		RS6000/540		Convex C–1	
	CPU	Wall	CPU	Wall	CPU	Wall
UVSRT	9.24	26	7.48	17	36.72	50
UVDIF	5.51	24	4.57	9	19.04	33
CCMRG	2.43	8	2.03	5	10.80	18
SUBIM	1.74	5	1.35	4	6.50	8
COMB	13.76	62	11.54	31	60.58	85
UVMAP	10.03	26	8.17	20	31.16	40
APCLN	230.68	242	192.56	203	200.95	240
APRES	8.39	14	6.82	11	25.88	34
ASCAL	123.85	129	103.43	108	130.70	152
MXMAP	15.95	31	13.22	25	46.48	60
MXCLN	284.95	312	237.94	260	294.04	349
VTESS	36.64	54	30.26	42	108.14	135
Total:	743.17	933	619.37	735	970.99	1204
Units of C1's:	1.31	1.29	1.57	1.64	1.00	1.00

From earlier results, one can arrive at a rough figure for the relative performance of the IBM RS/6000 model 320. This system was able to run the `MXCLN` phase of the DDT in 368.12 seconds. Using this result to scale the model 320 CPU time relative to the other IBM RS/6000 models, we can infer a performance factor of roughly 1.01 relative to a Convex C-1 over all the tests in CPU time.

The figures above clearly show that the IBM RS/6000 systems are considerably faster than the Convex at the typically scalar tasks, but the Convex of course excels in the tasks such as `APCLN` and `MXCLN` that vectorize well on its architecture. Despite this, the Convex can only beat the RS/6000 model 530 on one task (`APCLN`) in CPU usage, and even there the wallclock time is almost equal.

OPTIMIZATION TIMING

While running the `15JAN91` version of *AIPS* on the RS6000/540 system, various levels of optimization were used while running the large DDT test several times. The levels were:

- No optimization at all, everything compiled with debug on
- Only the Q routines optimized
- Everything optimized

The only other difference was that `EDGSKP` was accidentally set to 4 instead of 16 for the non-optimized (first) case. It is very unlikely that this error would affect the numbers presented below to any large extent. The conclusions would certainly not be affected at all. The AP size used for these tests was 256 kwords.

Table 4. Times for Large DDT, 15JAN91, RS6000/540.

Task	No Optimization		Q opt.only		Everything optimized	
	CPU	Wall	CPU	Wall	CPU	Wall
UVSRT	46.12	80	45.38	70	19.27	76
UVDIF	25.81	37	22.81	36	22.76	23
CCMRG	5.50	9	5.34	9	5.37	9
SUBIM	4.53	9	2.73	9	2.80	9
COMB	37.35	113	23.66	113	23.37	114
UVMAP	136.82	160	35.50	65	35.54	66
APCLN	2337.29	2361	1121.75	1145	1121.86	1145
APRES	144.24	154	32.50	43	32.51	43
ASCAL	3816.36	3849	1949.32	1965	1949.07	1965
MXMAP	202.52	238	53.05	81	52.99	76
MXCLN	3061.16	3143	1245.92	1301	1247.44	1307
VTESS	651.97	696	172.31	216	172.19	215
Total:	10469.67	10849	4710.27	5053	4685.17	5048
Ratio:	1.00	1.00	0.45	0.47	0.45	0.47
Speedup:	1.00	1.00	2.22	2.15	2.23	2.15

As with the previous tests, the results of the two runs of UVSRT, two of UVDIF, and eight¡ of COMB are summed together.

The conclusions are obvious. Optimizing the key set of "Q" routines makes a major difference to elapsed times for most of the *AIPS* tasks. Applying the optimization to the rest of the system makes little or no difference to either elapsed or CPU times.

Effects of NFS

The following table shows the details of one DDT medium run, performed on the rs530 with *everything* on a NFS-mounted disk: the executable binaries, the data disks, scratch files. The partition in question was on the rs540 disk.

Table 5. Times for MEDIUM DDT tests, *AIPS* 15JUL91.
IBM RS6000/530, data and binaries on NFS disk.

Task	CPU	Wall
UVSRT	10.50	156
UVDIF	6.25	24
CCMRG	3.69	39
SUBIM	2.56	22
COMB	23.50	183
UVMAP	13.68	90
APCLN	240.63	376
APRES	11.62	63
ASCAL	126.93	204
MXMAP	21.70	130
MXCLN	302.38	611
VTESS	42.17	190
Total:	805.61	2088
Units of C1's:	1.21	0.58

The results from this test indicate some slight degradation in CPU time; it is unclear what the origin of this effect is. It is not surprising that the wallclock time degrades significantly; this of course shows clearly that such a setup is severely network-bound. For reference, this test was performed at an off-peak hour, with both systems almost completely unloaded and very little network traffic.