Comparison of Origin 2000 and Origin 3000 Using NAS Parallel Benchmarks

Raymond D. Turney
Computer Sciences Corporation
Numerical Aerospace Simulation Facility
NASA Ames Research Center
MS 258-6
Moffett Field, CA. 94035
email: turney@nas.nasa.gov

Introduction

This report describes results of benchmark tests on the Origin 3000 system currently being installed at the NASA Ames National Advanced Supercomputing facility. This machine will ultimately contain 1024 R14K processors. The first part of the system, installed in November, 2000 and named mendel, is an Origin 3000 with 128 R12K processors. For comparison purposes, the tests were also run on lomax, an Origin 2000 with R12K processors.

The BT, LU, and SP application benchmarks in the NAS Parallel Benchmark Suite and the kernel benchmark FT were chosen to determine system performance and measure the impact of changes on the machine as it evolves. Having been written to measure performance on Computational Fluid Dynamics applications, these benchmarks are assumed appropriate to represent the NAS workload. Since the NAS runs both message passing (MPI) and shared-memory, compiler directive type codes, both MPI and OpenMP versions of the benchmarks were used. The MPI versions used were the latest official release of the NAS Parallel Benchmarks, version 2.3. The OpenMP versions used were PBN3b2, a beta version that is in the process of being released. NPB 2.3 and PBN 3b2 are technically different benchmarks, and NPB results are not directly comparable to PBN results.

Links to descriptions of the benchmarks themselves:

NPB description

PBN description

The benchmarks were run on mendel, an R12K Origin 3000, and on lomax, an R12K Origin 2000. While the processor chips were not of the same revision number,they have the same MHz ratings, and are all R12K chips.

All runs were Class C, and compiled with 64-bit addressing. The MPI programs were compiled with the -O2 compiler flag, described as extensive optimization by the SGI compiler man pages. The OpenMP runs were compiled with the -O3 compier flag, the highest level of optimization on the Origin. Different flags were used because the MPI BT Class C benchmark ran faster when compiled with the -O2 flag.

After running the MPI benchmarks it was discovered that this was not the case for all benchmarks, so the -O3 flag, normally the faster option, was used for the Open MP benchmarks. When the MPI benchmarks were run compiled with the -O3 flag on the O3K, the timings were within 5% of the times obtained compiling with -O2, so compiler flag choice did not significantly affect the results.

Summary of Results

The the average execution times of the MPI runs for each machine, as well as the sum of the averages, are listed in Table 1.

Table 1 - Summary of MPI Results

Machine	BT(sec)	FT(sec)	LU(sec)	SP(sec)	Total
Mendel - O3K 400MHz	1659.80	347.55	915.81	1018.20	3941.36
Lomax - O2K 400MHz	2613.65	430.73	922.04	1433.78	5400.20

The ratio of the total times is:

5400/3941=1.37

Averaging over all the MPI benchmark runs, the O3K was about a third faster than the O2K. There is considerable variation from one benchmark to another, and on LU the O3K was not significantly faster than the O2K. As explained in more detail below, lack of performance improvement on LU may be an effect of run time variation on the O3K. The minimum time to run LU on the O3K was significantly less than the minimum time on the O2K, 789.90 sec on the O3K compared to 910.50 sec on the O2K.

Table 2 lists the average execution times of the OpenMP runs for each machine, and the sum of the averages.

Table 2 - Summary of OpenMP Results

Machine	BT(sec)	FT(sec)	LU(sec)	SP(sec)	Total(sec)
Mendel - O3K 400MHz	825.59	231.90	764.37	827.36	2649.22
Lomax - O2K 400MHz	969.44	297.41	996.88	1178.87	3442.60

The ratio of the total times is:

3442 / 2649 = 1.30

Averaging over all the benchmarks, the O3K was also about a third faster on the OpenMP version of the benchmarks.

These results, MPI and OpenMP, suggest that codes represented by NPB and PBN Class C, should run about a third faster on an O3K than on an O2K. This performance improvement is on machines using the same CPUs (400 Mhz R12K), with 32Kb instruction and data caches, and 8Mb unified secondary caches. Wide variation from one benchmark to the next was observed. In one case, MPI LU, on average,

no performance improvement was observed.

Results

Class C was used because it is the largest size of these benchmarks, and the primary purpose of these machines is to run large jobs. Sixteen cpus was selected as representative of a small to medium size job, and a convenient number for the benchmarks and the machines. The O2 optimization level was selected for the MPI benchmarks because the MPI NPB BT Class C runs faster on the O2K when compiled with O2 optimization than with O3 optimization. The O3 optimization level was selected for the OpenMP benchmarks because it is the highest level of optimization on the Origin. 64-bit addressing was selected because it is impossible to compile Class C for some of these benchmarks using 32-bit addresses.

The timings and MOPS counts for each run are presented below in Tables 3, 4, 5 and 6. To get a reasonable sample, seven runs of each benchmark were done on each machine. In general run time variation proved insignificant, but MPI LU was an exception. One run on the O3K took significantly longer than any run on the O2K, and others took about the same amount of time on both machines. This created a collective result indicating that there was almost no performance improvement on MPI LU in going between the O2K and the O3K. This result has been replicated by another investigator (private communication), and similar lack of performance improvement has been observed for INS3D (private communication), a NASA legacy CFD code. Since the replication was done under PBS using cpusets, it is unlikely that the lack of performance improvement is an artifact of scheduling or cpusets. No explanation for either the run-time variation or lack of performance improvement on MPI LU is currently available. It is possible that this is merely an artifact of run time variation, since the minimum time for an O3K run of MPI LU was significantly less than the minimum time for an O2K run of MPI LU.

The O2K runs were done on a machine controlled by a custom PBS scheduler written by Ed Hook, which uses cpusets and an awareness of machine topology to insure execution on physically contiguous nodes. Because the O2K was space shared, not time shared, interference from other jobs was minimized. The lack of run time variation among benchmark runs on the O2K supports this hypothesis. The O3K runs were done interactively, one at a time, on an otherwise idle machine. Thus, run times were unaffected by the simultaneous execution of other jobs.

Table 3 - Origin 3000 MPI results

BT Class C	Seconds	MOPS
#1	1656.00	1730.84
#2	1664.21	1722.31
#3	1653.27	1733.70
#4	1662.06	1724.54
#5	1664.34	1722.17
#6	1655.53	1731.33
#7	1663.21	1723.34
Median	1662.06	1724.54
Mean	1659.80	1726.89

Std. Dev.	5.07	4.92
FT Class C	**Seconds**	**MOPS**
#1	387.12	1023.94
#2	327.23	1211.33
#3	327.42	1210.66
#4	342.23	1158.24
#5	357.87	1107.62
#6	331.31	1196.43
#7	359.69	1102.01
Median	342.23	1158.24
Mean	347.55	1144.32
Std. Dev.	22.10	70.02
LU Class C	**Seconds**	**MOPS**
#1	943.83	2160.35
#2	789.90	2581.33
#3	881.03	2314.32
#4	857.67	2377.36
#5	927.38	2198.66
#6	1026.98	1985.44
#7	983.86	2072.45
Median	927.38	2198.66
Mean	915.81	2241.42
Std. Dev.	79.95	200.75
SP Class C	**Seconds**	**MOPS**
#1	1025.12	1414.56
#2	1021.55	1419.51
#3	1017.83	1424.70
#4	1010.55	1434.96
#5	1022.39	1418.35
#6	1021.44	1419.66
#7	1008.54	1437.82
Median	1021.44	1419.66
Mean	1018.20	1424.22
Std. Dev.	6.32	8.80

Hardware info:
IRIX64 mendel 6.5 10120105 IP35

128 400 MHZ IP35 Processors
CPU: MIPS R12000 Processor Chip Revision: 3.5
FPU: MIPS R12010 Floating Point Chip Revision: 3.5
Main memory size: 32768 Mbytes
Instruction cache size: 32 Kbytes
Data cache size: 32 Kbytes
Secondary unified instruction/data cache size: 8 Mbytes

Table 4 - Origin 2000 MPI Results

BT Class C	Seconds	MOPS
#1	2601.94	1101.59
#2	2603.80	1100.81
#3	2671.59	1072.87
#4	2596.60	1103.86
#5	2617.09	1095.22
#6	2602.80	1101.23
#7	2601.71	1101.69
Median	2602.80	1101.23
Mean	2613.65	1096.90
Std. Dev.	26.31	10.93
FT Class C	**Seconds**	**MOPS**
#1	446.82	887.13
#2	416.03	952.78
#3	427.79	926.59
#4	461.62	858.69
#5	417.05	950.46
#6	418.66	946.79
#7	427.17	927.95
Median	427.17	927.95
Mean	430.73	921.48
Std. Dev.	17.24	35.71
LU Class C	**Seconds**	**MOPS**
#1	965.07	2112.80
#2	910.50	2239.43
#3	912.93	2233.45
#4	911.94	2235.89
#5	911.87	2236.07
#6	912.10	2235.50

#7	929.88	2192.76
Median	912.10	2235.50
Mean	922.04	2212.27
Std. Dev.	20.17	46.77
SP Class C	**Seconds**	**MOPS**
#1	1407.62	1030.18
#2	1409.46	1028.83
#3	1408.73	1029.37
#4	1407.31	1030.41
#5	1464.34	990.28
#6	1527.87	949.10
#7	1411.31	1027.49
Median	1409.46	1028.83
Mean	1433.78	1012.24
Std. Dev.	46.36	31.42

Hardware info:
IRIX64 lomax 6.5 07061118 IP27
Processors: 512 400 MHZ IP27 Processors
CPU: MIPS R12000 Processor Chip Revision: 3.5
FPU: MIPS R12010 Floating Point Chip Revision: 0.0
Main memory size: 196608 Mbytes
Instruction cache size: 32 Kbytes
Data cache size: 32 Kbytes
Secondary unified instruction/data cache size: 8 Mbytes

Table 5 - Origin 3000 OpenMP

>>

BT Class C	**Seconds**	**MOPS**
#1	813.59	3523.00
#2	863.04	3321.14
#3	818.46	3502.03
#4	814.87	3517.49
#5	816.48	3510.55
#6	840.13	3411.71
#7	812.58	3527.37
Median	816.48	3510.55
Mean	825.59	3473.33
Std. Dev.	19.04	77.95

FT Class C	Seconds	MOPS
#1	256.27	1546.77
#2	222.41	1782.26
#3	245.64	1613.72
#4	228.65	1733.62
#5	223.28	1775.30
#6	224.85	1762.93
#7	222.24	1783.60
Median	224.85	1726.93
Mean	231.90	1714.03
Std. Dev.	13.54	95.04
LU Class C	**Seconds**	**MOPS**
#1	781.22	2610.03
#2	733.51	2779.79
#3	731.03	2789.23
#4	771.50	2642.89
#5	773.38	2636.48
#6	783.43	2602.67
#7	776.50	2625.87
Median	773.38	2636.48
Mean	764.37	2669.61
Std. Dev.	22.32	79.75
SP Class C	**Seconds**	**MOPS**
#1	832.00	1742.91
#2	825.16	1757.35
#3	821.33	1765.54
#4	822.25	1763.58
#5	824.20	1759.41
#6	823.61	1760.66
#7	842.94	1720.30
Median	824.20	1759.41
Mean	827.36	1752.82
Std. Dev.	7.70	16.12

Hardware info:
IRIX64 mendel 6.5 10120105 IP35
128 400 MHZ IP35 Processors
CPU: MIPS R12000 Processor Chip Revision: 3.5

FPU: MIPS R12010 Floating Point Chip Revision: 3.5
Main memory size: 32768 Mbytes
Instruction cache size: 32 Kbytes
Data cache size: 32 Kbytes
Secondary unified instruction/data cache size: 8 Mbytes

Table 6 - Origin 2000 OpenMP

BT Class C	Seconds	MOPS
#1	972.48	2947.38
#2	966.62	2965.27
#3	967.13	2963.69
#4	971.31	2950.93
#5	967.04	2963.96
#6	967.25	2963.32
#7	974.28	2941.95
Median	967.25	2963.32
Mean	969.44	2956.64
Std. Dev.	2.63	9.66
FT Class C	**Seconds**	**MOPS**
#1	299.09	1325.33
#2	294.88	1344.25
#3	297.57	1332.08
#4	298.68	1327.14
#5	296.23	1338.10
#6	295.48	1341.48
#7	299.94	1321.56
Median	297.57	1332.08
Mean	297.41	1332.85
Std. Dev.	1.93	8.65
LU Class C	**Seconds**	**MOPS**
#1	996.36	2046.45
#2	995.66	2047.88
#3	997.13	2044.86
#4	996.52	2046.11
#5	997.69	2043.73
#6	994.19	2050.91
#7	1000.67	2037.64

Median	996.52	2046.11
Mean	996.88	2045.37
Std. Dev.	2.03	4.11
SP Class C	**Seconds**	**MOPS**
#1	1181.06	1227.80
#2	1176.46	1232.60
#3	1178.01	1230.98
#4	1178.71	1230.25
#5	1178.12	1230.86
#6	1179.35	1229.58
#7	1180.37	1228.51
Median	1178.71	1230.25
Mean	1178.87	1230.08
Std. Dev.	1.55	1.62

Hardware info:
IRIX64 lomax 6.5 07061118 IP27
Processors: 512 400 MHZ IP27 Processors
CPU: MIPS R12000 Processor Chip Revision: 3.5
FPU: MIPS R12010 Floating Point Chip Revision: 0.0
Main memory size: 196608 Mbytes
Instruction cache size: 32 Kbytes
Data cache size: 32 Kbytes
Secondary unified instruction/data cache size: 8 Mbytes

Related Work

Sheila Faulkner has also benchmarked the Origin 2000 and the Origin 3000 using NAS Parallel Benchmarks. She ran all the NPB MPI benchmarks, and investigated scaling. She did not use the OpenMP versions of the benchmarks, and her Origin 2000 numbers are for turing, a 195 Mhz Origin 2000 no longer used as a compute server at the NAS.

Future Work

This is the first of what is hoped to be a number of benchmark studies comparing various IPG and NAS hosts to each other.