The Performance of the NAS High-Speed Processors in 2Q93

Robert J. Bergeron

RND-94-005 March 1994

NAS Systems Development Branch
NAS Systems Division
NASA Ames Research Center
Mail Stop 258-6
Moffett Field, CA 94035-1000
bergeron@nas.nasa.gov

Abstract

During 2Q93, the NAS C-90 delivered an average throughput of 3.315 GFLOPS while the NAS Y-MP averaged 0.665 GFLOPS. The ratio of throughput performance for the two machines is 3315/665 = 4.98. A ratio based on number of CPUs and peak speed gives a theoretical ratio of 5.75. Although the users are employing similar codes on both machines, the C-90 user programs display vector lengths which are short relative to the longer C-90 hardware length. These programs do not fully exploit the vector hardware and lead to a lower efficiency relative to the Y-MP. The measured vectorization of 92% is higher than the 86% levels measured for the Y-MP. This increase appears due to the improvements in the compiler technology.

1.0 Introduction

The introduction of the C-90 in March 1993 motivated the daily monitoring of the hardware performance of the NAS High Speed Processors (HSPs). The results will be reported on a quarterly basis and this paper, covering the 2nd quarter of 1993, is the first report in the series.

The C-90 Hardware Performance Monitor (HPM) continuously delivers a full 32-counter record[1] for the NAS workload and the C-90 discussion will include a report of all 32 counters. The Y-MP HPM can report only a single group (8-counter) during any one period[1]. Since NAS has chosen to monitor the Group 0 performance of all programs in the Y-MP workload, the Y-MP discussion will include only the Group 0 counters. This group provides a general quantitative overview of program performance.

NAS records the daily average values of all HPM counters. In 2Q93, installation of the C-90 in late March allowed 81 days of measured user service. This report provides tables of counter values representing the average, maximum, and minimum values of the 81 daily reports in this quarter. The Y-MP counters represent 79 days of user service as daily Y-MP monitoring was also begun after the quarter had started.

The tables provide performance rate data per CPU for the actual time the CPU spent executing the user programs. System throughput, however, is based on wall clock time and total number of CPUs. A complete explanation of all counter data occurs in [2].

To provide a feel for the daily variation in each of the counters, the report also provides the standard deviation (STD) and coefficient of variation (COV). The coefficient of variation is the ratio of the standard deviation of a quantity divided by its average value. Typically, a COV of less than 0.1 indicates a well-defined distribution[3].

2.0 C-90 Counter Data

The C-90 CPUs have a clock period (CP) of 4.167 nanoseconds and a peak speed of 960 MFLOPS. This architecture has 128-element vector registers and double-width 64-element functional units.

The report divides the 32 C-90 counters into 4 functional groups for easy reference. Table 1 provides counter data giving a global overview of C-90 performance.

Table 1: NAS C-90 Hardware Measurements-Global Counters-2Q93

Measurement	Unit	Avg	STD	COV	Min	Max
CPU time	sec	60361.	11959.	0.320	11865.	82706.
Instruction Issue	M/sec	59.327	3.817	0.064	50.103	67.413
Average clock periods/inst	-----	4.062	0.271	0.067	3.560	4.790
CP holding issue	Percent	66.093	1.949	0.029	60.871	70.213
Instruction buffer fetches	M/sec	0.257	0.054	0.211	0.110	0.379
Floating Pt. Ops per CPU	M/sec	243.776	45.480	0.187	89.814	321.689
Vector Floating Pt. Ops	M/sec	239.900	45.151	0.188	88.820	318.607
CPU memory references	M/sec	244.606	19.271	0.079	199.209	291.795
CPU memory conflicts	avg/ref	0.223	0.048	0.217	0.098	0.335
VEC memory references	M/sec	239.067	19.761	0.083	192.352	288.911
B/T memory references	M/sec	1.426	0.337	0.236	0.645	2.704
I/O memory references	M/sec	2.344	1.397	0.596	0.190	5.744
I/O memory conflicts	avg/ref	0.265	0.021	0.081	0.205	0.317

Global Counters:

During 2Q93, each of the C-90s 16 CPUs performed at an average rate of 244 MFLOPS, with the vector units contributing 98% of the floating point operations. The memory reference rate is about the same as the CPU MFLOP rate, indicating that each floating point operation requires one memory reference.

The I/O memory reference rate of 2.34 Mwords/sec per CPU represents an average sustained I/O rate of 0.300 Gbytes for the machine. This rate includes I/O which is both internal and external to the machine. For rates internal to the machine the relevant targets are the SSD and the disks in the 4 IOCs (68 DD60s). Since the SSD can sustain 13.6 Gbytes/sec and the 68 DD60s can sustain 1.4 Gbytes/sec, the C-90 I/O does not, on the average, challenge the data transfer capacity of its disks. The large COV for I/O illustrates the bursty nature of this data transfer. The Cray SAR (System Activity Report) indicates that about 10% of the HPM-reported I/O goes to the disk devices.

Table 2: NAS C-90 Hardware Measurements-Instruction Holds-2Q93

Measurement	Unit	Avg	STD	COV	Min	Max
Waiting on A-registers	% CPU	5.091	0.528	0.104	3.453	6.208
Waiting on S-registers	% CPU	9.211	1.438	0.156	6.203	13.807
Waiting on V-registers	% CPU	24.039	5.791	0.241	16.514	43.911
Waiting on B/T-registers	% CPU	1.165	0.228	0.196	0.575	1.685
Waiting on F'nctnal Units	% CPU	21.821	3.825	0.175	8.994	26.135
Waiting on Shared Regs	% CPU	0.271	0.191	0.706	0.001	0.900
Waiting on Memory Ports	% CPU	16.012	1.408	0.088	12.783	19.815
Waiting on Miscellaneous	% CPU	2.444	0.134	0.055	2.114	2.716

Instruction Holds:

Instructions are fetched from the instruction buffer by the instruction processor. If any of the resources required to execute the instruction are reserved, the instruction issue logic prevents the instruction from issuing. The HPM records all CPs for which the instruction holds issue and the table presents these as the percent of total CPU time. Since there may be more than one resource reservation preventing an instruction issue, the sum of the percentages in this group can exceed 100%.

For the NAS C-90 workload, the major resources causing instruction issue delays (instruction holds) are busy vector registers and busy vector functional units. The instruction processor will not issue an instruction until operations in these units have completed. Calculations derived from counter data will show that other operations were in progress during these delays.

The approximately equal delays in vector registers and vector functional units indicates efficient register use and overlapping of vector functional units.

Memory references can lead to two kinds of delay in the Y-MP/C-90 architecture. A memory instruction hold occurs, for example, when a register is reserved by another instruction or a memory port is busy. Table 2 shows that the fraction of CPU time the processor held issue is about 16%.

The second type of memory delay is termed a memory conflict (or memory contention), and this delay occurs when a needed bank is busy. A CPU memory port accesses a section which accesses a memory bank. A user program executing on a single CPU can encounter conflicts when it continuously references the same bank. A workload can encounter conflicts when several CPUs simultaneously reference the same bank.

The total delay due to memory references includes both the delays due to memory contention as well as the delays due to memory instruction holds. Data from Table 1 indicates that each memory reference on the average experiences a memory contention delay of 0.223 CP. Data from Table 2 indicates that memory resources prevent the CPU from issuing about 16% of the time and converting this delay to a per reference basis yields a memory delay of 0.15 CP.

Since HPM measurements indicate that most of the workload references are vector, the theoretical vector memory reference rate can provide the basis for judging the workload memory delay. For the Y-MP, vector read and write memory references require startup periods of 19 CPs and 3 CPs, respectively. After the startup period, data arrives at a rate of 1 word per CP. For the measured vector length of 63 (Table 4), the Y-MP can load data from memory at a rate of 1.30 CP/word; the Y-MP can store data to memory at a rate of 1.05 CP/word. C-90 rates are expected to be similar.

Total memory delay is about 0.37 CP/reference and this delay is a fraction of the 1.05 CP minimum required for a C-90 workload vector memory reference. For the current fraction of vector operations, the C-90 memory does not constitute a bottleneck.

Table 3: NAS C-90 Hardware Measurements-2Q93-Instruction Issues

Measurement	Unit	Avg	STD	COV	Min	Max
(000-004)Special	M/sec	1.279	0.425	0.332	0.815	2.820
(005-017)Branch	M/sec	2.731	0.264	0.097	1.921	3.335
(02x,030-033)A Register	M/sec	25.544	3.834	0.150	15.948	31.729
(034-037)B/T Memory	M/sec	0.151	0.038	0.254	0.064	0.267
(040-043,071-077)S Register	M/sec	8.005	1.179	0.147	5.683	10.505
(044-061)Scalar Integer	M/sec	4.650	0.756	0.162	3.374	7.191
(062-070)Scalar Floating Pt.	M/sec	3.876	1.160	0.299	0.995	7.234
(10x-13x)Scalar Memory	M/sec	4.113	0.681	0.166	1.945	6.152
(140-177)All Vector	M/sec	8.978	0.749	0.083	6.784	10.643

Instruction Issues:

A-register instructions comprise about 43% of the instructions issued. These instructions compute memory addresses and indexes for memory, loop control, and I/O.

Scalar instructions constitute about 42% of NAS workload instructions. About 40% of these scalar instructions employ the S-registers for logical functions and special data

transfers. Integer, floating point, and memory instructions provide approximately equal contributions to the remainder of the scalar instructions.

Vector instructions are only 15% of the total instructions, but vector operations represent about 92% of the workload operations (Table 5). A single vector instruction can produce up to 128 vector operations.

Table 4: NAS C-90 Hardware Measurements-2Q93-Vector Operations

Measurement	Unit	Avg	STD	COV	Min	Max
Vector Logical	M/sec	49.101	55.180	1.124	18.492	236.207
Vector Shift/Pop/LZ	M/sec	8.148	2.142	0.263	3.439	17.823
Vector Integer Add	M/sec	25.196	20.835	0.827	11.715	96.056
Vector Floating Pt.Multiply	M/sec	122.812	24.110	0.196	42.914	165.086
Vector Floating Pt.Add	M/sec	110.797	20.248	0.183	43.326	152.906
Vector Floating Reciprocal	M/sec	6.292	1.453	0.231	2.355	9.217
Vector Memory Read	M/sec	163.279	12.980	0.079	133.123	192.241
Vector Memory Write	M/sec	75.788	13.215	0.174	59.229	123.685
Average Vector Length	------	62.930	8.187	0.130	48.680	85.050

Vector Operations:

The workload vector length is about 63 whereas the C-90 hardware vector length is 128. User programs with vector lengths closer to the hardware length can better exploit the vector performance.

The measured value of 62.9 is less than the program vector length. For example, a program with a logical vector length of 192 requires two vector instructions, one for the 128 element block and a second for the 64 element block. The HPM will report a vector length of about 96 for this example. While the logical length may exceed the HPM value, the logical length cannot be shorter than the HPM value. Thus, the average NAS program vector length exceeds 62.9.

Vector memory load rates exceed vector memory store rates by a factor of 2. A FLOP requires, on the average, one memory reference, but it is more likely to be a load than a store. The architecture provides each CPU with two memory paths for loading data from memory and one memory path for storage. Performance would suffer with only one memory load path.

Table 5: NAS C-90 Hardware Measurements-2Q93-Derived Data

Measurement	Unit	Avg	STD	COV	Min	Max
System Availability	Percent	85.6	12.7	1.49	27.0	98.0
System MFLOPS	M/sec	3315.529	748.620	0.226	1109.450	4558.330
Vector Operation Fraction	Percent	91.661	1.255	0.014	88.010	94.600
Scalar Operation Fraction	Percent	8.339	1.255	0.150	5.400	11.990
Vector Operation Rate	M/sec	561.411	59.232	0.106	439.700	741.330
Scalar Operation Rate	M/sec	50.349	3.752	0.075	41.980	59.890
Total Operation Rate	M/sec	611.759	56.201	0.092	499.580	783.610
Instruction Issue Fraction	Percent	25.139	1.293	0.051	20.878	28.091
Hold Issue Fraction	Percent	65.627	1.680	0.026	60.871	69.139
Null Instruction Fraction	Percent	9.235	0.698	0.076	7.783	11.141

Derived Data:

Table 5 lists several quantities obtained through calculations with the counter data. System MFLOPS denotes the system throughput. This rate is the product:

System MFLOPS = MFLOPS/CPU *CPUs*Availability.

Availability is the fraction of time the C-90 operated in user mode. The table shows the throughput rate to be 3315 MFLOPS or 20.7% of the theoretical peak rate.

The table indicates that 91.7% of the operations were performed in vector mode and a total operation rate of 612 MOPS per CPU. Since this rate is about 2.55 OPS/CP, the instruction processor is able to overlap operations despite the large number of hold issue CPs.

Of the 612 MOPS, 40% were memory operations. If typical machine operations followed the “Load, Load, Operate, and Store” pattern, 75% of the operations would be memory operations. The reduced amount of memory usage confirms that the compiler is successfully using registers to limit memory operations.

A complete accounting of all CPs accumulated by the C-90 CPU while in user mode includes the time spent issuing instructions, the time spent holding instruction issue, and the time spent preparing for the next instruction. Cray terms the latter quantity NIP (Next Instruction Parcel) time and includes in it the CPs spent jumping across instruction buffers, CPs spent fetching words from memory to load the buffers, and CPs spent processing instruction of more than one word in length[1]. Partition of total CP time into

these categories can illustrate the reasons for performance differences between two similar workloads. In 2Q93, the C-90 spent about 25% of the user time issuing instructions, 66% of the user time holding issue, and 9% of the user time preparing for the next instruction. This breakdown is consistent with a sequential instruction issue punctuated by delay periods due to resource reservation. The small amount of NIP time indicates relatively well-written code.

3.0 Y-MP Counter Data

Table 6 provides 2Q NAS Y-MP per-CPU counter data for Group 0. Table 7 provides values calculated from the counter data and several reference values from previous Y-MP monitoring.

Table 6: NAS Y-MP Hardware Measurements-2Q93-Global Counters

Measurement	Units	Average	STD	COV	Min	Max
CPU time	Sec	62720.	15093.	0.241	13447.	81497.
Instruction Issue	M/sec	37.549	2.387	0.064	30.220	43.560
Average clock periods/inst	-----	4.457	0.292	0.065	3.830	5.520
CP holding issue	Percent	67.338	2.411	0.036	60.860	75.530
Instruction buffer fetches	M/sec	0.293	0.040	0.138	0.160	0.420
Floating Pt. adds	M/sec	44.839	4.901	0.109	32.690	61.130
Floating Pt. multiplies	M/sec	49.022	4.782	0.098	37.170	63.460
Floating Pt. reciprocals	M/sec	2.920	0.414	0.142	1.900	4.250
Floating Ops/CPU	M/sec	96.782	9.432	0.097	72.530	126.490
CPU memory references	M/sec	99.332	7.359	0.074	77.330	123.630
I/O memory references	M/sec	0.719	0.711	0.989	0.150	4.730

Global Counters:

The Y-MP CPUs have a clock period (CP) of 6.000 nanoseconds and a peak speed of 333 MFLOPS. During 2Q93, each of the Y-MPs 8 CPUs performed at an average rate of 97 MFLOPS. The CPU memory reference rate of 99 million references per second is slightly more than the CPU MFLOP rate, indicating that each floating point operation requires slightly more than one memory reference.

The I/O memory reference rate of 0.719 Mwords/sec per CPU represents an average sustained I/O rate of 0.046 Gbytes for the machine. This rate includes I/O which is both internal and external to the machine. For rates internal to the machine, the relevant targets are again the SSD and the disks in the 2 IOSs (68 DD40s). Since the SSD can sustain 1.6 Gbytes/sec and the 40 DD40s and 8 DD49s can sustain 0.460 Gbytes/sec, the Y-MP I/O does not, on the average, challenge the data transfer capacity of its disks. Previous measurements have indicated that traffic to and from the SSD represents about 90% of the Y-MP I/O [2].

Table 7: NAS Y-MP Hardware Measurements-2Q93-Derived Data

Measurement	Units	Average	STD	COV	Min	Max
System Availability	Percent	86.0	5.20	0.060	53.0	96.0
System MFLOPS	M/sec	665.200	81.700	0.123	449.130	886.450
Vector Operation Fraction	Percent	86.				
Scalar Operation Fraction	Percent	14.				
Vector Operation Rate	M/sec	201.				
Scalar Operation Rate	M/sec	33.0				
Total Operation Rate	M/sec	234.68				
Instruction Issue Fraction	% CPU	22.534	1.454	0.065	18.132	26.136
Hold Issue Fraction	% CPU	67.331	2.450	0.036	60.860	75.530
Null Instruction Fraction	% CPU	10.335	1.070	0.106	6.338	13.004

Derived Data

Table 7 lists several quantities obtained through calculations with the counter data. System MFLOPS denotes the system throughput. This rate is the product:

System MFLOPS = MFLOPS/CPU *CPUs*Availability.

Availability is the fraction of time the Y-MP operated in user mode. The table shows the throughput rate to be 665 MFLOPS or 24.9% of the theoretical peak rate.

The table provides a value of 86% for the fraction of Y-MP operations performed in vector mode based on previous NAS measurements [2]. The total operation rate of 235 MOPS per CPU represents about 1.40 OPS/CP, indicating that the instruction processor is able to overlap operations despite the large number of hold issue CPs.

Of the 234 MOPS, 43% were memory operations. If typical machine operations followed the "Load, Load, Operate, and Store" pattern, 75% of the operations would be memory operations. The reduced amount of memory usage confirms that the compiler is successfully using registers to limit memory operations.

As with the C-90, a complete accounting of all CPs accumulated by the CPU while in user mode includes the time spent issuing instructions, the time spent holding instruction issue, and the time spent preparing for the next instruction. The Y-MP spent about 23% of the user time issuing instructions, 67% of the user time holding issue, and 10% of the user time preparing for the next instruction. The Null Instruction Fraction, while still constituting a small amount of user CPU time is about 10% larger than that of the C-90.

4.0 Discussion

The ratio of throughput performance for the two machines is 3315/665 = 4.98. A ratio based on number of CPUs and peak speed gives a theoretical ratio of 5.75. The ratio of measured CPU performance for the machines is 244/97 = 2.67, whereas the ratio based on peak speed alone is 1000/333 = 3.0. Although the users are employing similar codes on both machines, the C-90 user programs display vector lengths which are short relative to the longer C-90 hardware length. These programs do not fully exploit the vector hardware and lead to a lower efficiency relative to the Y-MP.

On a per CP basis, the C-90 issues slightly more instructions than the Y-MP (1 every 4.0 CPs on the C-90 vs. 1 every 4.4 CPs on the Y-MP). For the same amount of work per CP, the longer vector lengths of the C-90 should require fewer vector instructions. However, the C-90 CPUs deliver more FLOPS per CP:

C-90 FLOPS/CP = (244 MFLOPS/sec) * (1 sec/240 MCP) = 1.02

Y-MP FLOPS/CP = (97 MFLOPS/sec) * (1 sec/167 MCP) = 0.56

The C-90 CPU delivers more FLOPS/CP because each CPU has double-piped add and multiply functional units, i.e., each CPU can deliver a peak 4 FLOPS/CP. The Y-MP CPU add and multiply functional units were single-piped and the CPU could deliver a peak of 2 FLOPS/CP.

The C-90 CPUs also perform significantly more total operations (2.6 every CP on the C-90 vs. 1.4 every CP on the Y-MP). Since both workloads experience significant periods (about 65% of CPU time) of hold issue, the instruction processor is able to ensure that other operations are occurring during these periods.

The measured vectorization of 92% is higher than the 86% levels measured for the Y-MP in 1990 [2]. Comparison of C-90 HPM output generated for specific codes with the 1990 Y-MP HPM output for the same codes indicates the C-90 codes performed a larger fraction of vector operations. The effect is most noticeable for codes with lower than average vector fractions. The higher vectorization appears due to improvements in Cray compiler technology.

A C-90 memory reference experienced an average delay of 0.37 CP compared to 1.05 CP required to deliver the variable from memory. Previous measurements indicated that a Y-MP memory reference experienced a delay of 0.27 CP. C-90 memory hold issue delays and C-90 memory conflict delays exceeded those of the Y-MP. The C-90 delay increase may be due to the greater number of memory references per CP, i.e., 1.02 C-90 references per CP vs. 0.56 Y-MP references per CP. For both machines, the memory delays are small compared to the memory access time.

The CPUs for both architectures have a load/store memory bandwidth of 3 references/CP and the HPM measurements indicate the NAS workloads do not challenge this bandwidth.

The system availability, defined here to be the percent of time the CPU is executing user programs (as opposed to being in system routines or idle) is about 86% for both machines. This value includes periods of heavy interactive usage as well as periods of mostly batch usage.

5.0 Conclusions

The 3.4 GFLOP system throughput rate of the C-90 on the NAS workload in its first quarter has exceeded, on a daily basis, the 3.0 GFLOP rate specified in the NAS Program Plan. The 0.6 GFLOP rate of the NAS Y-MP, while a very successful machine, exceeded the 1.0 GFLOP NAS target rate only occasionally. The initial C-90 performance is thus very encouraging.

The 2Q93 HPM measurements disclosed no obvious resource bottlenecks. While memory utilization for both machines was well over 90%, the current NAS workload does not challenge the computing power of the C-90 because the workload vector lengths are less than half of the hardware maximum. The C-90 system was also fairly lightly used because this quarter was the first quarter of operation and because NAS users also had access to the Y-MP.

Although the NAS workloads strongly under utilize the massive I/O capacity of both machines, individual users employing the SSD may be making good use of the Cray I/O capability.

6.0 Acknowledgment

Thanks to David Barkai and Ken Stevens for reviewing this paper.

7.0 References

[1] Cray Research, Inc. UNICOS Performance Utilities Reference Manual, SR-2040 7.0, 1992.

[2] Bergeron, R.J. "Performance Analysis of the NAS Y-MP Workload", NASA Ames Research Center, NAS Systems Division Report RND-90-009, December, 1990.

[3] Jain, R. "The Art of Computer Systems Performance Analysis", *Wiley Professional Computing*, 1991.