NAS Parallel Benchmark Results 10-93

David H. Bailey, Eric Barszcz, Leonardo Dagum and Horst D. Simon

RNR Technical Report RNR-93-016

October 27, 1993

Abstract

The NAS Parallel Benchmarks have been developed at NASA Ames Research Center to study the performance of parallel supercomputers. The eight benchmark problems are specified in a "pencil and paper" fashion. In other words, the complete details of the problem to be solved are given in a technical document, and except for a few restrictions, benchmarkers are free to select the language constructs and implementation techniques best suited for a particular system.

This paper presents performance results of various systems using the NAS Parallel Benchmarks. These results represent the best results that have been reported to us for the specific systems listed.

Bailey and Barszcz are employees of NASA Ames Research Center. Dagum and Simon are employees of Computer Science Corp., and their work was funded by the NASA Ames Research Center under contract NAS2-12961. Mailing address for all authors: NASA Ames Research Center, Mail Stop T045-1, Moffett Field, CA 94035.

1 Introduction

The Numerical Aerodynamic Simulation (NAS) Program, which is based at NASA Ames Research Center, is dedicated to advance the science of computational aerodynamics. One key goal of the NAS organization is to demonstrate by the year 2000 an operational computing system capable of simulating an entire aerospace vehicle system within a computing time of one to several hours. It is currently projected that the solution of this grand challenge problem will require a computer system that can perform scientific computations at a sustained rate approximately one thousand times faster than 1990 generation supercomputers. Most likely such a computer system will employ hundreds or even thousands of processors operating in parallel.

In order to objectively measure the performance of various highly parallel computer systems and to compare them with conventional supercomputers, we along with other scientists in our organization have devised the NAS Parallel Benchmarks (NPB). Note that the NPB are distinct from the High Speed Processor (HSP) benchmarks and procurements. The HSP benchmarks are used for evaluating production supercomputers for procurement, whereas the NPB are for studying massively parallel processor (MPP) systems not necessarily tied to a procurement.

The NPB are a set of eight benchmark problems, each of which focuses on some important aspect of highly parallel supercomputing for aerophysics applications. Some extension of Fortran or C is required for implementations, and reasonable limits are placed on the usage of assembly code and the like, but otherwise programmers are free to utilize language constructs that give the best performance possible on the particular system being studied. The choice of data structures, processor allocation and memory usage are generally left open to the discretion of the implementer.

The eight problems consist of five "kernels" and three "simulated computational fluid dynamics (CFD) applications". Each of these is defined fully in [?]. The five kernels are relatively compact problems, each of which emphasizes a particular type of numerical computation. Compared with the simulated CFD applications, they can be implemented fairly readily and provide insight as to the general levels of performance that can be expected on these specific types of numerical computations.

The simulated CFD applications, on the other hand, usually require more effort to implement, but they are more indicative of the types of actual data movement and computation required in state-of-the-art CFD application codes. For example, in an isolated kernel a certain data structure may be very efficient on a certain system, and yet this data structure would be inappropriate if incorporated into a larger application. By comparison, the simulated CFD applications require data structures and implementation techniques that are more typical of real CFD applications.

Space does not permit a complete description of these benchmark problems. A more detailed description of these benchmarks, together with the rules and restrictions associated with the benchmarks, may be found in [?]. The full specification of the benchmarks is given in [?].

Sample Fortran programs implementing the NPB on a single processor system are available as an aid to implementors. These programs, as well as the benchmark document itself, are available from the following address: NAS Systems Division, Mail Stop 258-6, NASA Ames Research Center, Moffett Field, CA 94035, attn: NAS Parallel Benchmark Codes. The sample codes are provided on Macintosh floppy disks and contain the Fortran source codes, "README" files, input data files, and reference output data files for correct implementations of the benchmark problems. These codes have been validated on a number of computer systems ranging from conventional workstations to supercomputers.

In the following, each of the eight benchmarks will be briefly described, and then the best performance results we have received to date for each computer system will be given in Tables 2 through 9. These tables include run times and performance ratios. The performance ratios compare individual timings with the current best time on that benchmark achieved on one processor of a Cray Y-MP. The run times in each case are elapsed time of day figures, measured in accordance with the specifications given in [**?**].

There are now two standard sizes for the NAS Parallel Benchmarks; these will be referred to as the Class A and Class B size problems. The nominal benchmark sizes for the Class A and Class B are listed in Tables 1a and 1b respectively. These tables also give the standard floating point operation (flop) counts for the two classes of problems. We insist that those wishing to compute performance rates in millions of floating point operations per second (Mflop/s) use these standard flop counts. Table 1a also contains Mflop/s rates calculated in this manner for the current fastest implementation on one processor of the Cray Y-MP. Note, however, that in Tables 2 through 9, performance rates are *not* cited in Mflop/s; we present instead the actual run times (and, equivalently, the performance ratios). We suggest that these, and not Mflop/s, be examined when comparing different systems and implementations.

With the exception of the Integer Sort benchmark, these standard flop counts were determined by using the hardware performance monitor on a Cray Y-MP, and we believe that they are close to the minimal counts required for these problems. In the case of the Integer Sort benchmark, which does not involve floating-point operations, we selected a value approximately equal to the number of integer operations required, in order to permit the computation of performance rates analogous to Mflop/s rates. We reserve the right to change these standard flop counts in the future if deemed necessary.

The NAS organization reserves the right to verify any NPB results that are submitted to us. We may, for example, attempt to run the submitter's code on another system of the same configuration as that used by the submitter. In those instances where we are unable to reproduce the submitter's supplied results (allowing a 5% tolerance) our policy is to alert the submitter of the discrepancy and allow him or her until the next release of this report to resolve the discrepancy. If the discrepancy is not resolved to our satisfaction, then our own observed results, and not the submitter's results, will be reported. This policy will apply to all results we receive and publish.

Whenever possible, we have tried to credit the actual individuals and organizations who have contributed the performance results cited in the tables. In these citations, NAS

Benchmark Name	Abbreviation	Nominal Size	Operation Count ($\times 10^9$)	Mflop/s on Y-MP/1
Embarrassingly Parallel	EP	2^{28}	26.68	211
Multigrid	MG	256^3	3.905	176
Conjugate Gradient	CG	$14,000$	1.508	127
3-D FFT PDE	FT	$256^2 \times 128$	5.631	196
Integer Sort	IS	$2^{23} \times 2^{19}$	0.7812	68
LU Simulated CFD Application	LU	64^3	64.57	194
SP Simulated CFD Application	SP	64^3	102.0	216
BT Simulated CFD Application	BT	64^3	181.3	229

Table 1a: Standard Operation Counts and YMP/1 Mflop/s for Class A Size Problems

Benchmark Name	Abbreviation	Nominal Size	Operation Count ($\times 10^9$)	Mflop/s on C90
Embarrassingly Parallel	EP	2^{30}	100.9	na
Multigrid	MG	256^3	9.405	na
Conjugate Gradient	CG	75000	54.89	na
3-D FFT PDE	FT	512×256^2	71.37	na
Integer Sort	IS	$2^{25} \times 2^{21}$	3.150	na
LU Simulated CFD Application	LU	102^3	319.6	na
SP Simulated CFD Application	SP	102^3	447.1	na
BT Simulated CFD Application	BT	102^3	721.5	na

Table 1b: Standard Operation Counts and C90 Mflop/s for Class B Size Problems

denotes the NAS Applied Research Branch at NASA Ames (including both NASA civil servants and Computer Science Corp. contractors); RIACS denotes the parallel systems division of the Research Institute for Advanced Computer Science, which is located at NASA Ames; BBN denotes Bolt, Beranek and Newman; BCS denotes Boeing Computer Services; CRI denotes Cray Research, Inc.; KSR denotes Kendall Square Research Corp., Intel denotes the Supercomputer Systems Division of Intel Corp.; MasPar denotes MasPar Computer Corp.; Meiko denotes Meiko Scientific Corp.; and TMC denotes Thinking Machines, Inc. Where no individual citation is made for a specific model, the results are due to vendor staff.

This paper reports benchmark results on the following systems: TC2000 by Bolt, Beranek and Newman (BBN); YMP, EL, C90, and T3D by Cray Research Inc. (CRI); Paragon and iPSC/860 by Intel; SP-1 by International Business Machines (IBM); KSR1 by Kendall Square Research; MP-1 and MP-2 by MasPar Computer Corp.; CM-2, CM-200, and CM-5 by Thinking Machines Corp. (TMC); CS-1 by Meiko Scientific; nCUBE-2 by nCUBE; and clusters of distributed workstations including Sparcstation's by Sun; RS/6000's by IBM; and 4D25's by SGI. Entries in the tables are ordered alphabetically by vendor, except for distributed workstation results which appear last.

Unfortunately, the limited space in this report does not permit discussion of the methods used in any of these implementations. However, references to technical papers describing these methods have been included whenever such papers are available. In particular, details of the implementation of these benchmarks on the TC2000, the CM2, the CM200, and the SP-1 may be found in [?, ?, ?]. General discussion on architectural requirements for the benchmarks may be found in [?]. Readers are referred to these documents for full details.

This report includes a number of new results including previously unpublished Intel Paragon results, Thinking Machines CM-5 results, Cray EL98, C90 and T3D results, and IBM SP-1 results. Some results using Parallel Virtual Machine (PVM) on distributed workstations which have not previously appeared in this report are also included. In quite a few other instances, results are improved from previous listings, reflecting improvements both in compilers and implementations. Efforts are currently underway to port the NAS Parallel Benchmarks on other systems, and we hope to have more results in the future.

2 Kernel Results

2.1 Embarrassingly Parallel (EP) Benchmark

The first of the five kernel benchmarks is an "embarrassingly parallel" problem. In this benchmark, two-dimensional statistics are accumulated from a large number of Gaussian pseudorandom numbers, which are generated according to a particular scheme that is well-suited for parallel computation. This problem is typical of many "Monte-Carlo" applications. Since it requires almost no communication, in some sense this benchmark provides an estimate of the upper achievable limits for floating point performance on a particular system.

Results for the embarrassingly parallel benchmark are shown in Table 2. Not all systems exhibit high rates on this problem. This appears to stem from the fact that this benchmark requires references to several mathematical intrinsic functions, such as the Fortran routines AINT, SQRT, and LOG, and evidently these functions are not highly optimized on some systems.

Intel iPSC/860 and Paragon results are due to J. Baugh of Intel. CM-2, CM-200 and CM-5 results are due to J. Richardson of TMC. Distributed workstation results are due to S. White of Emory University [?] except for the SGI results which are due to D. Browning of the NAS System Development branch. The "Mixed-A" computer system consisted of 16 Sun Sparc 1's, one Sun IPC, one Sun Sparc2, 11 Sun SLC's, three IBM RS/6000 model 550's, one IBM RS/6000 model 530, and one NeXT machine. All distributed systems results for EP are for PVM 2.4 and Ethernet.

2.2 Multigrid (MG) Benchmark

The second kernel benchmark is a simplified multigrid kernel, which solves a 3-D Poisson PDE. This problem is simplified in the sense that it has constant rather than variable coefficients as in a more realistic application. This code is a good test of both short and long distance highly structured communication.

Results for this benchmark are shown in Table 3. Intel iPSC/860 and Paragon results are due to J. Patterson of BCS. CM-2 and CM-200 results are due to J. Richardson at TMC. Distributed workstation results are due to S. White of Emory University [?] using PVM 2.4 and Ethernet except where noted otherwise.

2.3 The Conjugate Gradient Benchmark

In this benchmark, a conjugate gradient method is used to compute an approximation to the smallest eigenvalue of a large, sparse, symmetric positive definite matrix. This kernel is typical of unstructured grid computations in that it tests irregular long distance communication and employs sparse matrix vector multiplication.

The irregular communication requirement of this benchmark is evidently a challenge for all systems. Results are shown in Table 4. CM-2 results are due to J. Richardson of TMC. Intel iPSC/860 and nCUBE-2 results are by B. Hendrickson, R. Leland, and S. Plimpton of Sandia National Laboratory[?]. Paragon results are due to R. van de Geijn of U.T. Austin and John Lewis of BCS[?]. Cray EL and C90 results are due to M. Zagha of Carnegie Mellon University. Distributed workstation results are due to S. White of Emory University [?] using PVM 2.4 and Ethernet except where noted otherwise.

2.4 3-D FFT PDE (FT) Benchmark

In this benchmark a 3-D partial differential equation is solved using FFTs. This kernel performs the essence of many "spectral" codes. It is a good test of long-distance commu-

Computer System	Date Received	No. Proc.	Time (sec.)	Ratio to Y-MP/1
BBN TC2000	Dec 91	64	284.0	0.44
Cray Y-MP	Aug 92	1	126.2	1.00
		8	15.9	7.95
Cray EL	Sep 93	1	404.7	0.31
		4	104.1	1.21
		8	54.9	2.30
Cray C-90	Sep 93	1	39.5	3.20
		4	9.9	12.78
		16	2.5	50.27
Cray T3D	Sep 93	32	19.7	6.41
		64	9.7	12.97
		128	4.8	26.24
IBM SP-1	Sep 93	8	46.9	2.69
		16	25.4	5.35
		32	11.9	10.61
		64	6.07	20.79
Intel iPSC/860	May 92	32	102.7	1.23
		64	51.4	2.46
		128	25.7	4.91
Intel Paragon	Oct 93	64	33.5	3.77
		128	17.0	7.42
	Sep 93	256	9.9	12.75
		512	5.2	24.42
Kendall Square KSR1	Nov 92	32	69.8	1.81
		64	34.9	3.62
		96	23.4	5.39
		128	18.1	6.97
MasPar MP-1	Aug 92	4K	248.0	0.51
		16K	69.3	1.82
MasPar MP-2	Nov 92	16K	22.4	5.63
Meiko CS-1	Aug 92	16	116.8	1.08

Table 2: Results of the Embarrassingly Parallel (EP) Benchmark (cont'd)

Computer System	Date Received	No. Proc.	Time (sec.)	Ratio to Y-MP/1
Thinking Machines CM-2	Oct 91	8K	126.6	1.00
		16K	63.9	1.97
		32K	33.7	3.74
		64K	18.8	6.71
Thinking Machines CM-200	Oct 91	8K	76.9	1.64
		16K	39.2	3.22
		32K	20.7	6.10
		64K	10.9	11.58
Thinking Machines CM-5	Nov 92	16	42.4	2.98
		32	21.5	5.88
		64	10.9	11.62
		128	5.4	23.49
		256	2.7	46.84
		512	1.4	90.47
PVM Sparcs (Ethernet)	Sep 93	16	1670.0	0.08
PVM RS/6000/550 (Ethernet)	Sep 93	4	890.0	0.14
PVM Mixed-A (Ethernet)	Sep 93	34	494.0	0.26
PVM SGI 4D25 (Ethernet)	Sep 93	4	2536.4	0.05

Table 2: (cont'd) Results of the Embarrassingly Parallel (EP) Benchmark

Computer System	Date Received	No. Proc.	Time (sec)	Ratio to Y-MP/1
Cray Y-MP	Aug 92	1	22.22	1.00
		8	2.96	7.51
Cray EL	Aug 92	1	89.19	0.25
		4	27.94	0.80
		8	22.30	0.95
Cray C-90	Aug 92	1	8.65	2.57
		4	2.42	9.18
		16	0.96	23.14
Cray T3D	Oct 93	64	3.03	7.33
		128	1.57	14.15
Intel iPSC/860	Aug 92	128	8.6	2.58
Intel Paragon	Sep 93	128	5.8	3.84
		256	4.2	5.30
Kendall Square KSR1	Nov 92	32	20.6	1.08
MasPar MP-1	Aug 92	16K	12.0	1.85
MasPar MP-2	Nov 92	16K	4.36	5.10
Meiko CS-1	Aug 92	16	42.8	0.52
Thinking Machines CM-2	Dec 91	16K	45.8	0.49
		32K	26.0	0.85
		64K	14.1	1.58
Thinking Machines CM-200	Dec 91	16K	30.2	0.74
		32K	17.2	1.29
Thinking Machines CM-5	Aug 93	32	19.5	1.14
		64	10.9	2.04
		128	6.1	3.64
PVM RS/6000/550 (Ethernet)	Sep 93	4	293.0	0.08
PVM RS/6000/560 (FDDI)	Sep 93	4	184.0	0.12
	Sep 93	8	110.4	0.20

Table 3: Results of the Multigrid (MG) Benchmark

Computer System	Date Received	No. Proc.	Time (sec.)	Ratio to Y-MP/1
BBN TC2000	Dec 91	40	51.4	0.23
Cray Y-MP	Aug 92	1	11.92	1.00
		8	2.38	5.01
Cray EL	Sep 93	1	45.24	0.26
		4	14.29	0.83
		8	10.14	1.18
Cray C-90	Sep 93	1	3.55	3.36
		4	0.96	12.42
		16	0.34	35.06
Cray T3D	Sep 93	16	21.89	0.54
		32	11.41	1.04
		64	6.14	1.94
		128	3.74	3.19
Intel iPSC/860	Sep 93	128	7.0	1.71
Intel Paragon	Sep 93	32	27.5	0.43
		64	16.5	0.72
		128	12.4	0.96
		256	11.1	1.08
Kendall Square KSR1	Nov 92	32	21.7	0.55
MasPar MP-1	Aug 92	4K	64.5	0.18
		16K	14.6	0.82
MasPar MP-2	Nov 92	16K	11.0	1.08
Meiko CS-1	Aug 92	16	67.5	0.18
nCUBE-2	Dec 92	1024	6.1	1.96
Thinking Machines CM-2	Mar 92	8K	25.6	0.47
		16K	14.1	0.85
		32K	8.8	1.35
Thinking Machines CM-200	Mar 92	8K	15.0	0.79
Thinking Machines CM-5	Aug 93	32	20.7	0.58
		64	10.6	1.12
		128	6.2	1.92
PVM RS/6000/550 (Ethernet)	Sep 93	4	203.2	0.06
PVM RS/6000/560 (FDDI)	Sep 93	4	81.5	0.15

Table 4: Results of the Conjugate Gradient (CG) Benchmark

nication performance.

The rules of the NAS Parallel Benchmarks specify that assembly-coded, library routines may be used to perform matrix multiplication and one-dimensional, two-dimensional or three-dimensional FFTs. Thus this benchmark is somewhat unique in that computational library routines may be legally employed.

Results are shown in Table 5. Intel iPSC/860 and Paragon results are due to E. Kushner of Intel. CM-2 and CM-200 results are due to J. Richardson of TMC.

2.5 Integer Sort (IS) Benchmark

This benchmark tests a sorting operation that is important in "particle method" codes. This type of application is similar to "particle in cell" applications of physics, wherein particles are assigned to cells and may drift out. The sorting operation is used to reassign particles to the appropriate cells. This benchmark tests both integer computation speed and communication performance.

This problem is unique in that floating point arithmetic is not involved. Significant data communication, however, is required. Results are shown in Table 6. Intel iPSC/860 and Paragon results are due to to J. Baugh of Intel. CM-2, CM-200 and MasPar results use a library sorting routine. Cray Y-MP results are due to CRI. Cray C-90 and EL results are due to M. Zagha of Carnegie Mellon University using a radix sort optimized for interleaved memories [?].

3 Simulated CFD Application Benchmarks

The three simulated CFD application benchmarks are intended to accurately represent the principal computational and data movement requirements of modern CFD applications.

The first of these is the called the lower-upper diagonal (LU) benchmark. It does not perform a LU factorization but instead employs a symmetric successive over-relaxation (SSOR) numerical scheme to solve a regular-sparse, block (5×5) lower and upper triangular system. This problem represents the computations associated with a newer class of implicit CFD algorithms, typified at NASA Ames by the code "INS3D-LU". This problem exhibits a somewhat limited amount of parallelism compared to the next two. Discussion of the serial algorithm underlying this benchmark may be found in [?]. Discussion of the parallel algorithms may be found in [?].

The second simulated CFD application is called the scalar pentadiagonal (SP) benchmark. In this benchmark, multiple independent systems of non-diagonally dominant, scalar pentadiagonal equations are solved. The third simulated CFD application is called the block tridiagonal (BT) benchmark. In this benchmark, multiple independent systems of non-diagonally dominant, block tridiagonal equations with a 5×5 block size are solved.

SP and BT are representative of computations associated with the implicit operators of CFD codes such as "ARC3D" at NASA Ames. SP and BT are similar in many respects,

Computer System	Date Received	No. Proc.	Time (sec.)	Ratio to Y-MP/1
Cray Y-MP	Aug 92	1	28.77*	1.00
		8	4.19*	6.87
Cray EL	May 93	1	105.1*	0.27
		4	27.9*	1.03
		8	18.5*	1.56
Cray C-90	Aug 92	1	10.28*	2.80
		4	2.58*	11.20
		16	0.91*	31.60
Cray T3D	Oct 93	64	4.53*	6.35
		128	2.33*	12.35
IBM SP-1	Sep 93	8	48.1*	0.60
		16	25.4*	1.13
		32	13.3*	2.16
		64	7.4*	3.89
Intel iPSC/860	Dec 91	64	20.9*	1.37
	Apr 92	128	9.7*	2.96
Intel Paragon	Aug 93	64	9.2*	3.12
		128	5.1*	5.60
Kendall Square KSR1	Dec 92	32	13.6*	2.12
		64	8.4*	3.43
MasPar MP-1	Aug 92	16K	18.3*	1.57
MasPar MP-2	Nov 92	16K	8.0*	3.60
Meiko CS-1	Aug 92	16	170.0*	0.17
Thinking Machines CM-2	Dec 91	16K	37.0*	0.78
		32K	18.2*	1.58
		64K	11.4*	2.52
Thinking Machines CM-200	Dec 91	8K	45.6*	0.63
Thinking Machines CM-5	Aug 93	32	14.9*	1.93
		64	7.9*	3.64
		128	6.6*	4.36

Table 5: Results of the 3-D FFT PDE (FT) Benchmark (* indicates library result).

Computer System	Date Received	No. Proc.	Time (sec.)	Ratio to Y-MP/1
Cray Y-MP	Aug 92	1	11.46	1.00
		8	1.85	6.19
Cray EL	Sep 93	1	43.76	0.26
		4	12.99	0.88
		8	8.45	1.35
Cray C-90	Sep 93	1	3.33	3.44
		4	0.85	13.46
		16	0.27	42.38
Cray T3D	Sep 93	32	7.04	1.62
		64	3.42	3.35
		128	1.75	6.54
IBM SP-1	Sep 93	8	11.2	1.03
		16	8.0	1.44
		32	6.7	1.70
		64	6.6	1.73
Intel iPSC/860	May 92	32	25.7	0.45
		64	17.3	0.66
		128	13.6	0.84
Intel Paragon	Sep 93	32	21.5	0.53
		64	15.1	0.75
		128	13.3	0.86
Kendall Square KSR1	Nov 92	32	40.2	0.29
MasPar MP-1	Jan 93	16K	11.5*	1.00
MasPar MP-2	Jan 93	16K	7.7*	1.49
Meiko CS-1	Aug 92	16	62.7	0.18
Thinking Machines CM-2	Dec 91	16K	35.8*	0.32
		32K	21.0*	0.55
		64K	14.9*	0.77
Thinking Machines CM-200	Dec 91	64K	5.7*	2.01
Thinking Machines CM-5	Aug 93	32	43.1	0.27
		64	24.2	0.47
		128	12.0	0.96

Table 6: Results of the Integer Sort (IS) Benchmark (* indicates library result).

but there is a fundamental difference with respect to the communication to computation ratio. Discussion of the serial algorithm underlying this benchmark may be found in [?].

Performance figures for the three simulated CFD applications are shown in Tables 7, 8 and 9. Timings are cited as complete run times, in seconds, as with the other benchmarks. A complete solution of the LU benchmark requires 250 iterations. For the SP benchmark, 400 iterations are required. For the BT benchmark, 200 iterations are required.

For LU, credits are as follows: iPSC/860 and CM-2 results are due to S. Weeratunga, R. Fatoohi, E. Barszcz and V. Venkatakrishnan of NAS; CM-5 results are due to J. Richardson and D. Sandee of TMC; MP-1 and MP-2 results are due to J. McDonald of MasPar; Intel Paragon results are due to T. Phung of Intel; KSR1 results are due to S. Breit of KSR; SP-1 results are due to V. Naik of IBM. For SP, credits are as follows: CM-2 results employ a library scalar pentadiagonal solver; CM-5 results are due to J. Richardson and D. Sandee of TMC; iPSC/860 results are due to J. Patterson of BCS; Paragon results are due to T. Phung of Intel; MP-1 and MP-2 results are due to J. McDonald of MasPar; KSR1 results are due to S. Breit of KSR; SP-1 results are due to V. Naik of IBM. For BT, credits are as follows: CM-2 and CM-200 results employ a library block tridiagonal solver; CM-5 results are due to J. Richardson and D. Sandee of TMC; iPSC/860 results are due to J. Patterson of BCS; Paragon results are due to T. Phung of Intel; MP-1 and MP-2 results are due to J. McDonald of MasPar; KSR1 results are due to S. Breit of KSR; SP-1 results are due to V. Naik of IBM.

4 Sustained Performance Per Dollar

One aspect of the relative performance of these systems has not been addressed so far, namely the differences in price between these systems. One should not be too surprised that the Cray C-90 system, for example, exhibits superior performance rates on these benchmarks, since its current list price is much greater than that of any other system tested.

One way to compensate for these price differences is to compute sustained performance per million dollars, i.e. the performance ratio figures shown in Tables 2 through 9 divided by the list price in millions. Some figures of this type are shown in Table 10 for two of the benchmarks (the EP and the SP benchmarks) for the most recent of the systems tested. The table includes the list price of the minimal system (in terms of memory per node, disk space, etc.) required to run the full NPB as implemented by the vendor. These prices were provided by the vendors and include any associated software costs (i.e. operating system, compilers, scientific libraries as required, etc.) but do not include maintenance. Prices for October, 1993 are as follows: Cray C90 with 16 processors, 256 Mwords is $30.65 million; Cray EL98 with 1 GB of memory and 12 GB disk is $1.11 million; Cray T3D with 128 nodes, 16 MB/node, no front end is $5.55 million; IBM SP-1 with 64 nodes, 64 MB/node, 64 GB disk is $2.66 million; Intel Paragon with 128 nodes, 16 MB/node, 5 GB disk, one 32 MB service node is $3.5 million; Kendall Square KSR1 with 128 nodes, 32 MB/node,

Computer System	Date Received	No. Proc.	Time (sec.)	Ratio to Y-MP/1
BBN TC2000	Dec 91	62	3032.0	0.11
Cray Y-MP	Aug 92	1	333.5	1.00
		8	49.5	6.74
Cray EL	Aug 92	1	1449.0	0.23
		4	522.3	0.64
		8	351.6	0.95
Cray C-90	Aug 92	1	157.6	2.12
		4	43.9	7.59
		16	17.6	18.93
Cray T3D	Oct 93	32	195.3	1.71
		64	102.0	3.27
		128	55.4	6.02
IBM SP-1	Sep 93	8	516.1	0.65
		16	295.2	1.13
		32	170.8	1.95
		64	102.5	3.25
Intel iPSC/860	Mar 91	64	690.8	0.48
		128	442.5	0.75
Intel Paragon	Aug 93	64	523.8	0.64
		128	378.0	0.88
Kendall Square KSR1	Nov 92	32	1041.3	0.32
MasPar MP-1	Aug 92	4K	1580.0	0.21
MasPar MP-2	Nov 92	4K	463.5	0.72
Meiko CS-1	Aug 92	16	2937.0	0.11
Thinking Machines CM-2	Mar 91	8K	1307.0	0.26
		16K	850.0	0.39
		32K	546.0	0.61
Thinking Machines CM-5	Aug 93	32	418.0	0.80
		64	272.0	1.23
		128	171.0	1.95

Table 7: Results for the LU Simulated CFD Application

Computer System	Date Received	No. Proc.	Time (sec.)	Ratio to Y-MP/1
BBN TC2000	Dec 91	112	880.0	0.54
Cray Y-MP	Aug 92	1	471.5	1.00
		8	64.6	7.30
Cray EL	Aug 92	1	2025.7	0.23
		4	601.9	0.78
		8	488.4	0.97
Cray C-90	Aug 92	1	184.70	2.55
		4	49.74	9.48
		16	13.06	36.10
Cray T3D	Sep 93	32	312.1	1.51
		64	159.9	2.95
		128	82.8	5.70
IBM SP-1	Sep 93	8	563.0	0.84
		16	329.6	1.43
		32	200.5	2.35
		64	121.5	3.88
Intel iPSC/860	Aug 92	64	667.3	0.71
		128	449.5	1.05
Intel Paragon	Oct 93	64	274.5	1.72
		128	179.9	2.62
		256	161.2	2.92
Kendall Square KSR1	Dec 92	32	377.7	1.25
		64	228.8	2.06
		96	170.2	2.77
		128	150.0	3.14
MasPar MP-1	Aug 92	4K	1772	0.27
MasPar MP-2	Nov 92	4K	615	0.77
Meiko CS-1	Aug 92	16	2975	0.16
Thinking Machines CM-2	Dec 91	16K	1444.0*	0.33
		32K	917.0*	0.51
		64K	640.0*	0.74
Thinking Machines CM-5	May 93	32	289.0	1.63
		64	170.0	2.77
		128	119.0	3.96

Table 8: Results for the SP Simulated CFD Application (* indicates library result).

Computer System	Date Received	No. Proc.	Time (sec.)	Ratio to Y-MP/1
BBN TC2000	Dec 91	112	1378.0	0.58
Cray Y-MP	Aug 92	1	792.4	1.00
		8	114.0	6.95
Cray EL	May 93	1	3832.8	0.21
		4	1090.2	0.73
		8	764.1	1.04
Cray C-90	Aug 92	1	356.9	2.22
		4	96.1	8.25
		16	28.4	27.91
Cray T3D	Sep 93	32	335.5	2.36
		64	153.7	5.16
		128	78.9	10.04
IBM SP-1	Sep 93	8	884.4	0.90
		16	481.6	1.65
		32	267.0	2.97
		64	152.6	5.19
Intel iPSC/860	Aug 92	64	714.7	1.11
		128	414.3	1.91
Intel Paragon	Sep 93	64	242.4	3.27
		128	139.4	5.68
		256	97.6	8.12
Kendall Square KSR1	Dec 92	32	439.0	1.81
		64	239.4	3.31
		96	167.9	4.72
		128	134.5	5.89
MasPar MP-1	Aug 92	4K	2396.0	0.33
MasPar MP-2	Nov 92	4K	789.0	1.00
Meiko CS-1	Aug 92	16	2984.0	0.27
Thinking Machines CM-2	Dec 91	16K	1118.0*	0.71
		32K	634.0*	1.25
		64K	370.0*	2.14
Thinking Machines CM-200	Dec 91	16K	832.0*	0.95
		32K	601.0*	1.32
Thinking Machines CM-5	May 93	32	284.0	2.79
		64	175.0	4.50
		128	119.0	6.66

Table 9: Results for the BT Simulated CFD Application (* indicates library result).

10 GB disk is $7.14 million; MasPar MP-2 with 16K processors, 1 GB memory, DEC front end is $ 1.61 million; TMC CM-5 with 128 nodes, 32 MB/node, 5 GB disk, Sparc front end is $4.92 million. Note that some vendor standard configurations may include substantially more hardware than required for the benchmarks (for example, the IBM SP-1). Also note that the KSR1 list price is from July 1992 and expected to change dramatically in the near future. Finally, be aware that list prices are similar to peak performance in that they are guaranteed not to be exceeded.

B'mark	Computer System	No. Proc.	Ratio to Y-MP/1	Nominal cost ($)	Perf. per million $
EP	Cray C-90	16	50.27	30.65M	1.64
	Cray EL	8	2.30	1.11M	2.07
	Cray T3D	128	26.24	5.55M	4.73
	IBM SP-1	64	20.79	2.66M	7.82
	Intel Paragon	128	7.42	3.51M	2.11
	Kendall Square KSR1	128	6.97	7.14M	0.98
	MasPar MP-2	16K	5.63	1.61M	3.50
	Thinking Machines CM-5	128	23.49	4.92M	4.77
SP	Cray C-90	16	36.10	30.65M	1.18
	Cray EL	8	0.97	1.11M	0.87
	Cray T3D	128	5.70	5.55M	1.03
	IBM SP-1	64	3.88	2.66M	1.46
	Intel Paragon	128	2.62	3.51M	0.75
	Kendall Square KSR1	128	3.14	7.14M	0.44
	MasPar MP-2	4K	0.77	0.43M	1.79
	Thinking Machines CM-5	128	3.96	4.92M	0.80

Table 10: Approximate Sustained Performance Per Dollar

5 Discussion and Conclusions

With some algorithmic experimentation and tuning, respectable NPB performance rates have been achieved on several multiprocessor systems. Except for the EP benchmark, the 16 processor Cray C-90 system is still the highest performing system tested. It also remains the highest priced system tested. MPP systems of comparable price do exist however the Class A size benchmarks do not offer sufficient parallelism to do justice to such systems. With the possible exception of the Cray T3D, message latency on current systems is such that beyond 128 nodes nodes the performance on the Class A size benchmarks begins to severely degrade. The Class B size benchmarks, however, should offer parallelism up to 512 nodes on current systems.

It is well understood that scaling problem size with machine size often leads to nearly linear speed up. However, in real world applications problem sizes are not so flexible. The simulated CFD bencharks in the NPB reflect the operations of CFD codes on structured grids. In practice, a single structured grid rarely exceeds one million points. Usually the components of a complicated geometry must be gridded individually in "grid blocks" and the calculation must be carried out on a per block basis. For this reason, it is unlikely that further scaling of the CFD benchmarks will proceed through a direct scaling of the grid.

For historical reasons, the Class A performance results are still provided in terms of equivalent single processor Cray Y-MP performance. The Class B performance results, however, will be provided in terms of equivalent single processor Cray C90 performance. The implication should be clear. Traditional vector supercomputers continue to improve and today's MPP systems must compete with the latest vector systems. In the following, the parallel system performance is discussed in terms of the "equivalent" number of C90 processors. Note, however, that the single processor C90 results were all obtained on a 16 processor system with 256 MW of multiported memory, and as such is not necessarily representative of a "typical" single processor C90 system.

The parallel systems continue to show excellent performance on EP and MG, as expected. Somewhat disappointing however, are the PVM results for EP. Clearly there is room for improvement even on the Ethernet connected systems. The PVM implementations are still evolving, and it is difficult to draw any conclusions from these PVM results because the distributed systems tested were not consistent across benchmarks.

The CG benchmark is challenging for all the parallel systems. In addition, newly reported and much improved EL and C90 results further diminish the relative performance of the parallel systems. None of the MPP systems tested showed better performance than a single processor C90 for the CG benchmark. Parallel algorithms for CG are still evolving [?, ?], and implementations of the newer algorithms have appeared only on the nCUBE-2 and the iPSC/860. For this reason, an 128 node iPSC/860 is outperforming a comparably sized Paragon by almost a factor of 2 on the CG benchmark, whereas on LU for example (another difficult parallel benchmark) the Paragon is superior.

On the FT benchmark, the T3D is showing the best performance. For this benchmark, a 64 node T3D is roughly equivalent to two C90 processors, whereas 64 node SP-1, Paragon, KSR1, CM-5, and 16K processor MP-2 systems all achieve only the performance of roughly one C90 processor. The CM-5 is showing poor scalability beyond 64 nodes, and the same appears true for the KSR1 although data has been reported only to 64 nodes. There is no obvious reason for this result. One would expect better scalability from the FT benchmark since is transposed based like SP and BT but with a significantly larger grid and correspondingly greater parallelism. The allowed use of library routines, however, makes interpretation of the results more difficult.

As with the CG benchmark, not all the vendors have implemented the same IS algorithm thus making results on this benchmark difficult to interpret. Algorithmic improvements for the vector systems (see [?]) have led to improved results being reported for the C90 (and the EL) thus further diminishing the relative performance for the MPP systems.

Nevertheless, the T3D again is showing the best performance, with 128 nodes achieving roughly the equivalent of two C90 processors. The Cray T3D results are based on a cyclic sort algorithm, because the key values are nearly uniformly distributed on the higher order bits they can be distributed in a cyclic fashion across processors for a uniform load balance. Of the remaining MPP systems, a 64 node SP-1 and a 16K processor MP-2 succeed in achieving the equivalent of half a C90 processor, and the rest do poorer still.

The most challenging CFD benchmarkd for the MPP systems is the LU benchmark. The best parallel performance is achieved by an 128 node T3D with the equivalent of almost three C90 processors. At 64 nodes, the T3D and the SP-1 achieve similar performance (roughly 1.5 C90 processors). The remaining parallel systems do poorly in comparison.

The SP benchmark offers somewhat greater parallelism than LU, resulting in comparatively better parallel system performance. Surprisingly, the T3D does better on the LU than on the SP benchmark, however this is due to a more recent result for the LU than for the SP. Presumably the SP (and BT) results for the T3D will also improve in the future. For the SP benchmark, the 128 node T3D does slightly better than two C90 processors; the 64 node SP-1 and 128 node CM-5 are roughly equal in performance at 1.5 C90 processors; and the 128 node KSR1 and Paragon achieve 1.2 and 1.0 C90 processors respectively. Good performance on this benchmark tends to rely equally on good node performance and a fast transpose, at least up to 128 nodes on the Class A size problem. The only system that appears to scale much beyond 128 nodes for this size is the T3D.

The BT benchmark is very similar to SP but with a greater computation to communication ratio. All the parallel systems show good speed ups at least to 128 nodes and even up to 256 nodes in the case of the Paragon. The highest performing parallel systems are the 128 node T3D, 256 node Paragon, and 128 node CM-5 systems with 4.5, 3.7 and 3.0 equivalent C90 processors. The 64 node T3D and SP-1 systems, and the 128 node Paragon and KSR1 systems are each roughly equivalent to 2.5 C90 processors.

Of the MPP systems, the T3D is consistently achieving the greatest performance on a per node basis. The excellent performance demonstrated by the T3D proves that distributed memory architectures are quite suitable for general purpose scientific computing and not destined just to fill a niche in the field.

Unfortunately, when normalized by price, the T3D no longer is the leader. For the EP benchmark, the SP-1 is the price performance leader. For the SP benchmark the MP-2 achieves the best price performance, however this result is tempered by the fact that the MasPar implementation of the simulated CFD benchmarks does not scale beyond 4K processors. The SP-1 is a close second in price performance for the SP benchmark, and third is the C90. This last result may at first seem surprising given that vector supercomputers are traditionally viewed as expensive "grand challenge" systems. Closer examination, however, reveals that the C90 is benefiting from economies of scale. The vector systems are not scalable like the parallel systems, so a two processor C90 lists for substantially more than an eighth of the list on a 16 processor system (especially if the same memory system is retained). Therefore price performance of the vector systems benefits most when the problem size is fixed and the number of processors is increased, as is the case for the results in Table 10. In contrast, price performance of the MPP systems benefits most when the number of processors is fixed and the problem size is increased, or conversely, problem size is fixed and the number of processors is decreased. Neither of these cases appear in Table 10.

References

[1] D. H. Bailey, E. Barszcz, J. T. Barton, D. S. Browning, R. L. Carter, L. Dagum, R. A. Fatoohi, P. O. Frederickson, T. A. Lasinski, R. S. Schreiber, H. D. Simon, V. Venkatakrishnan, and S. K. Weeratunga, "The NAS Parallel Benchmarks", *Intl. Journal of Supercomputer Applications*, v. 5, no. 3 (Fall 1991), pp. 63 – 73.

[2] D. Bailey, J. Barton, T. Lasinski, and H. Simon, eds., "The NAS Parallel Benchmarks", NASA Technical Memorandum 103863, Ames Research Center, Moffett Field, CA 94035-1000, July 1993.

[3] E. Barszcz, R. Fatoohi, V. Venkatakrishnan, and S. Weeratunga, "Solution of Regular Sparse Triangular Linear Systems on Vector and Distributed Memory Multiprocessors", Tech Report RNR-93-07, NASA Ames Research Center, Moffett Field, CA 94035, April 1993.

[4] G. Bhanot, K. Jordan, J. Kennedy, J. Richardson, D. Sandee and M. Zagha, "Implementing the NAS Parallel Benchmarks on the CM-2 and CM200 Supercomputers", Thinking Machines Corp, Cambridge, MA 02142.

[5] S. Breit, W. Celmaster, W. Coney, R. Foster, B. Gaiman, G. Montry and C. Selvidge, "The Role of Computational Balance in the Implementation of the NAS parallel Benchmarks on the BBN TC2000 Computer", submitted to *Concurrency*, April 1991.

[6] M. Fillo, *Architectural Support for Scientific Applications on Multicomputers*, Hartung Gorre Verlag, Series in Microelectronics, Volume 27, Konstanz, Germany, 1993.

[7] B. Hendrickson, R. Leland, and S. Plimpton, "An Efficient Parallel Algorithm for Matrix-Vector Multiplication", Sandia Report SAND92-2765, Sandia National Lab, Albuquerque, NM 87185, March 1993.

[8] J. Lewis, and R. van de Geijn, "Distributed Memory Matrix-vector Multiplication and Conjugate Gradient Algorithms", in Proceedings Supercomputing '93, Portland, OR, Nov. 15-19, 1993.

[9] V. K. Naik, "Performance Issues in Implementing NAS Parallel Benchmark Applications on IBM SP-1", Research Report, T.J. Watson Research Center, IBM, (in preparation) 1993.

[10] T. H. Pulliam. "Efficient Solution Methods for the Navier-Stokes Equations", Lecture Notes for The Von Karman Institute for Fluid Dynamics Lecture Series, Jan. 20 - 24, 1986.

[11] S. White, "NAS Benchmark on Virtual Parallel Machines", Master's thesis, Emory University, 1993.

[12] S. Yoon, D. Kwak, and L. Chang, "LU-SGS Implicit Algorithm for Implicit Three Dimensional Navier-Stokes Equations with Source Term", AIAA Paper 89-1964-CP, American Institute of Aeronautics and Astronautics, Washington, D.C., 1989.

[13] M. Zagha and G.E. Blelloch, "Radix Sort for Vector Multiprocessors", Proceedings of Supercomputing '91, pp. 712-721, New York, NY, Nov. 1991.