ORNL/TM–13006

ornl

OAK RIDGE
NATIONAL
LABORATORY

LOCKHEED MARTIN

Message-Passing Performance of Various Computers

Jack J. Dongarra
Thomas H. Dunigan

MANAGED BY
LOCKHEED MARTIN ENERGY SYSTEMS, INC.
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

UCN-13673 (36 6-95)

ORNL/TM-13006

Computer Science and Mathematics Division

Mathematical Sciences Section

MESSAGE-PASSING PERFORMANCE OF VARIOUS COMPUTERS

Jack J. Dongarra and Thomas H. Dunigan

Mathematical Sciences Section
Oak Ridge National Laboratory
P.O. Box 2008, Bldg. 6012
Oak Ridge, TN 37831-6367
dongarra@cs.utk.edu thd@ornl.gov

Date Published: February 1996

Research was supported by the Applied Mathematical Sciences Research Program of the Office of Energy Research, U.S. Department of Energy.

Prepared by the
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831
managed by
Lockheed Martin Energy Research Corp.
for the
U.S. DEPARTMENT OF ENERGY
under Contract No. DE-AC05-96OR22464

Contents

MESSAGE-PASSING PERFORMANCE OF VARIOUS COMPUTERS

Jack J. Dongarra and Thomas H. Dunigan

Abstract

This report compares the performance of different computer systems for basic message passing. Latency and bandwidth are measured on Convex, Cray, IBM, Intel, KSR, Meiko, nCUBE, NEC, SGI, and TMC multiprocessors. Communication performance is contrasted with the computational power of each system. The comparison includes both shared and distributed memory computers as well as networked workstation clusters.

1. Introduction and Motivation

1.1. The Rise of the Microprocessor

The past decade has been one of the most exciting periods in computer development that the world has ever experienced. Performance improvements have been dramatic, and that trend promises to continue for the next several years.

In particular, microprocessor technology has changed rapidly. Microprocessors have become smaller, denser, and more powerful. Indeed, microprocessors have made such progress that, if cars had made equal progress since the day they were invented, we would now be able to buy a car for a few dollars, drive it across the country in a few minutes, and not worry about parking because the car would fit into one's pocket. The result is that the vendors of high-performance computing have turned to RISC microprocessors for performance.

Collections of these processors are interconnected by hardware and software to attack various applications. The physical interconnection of these processors may be contained in one or more cabinets as part of a multiprocessor, or the processors may be standalone workstations dispersed across a building or campus interconnected by a local area network. The effectiveness of using a collection of processors to solve a particular application is constrained by the amount of parallelism in the application, compiler technology, message passing software, amount of memory, and by the speed of the processors and of the interconnecting network.

1.2. Communications and Parallel Processing Systems

This report compares the results of a set of benchmarks for measuring communication time on a number of NUMA computers ranging from a collection of workstations using PVM [5] to machines like the IBM SP-2 and the Cray T3D using their native communication library, MPI [4], or PVM. We are interested in the communication performance for a number of reasons. First, our main interest is to obtain fundamental parameters on a given hardware platform to help in building models of execution. Second, performance data can be used to compare machines and help to evaluate new machines and architectures as they become available.

The following section describes the critical parameters in evaluating message passing systems. The techniques to measure these parameters are described. In section 3, the message passing performance of several multiprocessors and networks are presented. Communication and computational performance are contrasted. Section 4 provides details for obtaining the test software.

2. Message Passing

2.1. Programming Model

Processes of a parallel application distributed over a collection of processors must communicate problem parameters and results. In distributed memory multiprocessors or workstations on a network, the information is typically communicated with explicit message-passing subroutine calls. To send data to another process, a subroutine is usually provided that requires a destination address, message, and message length. The receiving process usually provides a buffer, a maximum length, and the senders address. The programming model is often extended to include both synchronous and asynchronous communication, group communication (broadcast and multicast), and aggregate operations (e.g., global sum).

Message passing performance is usually measured in units of time or bandwidth (bytes per second). In this report, we choose time as the measure of performance for sending a small message. The time for a small, or zero length, message is usually bounded by the speed of the signal through the media (latency) and any software overhead in sending/receiving the message. Small message times are important in synchronization and determining optimal granularity of parallelism. For large messages, bandwidth is the bounding metric, usually approaching the maximum bandwidth of the media. Choosing two numbers to represent the performance of a network can be misleading, so the reader is encouraged to plot communication time as function of message length to compare and understand the behavior of message passing systems.

Message passing time is usually a linear function of message size for two processors that are directly connected. For more complicated networks, a per-hop delay may increase the message passing time. Message-passing time, t_n, can be modeled as

$$t_n = \alpha + \beta n + (h-1)\gamma$$

with a start-up time, α, a per-byte cost, β, and a per-hop delay, γ, where n is the number of bytes per message and h the number of hops a message must travel. On most current message-passing multiprocessors the per-hop delay is negligible due to "worm-hole" routing techniques and the small diameter of the communication network [3]. The results reported in this report reflect nearest-neighbor communication. A linear least-squares fit can be used to calculate α and β from experimental data of message-passing times versus message length. The start-up time, α, may be slightly different than the zero-length time, and $1/\beta$ should be asymptotic bandwidth. The message length at which half the maximum bandwidth is achieved, $n_{1/2}$, is another metric of interest and is equal

to α/β [6]. As with any metric that is a ratio, any notion of "goodness" or "optimality" of $n_{1/2}$ should only be considered in the context of the underlying metrics α and β.

There are a number of factors that can affect the message passing performance. The number of times the message has to be copied or touched (e.g., checksums) is probably most influential and obviously a function of message size. The vendor may provide hints as to how to reduce message copies, for example, by posting the receive before the send. Second order effects of message size may also affect performance. Message lengths that are powers of two or cache-line size may provide better performance than shorter lengths. Buffer alignment on word, cache-line, or page may also affect performance. For small messages, context-switch times may contribute to delays. Touching all the pages of the buffers can reduce virtual memory effects. For shared media, contention may also affect performance. There also may be some first-time effects that can be identified or eliminated by performing some "warm up" tests before collecting performance data.

There are of course other parameters of a message-passing system that may affect performance for given applications. The aggregate bandwidth of the network, the amount of concurrency, reliability, scalability, and congestion management may be issues.

2.2. Measurement Methodology

To measure latency and bandwidth, we use a simple echo test between two adjacent nodes. A receiving node simply echos back whatever it is sent, and the sending node measures round-trip time. Times are collected for some number of repetitions (100 to 1000) over various messages sizes (0 to 1,000,000 bytes). Times can be collected outside the repetition loop as illustrated in Figure 2.1. If the system has high resolution timers then a more detailed analyses can be made by timing each send-receive pair. The time for each send-receive is saved in a vector and printed at the end of the test. For small message sizes, clock resolution may not be adequate, and clock jitter from time-sharing interrupts in the underlying OS may be observed. Unidirectional transfer time, or latency, is calculated as the minimum send-receive time (divided by two) for zero-length messages. Data rate, or bandwidth, is calculated from the number of bytes sent divided by half the round-trip time.

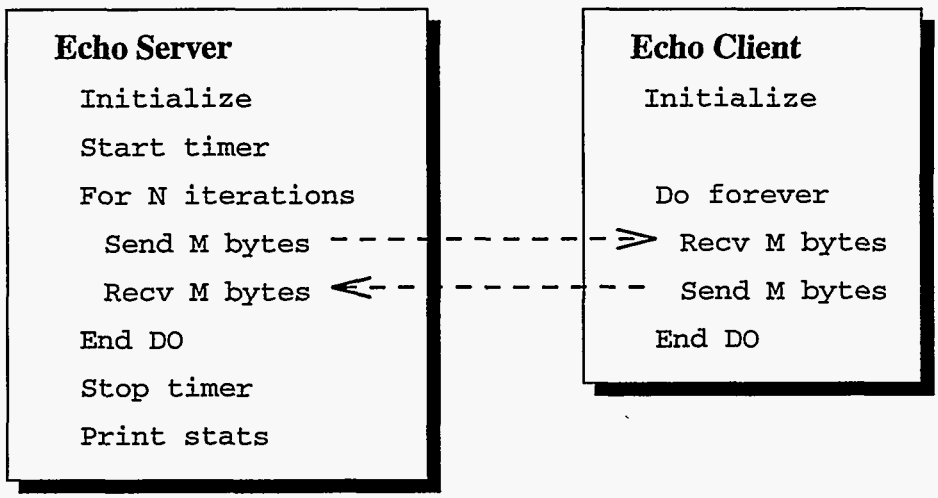

Figure 2.1: *Echo test pseudo-code.*

2.3. Latency and Bandwidth

Latency and bandwidth were measured on a number of different multiprocessors. Each architecture is briefly summarized in Appendix A. Table 2.3 shows the measured latency, bandwidth, and $n_{1/2}$ for nearest neighbor communication. The table also includes the peak bandwidth as stated by the vendor. For comparison, typical data rates and latencies are reported for several local area network technologies.

Table 2.1: *Multiprocessor Latency and Bandwidth.*

Machine	OS	Latency $n = 0$ (μs)	Bandwidth $n = 10^6$ (MB/s)	$n_{1/2}$ bytes	Theoretical Bandwidth (MB/s)
Convex SPP1000 (PVM)	SPP-UX 3.0.4.1	76	11	1000	250
Convex SPP1000 (sm 1-n)	SPP-UX 3.0.4.1	2.5	82	1000	250
Convex SPP1000 (sm m-n)	SPP-UX 3.0.4.1	12	59	1000	250
Convex SPP1200 (PVM)	SPP-UX 3.0.4.1	63	15	1000	250
Convex SPP1200 (sm 1-n)	SPP-UX 3.0.4.1	2.2	92	1000	250
Convex SPP1200 (sm m-n)	SPP-UX 3.0.4.1	11	71	1000	250
Cray T3D (sm)	MAX 1.2.0.2	3	128	363	300
Cray T3D (PVM)	MAX 1.2.0.2	21	27	1502	300
Intel Paragon	OSF 1.0.4	29	154	7236	175
Intel Paragon	SUNMOS 1.6.2	25	171	5856	175
Intel Delta	NX 3.3.10	77	8	900	22
Intel iPSC/860	NX 3.3.2	65	3	340	3
Intel iPSC/2	NX 3.3.2	370	2.8	1742	3
IBM SP-1	MPL	270	7	1904	40
IBM SP-2	MPI	35	35	3263	40
KSR-1	OSF R1.2.2	73	8	635	32
Meiko CS2 (sm)	Solaris 2.3	11	40	285	50
Meiko CS2	Solaris 2.3	83	43	3559	50
nCUBE 2	Vertex 2.0	154	1.7	333	2.5
nCUBE 1	Vertex 2.3	384	0.4	148	1
NEC Cenju-3	Env. Rel 1.5d	40	13	900	40
NEC Cenju-3 (sm)	Env. Rel 1.5d	34	25	400	40
SGI	IRIX 6.1	10	64	799	1200
TMC CM-5	CMMD 2.0	95	9	962	10
Ethernet	TCP/IP	500	0.9		1.2
FDDI	TCP/IP	900	9.7		12
ATM-100	TCP/IP	900	3.5		12

Figure 2.2 details the message-passing times of various multiprocessors over a range of message sizes. For small messages, the fixed overhead and latency dominate transfer time. For large message, the transfer time rises linearly with message size. Figure 2.3 illustrates the asymptotic behavior of bandwidth for large message sizes. It is possible to reduce latency on the shared-memory architectures by using shared-memory copy operations. These operations usually involve only one-processor and assume that the message is ready to be retrieved on the other processor. Figure 2.4 compares the message transfer times for the shared-memory *get* and explicit message passing for the Cray T3D, Meiko, and NEC. Current research in "active messages" is seeking ways to reduce message-passing overhead by eliminating context switches and message copying. Finally, Figure 2.5 graphically summarizes the communication performance of the various multiprocessors in a two-dimensional message-passing metric space. The upper-left region is the high performance area, lower performance and LAN networks occupy the lower performance region in the lower right.

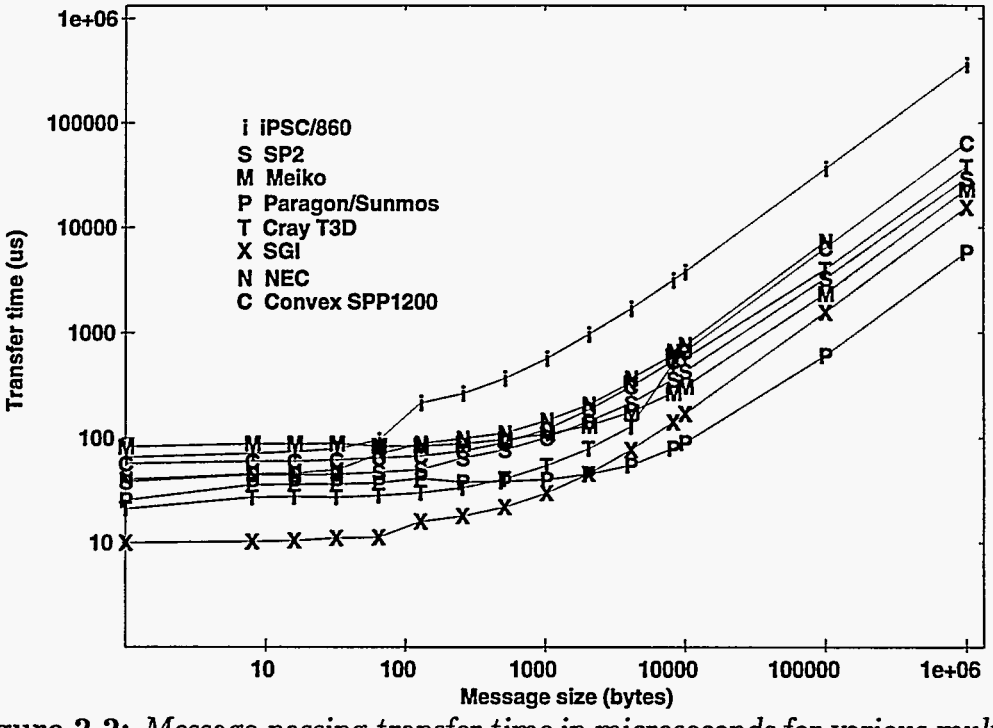

Figure 2.2: *Message-passing transfer time in microseconds for various multiprocessors and messages sizes.*

Since clusters of workstations on a network are often used as a virtual parallel machine, it is interesting to compare latency and bandwidths for various local area networks. Most communications over local area networks is done with the

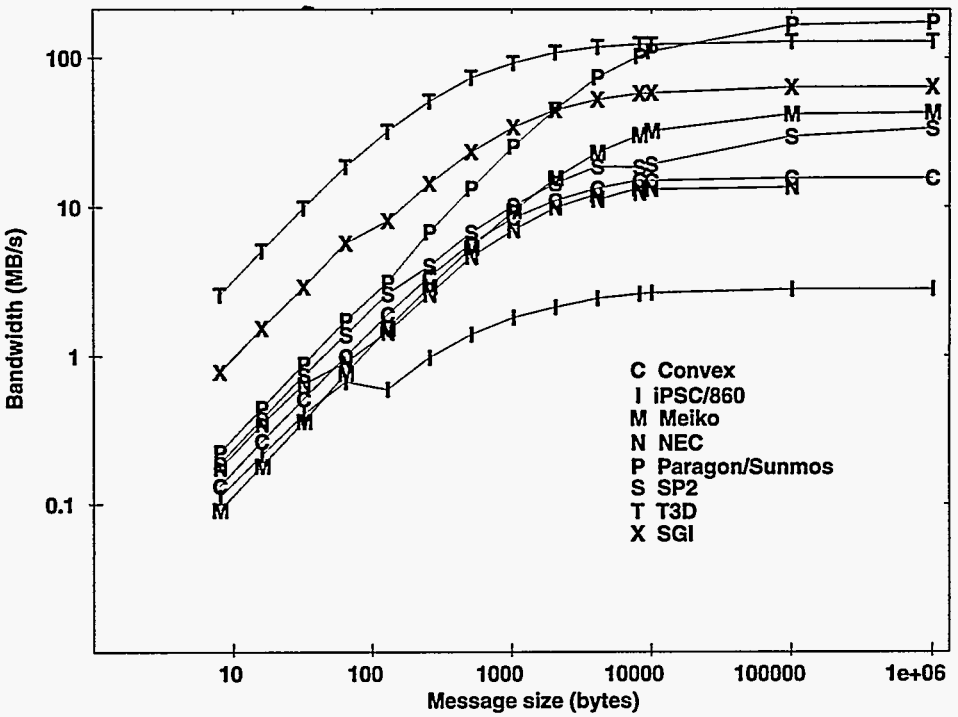

Figure 2.3: *Message-passing bandwidth in megabytes/second for various multiprocessors and messages sizes.*

TCP/IP protocols, though proprietary API's may exist. We measured latency for small messages using a UDP echo test. TCP bandwidth was measured at the receiver with the *ttcp* program using 50,000 byte messages and 50,000 byte window sizes. Some newer operating systems support even larger window sizes, which could provide higher bandwidths. Most high-end workstations can transmit network data at or near media data rates (e.g., 12 MB/second for FDDI). Data rates of 73 MB/second for UDP have been reported between Crays on HiPPI (and even over a wide-area using seven OC3's) [1]. Latency and bandwidth will depend as much on the efficiency of the TCP/IP implementation as on the network interface hardware and media. As with multiprocessors, the number of times the message is touched is a critical parameter as is context-switch time. Latencies for local area networks (Ethernet, FDDI, ATM, HiPPI) are typically on the order of 500 μs. For wide-area networks, latency is usually dominated by distance (speed of light) and is on the order of tens of milliseconds.

Message passing is often the limiting factor in performance of a parallel computer, so it is insightful to compare communication performance with computational performance for various architectures. The relative speed of computation and communication can be used in choosing the granularity of parallelism in im-

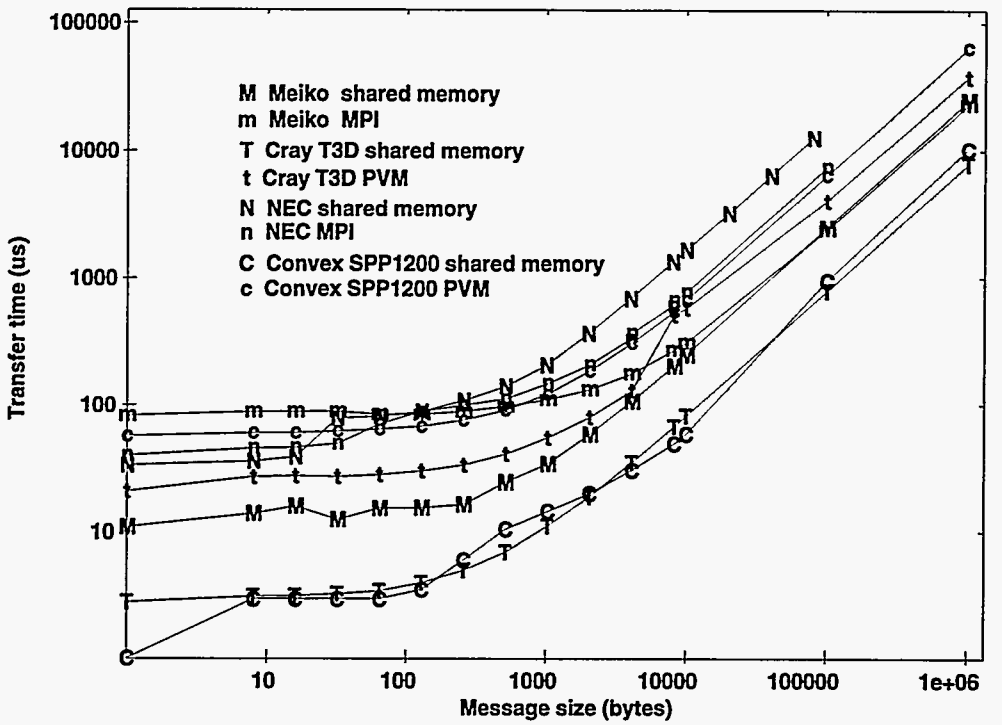

Figure 2.4: *Transfer time in microseconds for both shared-memory operations and explicit message passing.*

plementing a given application, or in optimizing the movement of data between processors or between levels of the memory hierarchy.

3. Computation and Communication

3.1. Performance

The performance of a computer is a complicated issue, a function of many interrelated quantities. These quantities include the application, the algorithm, the size of the problem, the high-level language, the implementation, the human level of effort used to optimize the program, the compiler's ability to optimize, the age of the compiler, the operating system, the architecture of the computer, and the hardware characteristics. The results presented for benchmark suites should not be extolled as measures of total system performance (unless enough analysis has been performed to indicate a reliable correlation of the benchmarks to the workload of interest) but, rather, as reference points for further evaluations.

Computational performance is often measured in terms of Megaflops, millions of floating point operations per second (Mflop/s). We usually include both additions and multiplications in the count of Mflop/s, and the reference to an

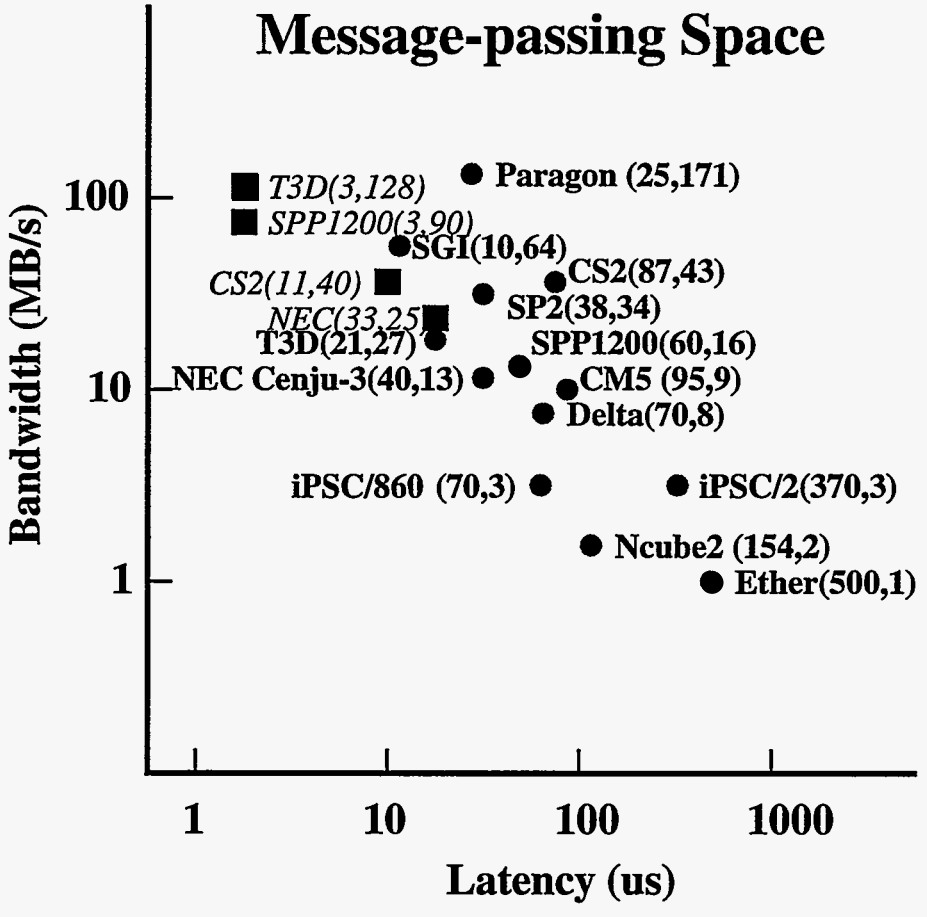

Figure 2.5: *Latency/bandwidth space for 0-byte message (latency) and 1 MB message (bandwidth). Block points represent shared-memory copy performance.*

operation is assumed to be on 64-bit operands.

The manufacturer usually refers to peak performance when describing a system. This peak performance is arrived at by counting the number of floating-point additions and multiplications that can be performed in a period of time, usually the cycle time of the machine. As an example, the IBM SP-1 processor, has a cycle time of 62.5 MHz. During a cycle the results of the multiply/add instruction can be completed giving:

$$2\ operations/1\ cycle\ *\ 1\ cycle/16nsec\ =\ 125\ Mflop/s.$$

Table 3.1 displays the performance for a single processor of various parallel computers using the LINPACK Benchmark [2]. The floating point execution rates have been converted to operations per cycle and also calculated the number of cycles consumed, as overhead (latency), during communication. At one time, a programmer had to go out of his way to code a matrix routine that would not run at nearly top efficiency on any system with an optimizing compiler. Owing to the proliferation of exotic computer architectures, this situation is no longer true.

Table 3.1: *Computation Performance.*

		Clock cycle		Linpack 100		Linpack 1000		Latency	
Machine	OS	MHz	(nsec)	Mfls	(ops/cl)	Mfls	(ops/cl)	us	(cl)
Convex SPP1000 (PVM)	SPP-UX 3.0.4.1	100	(10)	48	(.48)	123	(1.23)	76	(7600)
Convex SPP1000 (sm 1-n)								2.6	(260)
Convex SPP1000 (sm m-n)								11	(1080)
Convex SPP1200 (PVM)	SPP-UX 3.0.4.1	100	(8.33)	65	(.54)	123	(1.02)	63	(7560)
Convex SPP1200 (sm 1-n)								2.2	(264)
Convex SPP1200 (sm m-n)								11	(1260)
Cray T3D (sm)	MAX 1.2.0.2	150	(6.67)	38	(.25)	94	(.62)	3	(450)
Cray T3D (PVM)								21	(3150)
Intel Paragon	OSF 1.0.4	50	(20)	10	(.20)	34	(.68)	29	(1450)
Intel Paragon	SUNMOS 1.6.2							25	(1250)
Intel Delta	NX 3.3.10	40	(25)	9.8	(.25)	34	(.85)	77	(3080)
Intel iPSC/860	NX 3.3.2	40	(25)	9.8	(.25)	34	(.85)	65	(2600)
Intel iPSC/2	NX 3.3.2	16	(63)	.37	(.01)	–	(–)	370	(5920)
IBM SP-1	MPL	62.5	(16)	38	(.61)	104	(1.66)	270	(16875)
IBM SP-2	MPI	66	(15.15)	130	(1.97)	236	(3.58)	35	(2310)
KSR-1	OSF R1.2.2	40	(25)	15	(.38)	31	(.78)	73	(2920)
Meiko CS2 (MPI)	Solaris 2.3	90	(11.11)	24	(.27)	97	(1.08)	83	(7470)
Meiko CS2 (sm)								11	(990)
nCUBE 2	Vertex 2.0	20	(50)	.78	(.04)	2	(.10)	154	(3080)
nCUBE 1	Vertex 2.3	8	(125)	.10	(.01)	–	(–)	384	(3072)
NEC Cenju-3	Env Rev 1.5d	75	(13.3)	23	(.31)	39	(.52)	40	(3000)
NEC Cenju-3(sm)	Env Rev 1.5d	75	(13.3)	23	(.31)	39	(.52)	34	(2550)
SGI Power Challenge	IRIX 6.1	90	(11.11)	126	(1.4)	308	(3.42)	10	(900)
TMC CM-5	CMMD 2.0	32	(31.25)	–	(–)	–	(–)	95	(3040)

3.2. The LINPACK Benchmark

The LINPACK benchmark features solving a system of linear equation, $Ax = b$. The benchmark results examined here are for two distinct benchmark problems.

The first problem uses Fortran software from the LINPACK software package to solve a matrix problem of order 100. That is, the matrix A has 100 rows and columns and is said to be of size 100×100. The software used in this experiment is based on two routines from the LINPACK collection: DGEFA and DGESL. DGEFA performs the decomposition with partial pivoting, and DGESL uses that decomposition to solve the given system of linear equations. Most of the time - $O(n^3)$ floating-point operations - is spent in DGEFA. Once the matrix has been decomposed, DGESL is used to find the solution; this requires $O(n^2)$ floating-point operations.

DGEFA and DGESL in turn call three BLAS routines: DAXPY, IDAMAX, and DSCAL. For the size 100 benchmark, the BLAS used are written in Fortran. By far the major portion of time - over 90% at order 100 - is spent in subroutine DAXPY. DAXPY is used to multiply a scalar, α, times a vector, x, and add the results to another vector, y. It is called approximately $n^2/2$ times by DGEFA and $2n$ times by DGESL with vectors of varying length. The statement $y_i \leftarrow y_i + \alpha x_i$, which forms an element of the DAXPY operation, is executed approximately $n^3/3 + n^2$ times, which gives rise to roughly $(2/3)n^3$ floating-point operations in the solution. Thus, the benchmark requires roughly 2/3 million floating-point operations.

For the LINPACK 100 test, many processors achieve one floating point operation every four cycles, even though the process has the ability to deliver much more than this. The primary reason for this lack of performance relates to the poor compiler generated code and the algorithm's ineffective use of the memory hierarchy. There are a few exceptions, most notably the IBM SP-2's processor. The RS/6000-590 processor is able to achieve two floating point operations per cycle for the LINPACK 100 test, because the compiler and the cache structure work together. There are also examples of poor performance on some of the first generation parallel machines, such as the nCUBE 1 and 2 and the Intel iPSC/2. These processors are able to achieve only .01 to .04 floating point operations per cycle.

In the second benchmark, the problem size is larger (matrix of order 1000), and modifying or replacing the algorithm and software is permitted to achieve as high an execution rate as possible. (The optimized programs must still maintain the same relative accuracy as standard techniques.) The algorithm used for the $n = 1000$ problem makes better use of the memory hierarchy by utilizing the data in cache. Thus, the hardware had more opportunity for reaching near-asymptotic rates. Most of the processors achieve 70 to 80 % of their peak.

If we examine the algorithm used in LINPACK and look at how the data

are referenced, we see that at each step of the factorization process there are operations that modify a full submatrix of data. This update causes a block of data to be read, updated, and written back to central memory. The number of floating-point operations is $(2/3)n^3$, and the number of data references, both loads and stores, is $(2/3)n^3$. Thus, for every *add/multiply* pair we must perform a load and store of the elements, unfortunately obtaining no reuse of data. Even though the operations are fully vectorized, there is a significant bottleneck in data movement, resulting in poor performance. To achieve high-performance rates, this *operation-to-memory-reference rate* must be higher.

Just as the operation-to-memory-reference rate affects performance at the instruction level, the computation-to-communication ratio for message passing will affect application performance at a coarser level. The compiler and/or application programmer can improve application performance by algorithm restructing to increase data re-use and by reducing or pipelining message passing.

4. Future comparisons

In order that the reader might perform these same tests, both LINPACK and the message-passing tests are available over the Internet.

4.1. Rules for Running the Tests

The message-passing test software intentionally has been kept simple so that it will be easy for an experienced programmer to adapt the program, or parts of it, to a specific architecture with only a modest effort. In running the tests, the user is allowed to change the message passing calls to the appropriate call on the specific system the program is to be run on. We have provided both PVM and MPI [4] implementations in netlib.

4.2. Obtaining the Software

The software used to generate the data for this report can be obtained by sending electronic mail to *netlib@www.netlib.org* . To receive the single-precision software for this benchmark, in the mail message to *netlib@www.netlib.org* type *send comm.shar from benchmark* . To receive the double-precision software for this benchmark, type *send comm.shar from benchmark* .

A web browser can be used as well. With the url

http://www.netlib.org/benchmark/index.html

click on "benchmark/comm.shar".

5. References

[1] HPCwire No. 4912 12/2/94, 1994. Email exchange.

[2] J. Dongarra. Performance of various computers using standard linear equations software in a Fortran environment. Technical Report CS-89-85, University of Tennessee, 1995.

[3] T. H. Dunigan. Early experiences and performance of the intel paragon. Technical report, Oak Ridge National Laboratory, 1993. ORNL/TM-12194.

[4] Message Passing Interface Forum. MPI: A Message-Passing Interface Standard . *International Journal of Supercomputer Applications and High Perf ormance Computing*, 8(3/4), 1994. Special issue on MPI. Also available electronically, the url is `ftp://www.netlib.org/mpi/mpi-report.ps`.

[5] A. Geist, A. Beguelin, J. Dongarra, W. Jiang, R. Manchek, and V. Sunderam. *PVM: A Users' Guide and Tutorial for Networked Parallel Computing.* MIT Press, 1994.

[6] Roger Hockney. The communication challenge for mpp. *Parallel Computing*, 20:389–398, 1994.

Appendix

A. Appendix: Machine Configurations for Echo Tests

A summary of the various architectures and configurations used when these performance figures were measured follows. Unless otherwise noted, the test programs were compiled with *cc -O*.

The **Convex SPP1000** and SPP1200 consist of SCI-ring connected nodes (160 MB/second). Each SPP1000 node consists of eight 100 MHz HP PA RISC 7100 processors (120 MHz for the SPP1200) with a cross-bar memory interconnect (250 MB/second). The tests were run under SPP-UX 3.0.4.1 and ConvexPVM 3.3.7.1.

The **Cray T3D** is 3-D-torus multiprocessor using the 150 MHz DEC Alpha processor. Communication channels have a peak rate of 300 MB/second. Tests were performed using MAX 1.2.0.2. A special thanks to Majed Sidani of Cray for running our communication tests on the T3D using PVM. The PVM communication was with pvm_psend and pvm_precv.

The **Intel iPSC/860** is Intel's third generation hypercube. Each node has a 40 MHz i860 with 8 KB cache and at least 8 MB of memory. Communication channels have a peak rate of 2.8 MB/second. Tests were performed using NX 3.3.2. The **Intel iPSC/2** uses the same communication hardware as the iPSC/860 but uses 16 MHz 80386/7 for computation.

The **Intel Delta** is a 512-node mesh designed as a prototype for the Intel Paragon family. Each node has a 40 MHz i860 with 8 KB cache and 16 MB of memory. Communication channels have a peak rate of 22 MB/second. Tests were performed using NX 3.3.10.

The **Intel Paragon** is a mesh-based multiprocessor. Each node has at least two 50 MHz i860XP processors with 16 KB cache and at least 16 MB of memory. One processor is usually dedicated to communications. Communication channels have a peak rate of 175 MB/second. Test were run under OSF 1.0.4 Server 1.3/WW48-02 and SUNMOS 1.6.2 (using NX message passing).

The **IBM SP1** is an omega-switch-based multiprocessor using 62.5 MHz RS6000 processors. Communication channels have a peak rate of 40 MB/second. Tests were run using MPL.

The **IBM SP2** is an omega-switch-based multiprocessor using 66 MHz RS6000 processors with L2 cache. Communication channels have a peak rate of 40 MB/second. Tests were run using MPI. The MPI communication was with mpi_send and mpi_recv.

The **Kendall Square** architecture is a shared-memory system based on a hierarchy of rings using a custom 20 MHz processor. Shared-memory latency is about 7 μs, and bandwidth is about 32 MB/second. The message-passing performance was measured using Pacific Northwest Laboratory's *tcgmsg* library on one ring of a KSR1 running OSF R1.2.2.

The **Meiko CS2** uses SPARC processors with 200 Mflop/s vector co-processors. The communication topology is a fat tree with peak bandwidth of 50 MB/second. The MPSC message-passing library was used for the echo tests. Meiko notes that using point-to-point bidirectional channels in the echo test reduces latency from 82 microseconds to 14 microseconds. A special thanks to Jim Cownie of Meiko for running our communication tests.

The **Ncube** hypercube processors are custom processors with hypercube communication integrated into the chip. The first generation chip ran at 8 MHz, the second generation chip ran at 20 MHz.

The **NEC Cenju-3** results are from a 75 MHz VR4400SC MIPS processor with 32 KBytes of primary cache and 1 MByte of secondary cache using MPI under the Cenju Environment Release 1.5d. Communication channels have a peak rate of 40 MB/second through a multistage interconnection network.

The **SGI** results are from a 90 MHz PowerChallenge using MPI under IRIX 6.1. The SGI is a shared-memory multiprocessor using a 1.2 GB/s bus.

The **TMC CM5** is hypertree multiprocessor using 32 MHz SPARC processors with four vector units and 16 MB of memory per node. Communication channels have a peak rate of 20 MB/second. Tests were run using the message passing library CMMD 2.0.

INTERNAL DISTRIBUTION

1. T. S. Darland
2. J. J. Dongarra
3–7. T. H. Dunigan
8. G. A. Geist
9. K. L. Kliewer
10. C. E. Oliver
11. R. T. Primm
12–16. S. A. Raby
17–21. M. R. Leuze
22–26. R. F. Sincovec
27. P. H. Worley
28. Central Research Library
29. ORNL Patent Office
30. K-25 Appl Tech Library
31. Y-12 Technical Library
32. Laboratory Records - RC
33–34. Laboratory Records Department

EXTERNAL DISTRIBUTION

35. Cleve Ashcraft, Boeing Computer Services, P.O. Box 24346, M/S 7L-21, Seattle, WA 98124-0346
36. Lawrence J. Baker, Exxon Production Research Company, P.O. Box 2189, Houston, TX 77252-2189
37. Clive Baillie, Physics Department, Campus Box 390, University of Colorado, Boulder, CO 80309
38. Jesse L. Barlow, Department of Computer Science, 220 Pond Laboratory, Pennsylvania State University, University Park, PA 16802-6106
39. Edward H. Barsis, Computer Science and Mathematics, P. O. Box 5800, Sandia National Laboratories, Albuquerque, NM 87185
40. Chris Bischof, Mathematics and Computer Science Division, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439
41. Ake Bjorck, Department of Mathematics, Linkoping University, S-581 83 Linkoping, Sweden
42. Roger W. Brockett, Wang Professor EE and CS, Div. of Applied Sciences, 29 Oxford St., Harvard University, Cambridge, MA 02138
43. James C. Browne, Department of Computer Science, University of Texas, Austin, TX 78712
44. Bill L. Buzbee, Scientific Computing Division, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307
45. Donald A. Calahan, Department of Electrical and Computer Engineering, University of Michigan, Ann Arbor, MI 48109
46. Thomas A. Callcot, Director Science Alliance, 53 Turner House, University of Tennessee, Knoxville, TN 37996
47. Ian Cavers, Department of Computer Science, University of British Columbia, Vancouver, British Columbia V6T 1W5, Canada

48. Tony Chan, Department of Mathematics, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, CA 90024

49. Jagdish Chandra, Army Research Office, P.O. Box 12211, Research Triangle Park, NC 27709

50. Siddhartha Chatterjee, RIACS, MAIL STOP T045-1, NASA Ames Research Center, Moffett Field, CA 94035-1000

51. Eleanor Chu, Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, Canada N1G 2W1

52. Melvyn Ciment, National Science Foundation, 1800 G Street N.W., Washington, DC 20550

53. Tom Coleman, Department of Computer Science, Cornell University, Ithaca, NY 14853

54. Paul Concus, Mathematics and Computing, Lawrence Berkeley Laboratory, Berkeley, CA 94720

55. Andy Conn, IBM T. J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598

56. John M. Conroy, Supercomputer Research Center, 17100 Science Drive, Bowie, MD 20715-4300

57. Jane K. Cullum, IBM T. J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598

58. George Cybenko, Center for Supercomputing Research and Development, University of Illinois, 104 S. Wright Street, Urbana, IL 61801-2932

59. George J. Davis, Department of Mathematics, Georgia State University, Atlanta, GA 30303

60. Tim A. Davis, Computer and Information Sciences Department, 301 CSE, University of Florida, Gainesville, FL 32611-2024

61. John J. Dorning, Department of Nuclear Engineering Physics, Thornton Hall, McCormick Road, University of Virginia, Charlottesville, VA 22901

62. Larry Dowdy, Computer Science Department, Vanderbilt University, Nashville, TN 37235

63. Iain Duff, Numerical Analysis Group, Central Computing Department, Atlas Centre, Rutherford Appleton Laboratory, Didcot, Oxon OX11 0QX, England

64. Patricia Eberlein, Department of Computer Science, SUNY at Buffalo, Buffalo, NY 14260

65. Albert M. Erisman, Boeing Computer Services, Engineering Technology Applications, P.O. Box 24346, M/S 7L-20, Seattle, WA 98124-0346

66. Geoffrey C. Fox, Northeast Parallel Architectures Center, 111 College Place, Syracuse University, Syracuse, NY 13244-4100

67. Robert E. Funderlic, Department of Computer Science, North Carolina State University, Raleigh, NC 27650

68. Professor Dennis B. Gannon, Computer Science Department, Indiana University, Bloomington, IN 47401

69. David M. Gay, Bell Laboratories, 600 Mountain Avenue, Murray Hill, NJ 07974

70. C. William Gear, NEC Research Institute, 4 Independence Way, Princeton, NJ 08540

71. W. Morven Gentleman, Division of Electrical Engineering, National Research Council, Building M-50, Room 344, Montreal Road, Ottawa, Ontario, Canada K1A 0R8

72. J. Alan George, Vice President, Academic and Provost, Needles Hall, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

73. John R. Gilbert, Xerox Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, CA 94304

74. Gene H. Golub, Department of Computer Science, Stanford University, Stanford, CA 94305

75. Joseph F. Grcar, Division 8245, Sandia National Laboratories, Livermore, CA 94551-0969

76. John Gustafson, Ames Laboratory, Iowa State University, Ames, IA 50011

77. Michael T. Heath, National Center for Supercomputing Applications, 4157 Beckman Institute, University of Illinois, 405 North Mathews Avenue, Urbana, IL 61801-2300

78. Don E. Heller, Center for Research on Parallel Computation, Rice University, P.O. Box 1892, Houston, TX 77251

79. Robert E. Huddleston, Computation Department, Lawrence Livermore National Laboratory, P.O. Box 808, Livermore, CA 94550

80. Lennart Johnsson, Thinking Machines Inc., 245 First Street, Cambridge, MA 02142-1214

81. Harry Jordan, Department of Electrical and Computer Engineering, University of Colorado, Boulder, CO 80309

82. Malvyn H. Kalos, Cornell Theory Center, Engineering and Theory Center Bldg., Cornell University, Ithaca, NY 14853-3901

83. Hans Kaper, Mathematics and Computer Science Division, Argonne National Laboratory, 9700 South Cass Avenue, Bldg. 221, Argonne, IL 60439

84. Kenneth Kennedy, Department of Computer Science, Rice University, P.O. Box 1892, Houston, TX 77001

85. Richard Lau, Office of Naval Research, Code 1111MA, 800 Quincy Street, Boston, Tower 1, Arlington, VA 22217-5000

86. Alan J. Laub, Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106

87. Robert L. Launer, Army Research Office, P.O. Box 12211, Research Triangle Park, NC 27709

88. Charles Lawson, MS 301-490, Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109

89. Professor Peter Lax, Courant Institute for Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012

90. John G. Lewis, Boeing Computer Services, P.O. Box 24346, M/S 7L-21, Seattle, WA 98124-0346

91. Robert F. Lucas, Supercomputer Research Center, 17100 Science Drive, Bowie, MD 20715-4300

92. Franklin Luk, Electrical Engineering Department, Cornell University, Ithaca, NY 14853

93. Paul C. Messina, Mail Code 158-79, California Institute of Technology, 1201 E. California Blvd., Pasadena, CA 91125

94. James McGraw, Lawrence Livermore National Laboratory, L-306, P.O. Box 808, Livermore, CA 94550

95. Cleve Moler, The Mathworks, 325 Linfield Place, Menlo Park, CA 94025

96. Dr. David Nelson, Director of Scientific Computing ER-7, Applied Mathematical Sciences, Office of Energy Research, U. S. Department of Energy, Washington DC 20585

97. Professor V. E. Oberacker, Department of Physics, Vanderbilt University, Box 1807 Station B, Nashville, TN 37235

98. Dianne P. O'Leary, Computer Science Department, University of Maryland, College Park, MD 20742

99. James M. Ortega, Department of Applied Mathematics, Thornton Hall, University of Virginia, Charlottesville, VA 22901

100. Charles F. Osgood, National Security Agency, Ft. George G. Meade, MD 20755

101. Roy P. Pargas, Department of Computer Science, Clemson University, Clemson, SC 29634-1906

102. Beresford N. Parlett, Department of Mathematics, University of California, Berkeley, CA 94720

103. Merrell Patrick, Department of Computer Science, Duke University, Durham, NC 27706

104. Robert J. Plemmons, Departments of Mathematics and Computer Science, Box 7311, Wake Forest University, Winston-Salem, NC 27109

105. James Pool, Caltech Concurrent Supercomputing Facility, California Institute of Technology, MS 158-79, Pasadena, CA 91125

106. Alex Pothen, Department of Computer Science, Pennsylvania State University, University Park, PA 16802

107. Yuanchang Qi, IBM European Petroleum Application Center, P.O. Box 585, N-4040 Hafrsfjord, Norway

108. Giuseppe Radicati, IBM European Center for Scientific and Engineering Computing, via del Giorgione 159, I-00147 Roma, Italy

109. Professor Daniel A. Reed, Computer Science Department, University of Illinois, Urbana, IL 61801

110. John K. Reid, Numerical Analysis Group, Central Computing Department, Atlas Centre, Rutherford Appleton Laboratory, Didcot, Oxon OX11 0QX, England

111. John R. Rice, Computer Science Department, Purdue University, West Lafayette, IN 47907

112. Donald J. Rose, Department of Computer Science, Duke University, Durham, NC 27706

113. Edward Rothberg, Department of Computer Science, Stanford University, Stanford, CA 94305

114. Joel Saltz, ICASE, MS 132C, NASA Langley Research Center, Hampton, VA 23665

115. Ahmed H. Sameh, Center for Supercomputer R&D, 469 CSRL 1308 West Main St., University of Illinois, Urbana, IL 61801

116. Robert Schreiber, RIACS, Mail Stop 230-5, NASA Ames Research Center, Moffett Field, CA 94035

117. Martin H. Schultz, Department of Computer Science, Yale University, P.O. Box 2158 Yale Station, New Haven, CT 06520

118. David S. Scott, Intel Scientific Computers, 15201 N.W. Greenbrier Parkway, Beaverton, OR 97006

119. Kermit Sigmon, Department of Mathematics, University of Florida, Gainesville, FL 32611

120. Horst Simon, Mail Stop T045-1, NASA Ames Research Center, Moffett Field, CA 94035

121. Danny C. Sorensen, Department of Mathematical Sciences, Rice University, P. O. Box 1892, Houston, TX 77251

122. G. W. Stewart, Computer Science Department, University of Maryland, College Park, MD 20742

123. Paul N. Swartztrauber, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307

124. Robert G. Voigt, ICASE, MS 132-C, NASA Langley Research Center, Hampton, VA 23665

125. Phuong Vu, Cray Research, Inc., 19607 Franz Rd., Houston, TX 77084

126. Robert Ward, Department of Computer Science, 107 Ayres Hall, University of Tennessee, Knoxville, TN 37996-1301

127. Andrew B. White, Computing Division, Los Alamos National Laboratory, P.O. Box 1663 MS-265, Los Alamos, NM 87545

128. David Young, University of Texas, Center for Numerical Analysis, RLM 13.150, Austin, TX 78731

129. Office of Assistant Manager for Energy Research and Development, U.S. Department of Energy, Oak Ridge Operations Office, P.O. Box 2001 Oak Ridge, TN 37831-8600

130–131. Office of Scientific & Technical Information, P.O. Box 62, Oak Ridge, TN 37831