

**Commodity Workstations Clusters
and the
DAISy Cluster Project
(Distributed Array of Inexpensive Systems)**

by

John Laroco
Sandia National Laboratories
Livermore, CA

Thesis

Submitted in Partial Fulfillment of the Requirements for
the Degree of Master of Science
in Computer Engineering in the School of Engineering at
Santa Clara University, 1996

Santa Clara, CA

Computer Engineering
Master's Thesis

Acknowledgment

The following discussion of the DAISy project is provided by the Distributed Systems Research Department at Sandia National Laboratories, CA (SNL,CA).

Special thanks are extended to: Robert Armstrong (SNL, CA), Russell Carter (Geli Engineering), Alex Cheung (SNL, CA), and David Evensky (SNL, CA). The above, along with myself, were responsible for the original design and construction of the DAISy cluster.

The following are special mentions, for without their support, guidance, and patience this paper could not have been written: Larry Ranh (SNL, CA), Robert Armstrong, Ray Cline (), Jim Costa (SNL, CA), and Rich Palmer (SNL, CA).

Lastly, to Nicollette, Jonathan, and Gigi, for providing the kind of love and support a father and husbands needs.

Contents

1	Introduction	1
2	History	4
2.1	Performance Growth for Practical Parallel Computing	4
2.2	Distributed System	5
2.3	Distributed System Key Characteristics	6
2.4	Generalized Parallel Processor	8
2.5	Control vs. Data Parallelism	9
2.6	SPMD (Single Program, Multiple Data)	10
2.7	Distributed Computing Models/IPC Socket Method	11
2.7.1	Client/Server Model	11
2.7.2	Message-Passing Model	12
2.7.3	IPC Sockets Method	14
2.8	System Performance	15
3	System Hardware Configuration	16
3.1	DAISy	16
3.2	Memory	17
3.3	The P54C Pentium™ 90 MHz Processor	18
3.4	PCI (Peripheral Control Interface)	19
3.5	Frame Switched 100BASE-TX Fast Ethernet	19
4	System Software Configuration	21
4.1	FreeBSD	21
4.2	Message Passing Software	21
4.3	Directory Hierarchy	22
5	HEAT (Heterogeneous Environment And Testbed)	23
5.1	HEAT	23
6	Benchmark results	25
6.1	Theoretical Floating-point Performance	25
6.2	Matrix-Matrix Multiply (using assembly coded DGEMM)	26
6.3	LINPACK: Standard Linear Equations Software	26
6.4	STREAM: A Main Memory Benchmark	27
6.5	SPECx92: Standard Performance Evaluation Corporation	27
6.6	Imbench: Portable tools for performance analysis	28
6.6.1	Bandwidth Measurements	30
6.6.1.1	Memory Bandwidth	30
6.6.1.2	IPC Bandwidth	31
6.6.1.3	Cached I/O Bandwidth	32
6.6.2	Latency Measurements	33
6.6.2.1	Memory Latency	33
6.6.2.2	Basic Operating System Entry Cost	35
6.6.2.3	Process Creation Times	35
6.6.2.4	Context Switching	36
6.6.2.5	Inter Process Communication	37
6.6.2.6	File System Latency	39
6.6.2.7	Page Fault Latency	39
6.7	Bonnie: Disk Performance Benchmark	40
6.8	Netperf suite: A Network Performance Benchmark	40
6.8.1	Stream	41
6.8.1.1	TCP Range	41

6.8.1.2	TCP Stream	42
6.8.1.3	UDP Stream	42
6.8.2	Request/Response	42
6.8.2.1	TCP Request/Response	42
6.8.2.2	UDP Request/Response	43
6.9	The PVM Timing Example	43
6.10	NFS Performance.	43
6.11	NAS Parallel Benchmarks 1.0, PVM versions	44
6.11.1	Kernel EP	44
6.12	NAS Parallel Benchmarks 2.0, MPI versions.	45
6.12.1	Application Benchmark: LU	46
6.12.2	Application Benchmark: SP and BT.	47
6.13	Parallel Seismic Inverse Problem	48
7	Conclusion	51
7.1	Benchmark Results	51
7.2	Overall	52
8	References	54

Appendix

A	LINPACK: Standard Linear Equations Software	56
B	STREAM: A Main Memory Benchmark	57
C	SPECx92: Standard Performance Evaluation Corporation	58
D	lmbench 1.0: Portable Tools for Performance Analysis	61
E	Bonnie: Disk Performance Benchmark	80
F	Netperf: A Network Performance Benchmark	82
G	The PVM Timing Example	87
H	NAS Parallel Benchmarks 1.0, PVM	89

DAISy Cluster Project

(Distributed Array of Inexpensive Systems)

John Laroco
Sandia National Laboratories
Livermore, CA

Department of Computer Engineering
Santa Clara University, 1996

Abstract

DAISy is a 16 node PC cluster running a full UNIX compatible operating system. The network media used includes standard 10Mb/s (10BASE-2) Ethernet (used for client node NFS mounts and any client node interactive work users find necessary), and, switched 100Mb/s (100BASE-TX) Fast Ethernet. (used for user program message passing traffic). The DAISy cluster is used to investigate the viability of commodity PC technology to perform computation of scientific and engineering problems traditionally performed on "Supercomputers," and more recently high performance RISC workstations and clusters of RISC workstations. Performance analysis of the various single node subsystems were carried out, along with performance analysis of the cluster as a whole on a number of parallel applications. The results show that the current Pentium 90MHz CPU and motherboards used are well within that of many low-end workstations offered by traditional workstation vendors, yet lack the power to drive the networking hardware to the theoretical maximum. The viability of using DAISy as a parallel processing system was proven with the various parallel applications that were run. Computation intensive or "Embarrassingly Parallel" codes proved that (with disregard to the networking media used) the DAISy cluster showed exceptional performance. With applications that require small to medium memory requirements and relatively fast networking media DAISy proved itself to be a viable alternative. With little investment, upgrading to the now standard 200MHz Pentium Pros would increase single node performance by a factor of three and increase the available networking bandwidth to well over 70% with motherboards that support Piplined Burst SRAM. This type of performance increase at a cost effective upgrade proves that commodity based PC clusters ought to gain an increasingly large share of the distributed computing market.