UNIVERSITY OF CALIFORNIA, SAN DIEGO

Performance Prediction and Scheduling for Parallel Applications on Multi-User Clusters

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Computer Science

by

Jennifer Melinda Schopf

Committee in charge:

Professor Francine Berman, Chairperson
Professor Larry Carter
Professor Mark Ellisman
Professor William G. Griswold
Professor Peter Taylor

1998

The dissertation of Jennifer Melinda Schopf is approved, and it is acceptable in quality and form for publication on microfilm:

Chair

University of California, San Diego

1998

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my advisor, Fran Berman, for all her support, encouragement and guidance over the last 5 years. Fran's advice and insight have made an enormous contribution to this thesis, and life in general. Without her help, I'd be working for Rolling Stone right now.

Second only to Fran has been Rich Wolski's influence on this work. Rich has commented, complained, suggested, and clarified many many things over the last several years, and he always had a time to chat (or drink coffee) when it was most needed.

Additionally, I thank the members of my committee: Larry Carter, Bill Griswold, Mark Ellisman, Peter Taylor, Reagan Moore and Charlie Brooks for their interest in this work and much needed feedback. And additional thanks to Professors Allen Downey and Keith Maruzllo who gave invaluable assistance along the way.

Special thanks to the other members of the lab, whose friendship and support helped make my graduate career if not fun, then at least interesting. In particular, thanks for all the late nights with Steve Fink, Neil Spring, Silvia Figueira, Graziano Obertelli, and Dmitrii Zagorodnov.

A big thanks to everyone who helped me with access to the various platforms used in this thesis: Phil Andrews and Frank Dwyer, both at SDSC, for use of the Alpha Farm; Neil Spring, SDSC and UCSD, Jim Hayes, UCSD and Graziano Obertelli, UCSD, for the Linux Cluster (aka the Circus); and the entire PCL for giving me dedicated time on the machines there.

Also, thanks to the people who donated or developed codes that were part of the test suite I used: Silvia Figueria, UCSD and SCU, for the original implementation of the SOR code; Karan Bhatia, UCSD, for the original implementation of the GA code; Kelsey Anderson, Harvey Mudd College, for her analysis of the NAS Parallel Benchmark during her CRA Distributed Mentoring summer; Alex Cereghini and Wes Ingalls for the N-body implementation, part of their CS 160 class project; James Kohl, ORNL, for the HiTemp code.

Thanks for the use of the Network Weather Service, originally developed by Rich Wolski, implemented by Neil Spring, and maintained by Jim Hayes and Graziano Obertelli. And to Neil and Dmitrii for help with data analysis.

This research was supported in part by NASA GSRP Grant #NGT-1-52133, under the guidance of David Rudy, and DARPA Contract N66001-97-C-8531, many thanks to Fredricka Darema.

To my family, for all your support: Mom, Ken, Tracy, Carl and Erin. And thanks to my non-computer science friends who kept me sane and whole: Dan, Mike, Heather, Aaron, Kacy, Coryne, and everyone else who made sure I ate and slept at semi-regular intervals. And of course, to Tiger, for being Tiger.

VITA

August 3, 1970	Born, Woods Hole, MA
1988	B.A. Mathematics/Computer Science Vassar College, Poughkeepsie, NY
1994	M.S. Computer Science University of California, San Diego
1998	Doctor of Philosophy University of California, San Diego

PUBLICATIONS

Schopf, J.M., and Berman, F., Performance Prediction in Production Environments, *Proceedings of IPPS/SPDP '98*, April 1998. Also available as UCSD Technical Report #CS97-558, September 1997.

Schopf, J.M., Structural Prediction Models for High-Performance Distributed Applications, *Proceedings of the Cluster Computing Conference (CCC '97)*, March 1997. Also available as UCSD Technical Report #CS97-528.

Berman F., Wolski, R.M., Figueira, S., Schopf, J., and Shao, G., Application-Level Scheduling on Distributed Heterogeneous Networks, *Proceedings of Supercomputing '96*, November 1996.

Wolski, R.M., Anglano, C., Schopf, J., and Berman, F., Developing Heterogeneous Applications Using Zoom and HeNCE, *Proceedings of the Heterogeneous Processing Workshop, International Parallel Processing Symposium (IPPS)*, April 1995.

Schopf, J.M., and Berman, F., Performance Prediction Using Intervals, UCSD Technical Report #CS97-541, May 1997.

Anglano, C., Schopf, J., Wolski, R., and Berman, F., Zoom: A Hierarchical Representation for Heterogeneous Applications, UCSD Technical Report #CS95-451, October, 1995.

FIELDS OF STUDY

Major Field: Computer Science
    Studies in Distributed Parallel Computing.
    Professor Francine Berman

ABSTRACT OF THE DISSERTATION

Performance Prediction and Scheduling for
Parallel Applications on Multi-User Clusters

by

Jennifer Melinda Schopf
Doctor of Philosophy in Computer Science
University of California, San Diego, 1998
Professor Francine Berman, Chair

Current distributed parallel platforms can provide an efficient set of resources on which to execute many scientific applications. However, when these platforms are shared by multiple users, the performance of an application may be impacted in dynamic and unpredictable ways. In particular, it becomes challenging to achieve good performance for any single application in the system.

To efficiently use distributed parallel platforms, adaptive scheduling techniques are needed to leverage the dynamic performance characteristics of shared resources. A fundamental problem in developing adaptive schedulers is the inability to accurately predict application execution performance.

One promising approach to modeling distributed parallel applications is **structural modeling**. This approach decomposes application performance with respect to its functional structure into a top-level model and interacting component models, representing application tasks. Component models reflect the dynamic, time-dependent changes in effective capacities by allowing application or model developers to choose their parameterizations.

Many conventional models are parameterized by single (point) values. However, in shared environments, point values may provide an inaccurate representation of application behavior. An alternative to using point values is the use

of **stochastic values**, or distributions. Whereas a point value provides a single value representation of a quantity, a stochastic value provides a set of possible values weighted by probabilities to represent a range of likely behavior.

Parameters for structural models can be stochastic values in order to accurately reflect a range of observed behavior, and lend themselves to advanced scheduling techniques. Performance models that cannot adapt to changing resource conditions are often insufficient for the resource selection and scheduling needs of production distributed parallel applications. We take advantage of of stochastic information and develop a **stochastic scheduler** to maximize application performance on shared clusters of workstations.

This thesis presents three main contributions: an approach to define distributed parallel application performance models called **structural modeling**, the ability to parameterize these models with **stochastic values** in order to meet the prediction needs of multiple user clusters of workstations, and a **stochastic scheduling policy** that can make use of the resulting stochastic prediction to achieve more performance efficient application execution times and more predictable application behavior.