Operating System Resource Reservation for Real-Time and Multimedia Applications

Clifford W. Mercer

June 1997
CMU-CS-97-155

School of Computer Science
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213-3890

Submitted in partial fulfillment of the requirements for a degree of Doctor of Philosophy in Computer Science

Thesis Committee:

Elmootazbellah N. Elnozahy, Co-chair
Ragunathan Rajkumar, Co-chair
Mahadev Satyanarayanan
Hideyuki Tokuda
Kevin Jeffay, University of North Carolina-Chapel Hill

This research was supported in part by a National Science Foundation Graduate Fellowship, in part by Bell Communications Research, in part by U. S. Naval Research and Development, in part by the Office of Naval Research, in part by Northrop-Grumman, in part by Philips Research, and in part by Loral Federal Systems. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of NSF, Bell Communications Research, NRaD, ONR, Northrop-Grumman, Philips Research, or Loral Federal Systems.

Keywords: Resource reservation, Real-time systems, Operating systems, Multimedia applications, Scheduling, Resource management, Real-Time Mach

For Franke

Abstract

Increases in processor speeds and the availability of audio and video devices for personal computers have encouraged the development of interactive multimedia applications for teleconferencing and digital audio/video presentation among others. These applications have stringent timing constraints, and traditional operating systems are not well suited to satisfying such constraints. On the other hand, hard real-time systems that can meet these constraints are typically static and inflexible.

This dissertation presents an enforced operating system resource reservation model for the design and implementation of predictable real-time programs. Applications can reserve resources based on their timing constraints, and an enforcement mechanism ensures that they do not overrun their reservations. Thus, reserves isolate real-time applications from the temporal properties of other real-time (and non-real-time) applications just as virtual memory systems isolate applications from memory accesses by other applications. In addition, reserves are first class objects that are separated from control abstractions such as processes or threads. Therefore reserves can be passed between applications, and this model extends naturally to distributed systems.

Reserves support the development of hard real-time and soft real-time programs, and programming techniques based on reserves illustrate how to use them effectively. An implementation of processor reserves in Real-Time Mach shows that reserved multimedia applications can achieve predictable real-time performance.

Acknowledgments

I would like to express my appreciation to my advisors, Mootaz Elnozahy and Raj Rajkumar, for their guidance, technical expertise, and time which they so generously shared with me during my doctoral studies. Thanks also to M. Satyanarayanan, Hide Tokuda, and Kevin Jeffay for their insights, critical comments, and encouragement. I am especially indebted to Hide Tokuda who was my mentor for several years before I entered the graduate program at CMU. Hide started the Real-Time Mach Project, which provided the research environment for this work, and Raj took over the project in the final years of my thesis work. I am indebted to both of them for their support.

I would like to thank Jim Morris for his guidance and assistance. He helped me better understand how to do systems research and served as a role model for me. Thanks go to Roger Dannenberg for the opportunity to work with him and learn from him on several interesting computer music projects (and for the occasional trumpet duet and brass ensemble session).

A special thanks to Joan Maddamma who looked out for me in many ways during my years at CMU. From the time I was an undergraduate through my years as a graduate student, Joan has been a very special friend.

Many other people who I worked with over the years had a positive influence on me and on the work contained in this dissertation. I would like to thank Stefan Savage and Jim Zelenka for their help with pieces of the implementation of reserves in Real-Time Mach and for being good friends and colleagues. And I would like to thank Chen Lee, Prithvi Rao, Andrei Ghetie, Yutaka Ishikawa, Tatsuo Nakajima, Takuro Kitayama, Dan Katcher, Kevin Kettler, John Sasinowski, and Saurav Chatterjee for being part of the invigorating academic environment that makes CMU so rewarding.

I would like to thank my father, mother, and sister for their moral support during my years away at school.

Finally, I want to thank my wife, Franke, who supported me wholeheartedly in the pursuit of my graduate studies. Without her, this work would not have been possible. Thanks also to my daughters, Caroline and Leigh, for their love and patience with me during the final years of the thesis work.