FOUNDATIONS OF DEPENDABLE COMPUTING
Paradigms for Dependable Applications

Edited by

Gary M. Koob
Clifford G. Lau

KLUWER ACADEMIC PUBLISHERS

FOUNDATIONS OF DEPENDABLE COMPUTING

Paradigms for Dependable Applications

edited by

Gary M. Koob
Clifford G. Lau
Office of Naval Research

KLUWER ACADEMIC PUBLISHERS
Boston / Dordrecht / London

Distributors for North America:
Kluwer Academic Publishers
101 Philip Drive
Assinippi Park
Norwell, Massachusetts 02061 USA

Distributors for all other countries:
Kluwer Academic Publishers Group
Distribution Centre
Post Office Box 322
3300 AH Dordrecht, THE NETHERLANDS

Library of Congress Cataloging-in-Publication Data

A C.I.P. Catalogue record for this book is available
from the Library of Congress.

Printed on acid-free paper.

Printed in the United States of America

FOUNDATIONS OF DEPENDABLE COMPUTING

Paradigms for Dependable Applications

THE KLUWER INTERNATIONAL SERIES IN ENGINEERING AND COMPUTER SCIENCE

OFFICE OF NAVAL RESEARCH
Advanced Book Series

Consulting Editor

André M. van Tilborg

Other titles in the series:

FOUNDATIONS OF DEPENDABLE COMPUTING: Models and Frameworks for Dependable Systems, edited by Gary M. Koob and Clifford G. Lau
ISBN: 0-7923-9484-4

FOUNDATIONS OF DEPENDABLE COMPUTING: System Implementation, edited by Gary M. Koob and Clifford G. Lau
ISBN: 0-7923-9486-0

PARALLEL ALGORITHM DERIVATION AND PROGRAM TRANSFORMATION, edited by Robert Paige, John Reif and Ralph Wachter
ISBN: 0-7923-9362-7

FOUNDATIONS OF KNOWLEDGE ACQUISITION: Cognitive Models of Complex Learning, edited by Susan Chipman and Alan L. Meyrowitz
ISBN: 0-7923-9277-9

FOUNDATIONS OF KNOWLEDGE ACQUISITION: Machine Learning, edited by Alan L. Meyrowitz and Susan Chipman
ISBN: 0-7923-9278-7

FOUNDATIONS OF REAL-TIME COMPUTING: Formal Specifications and Methods, edited by André M. van Tilborg and Gary M. Koob
ISBN: 0-7923-9167-5

FOUNDATIONS OF REAL-TIME COMPUTING: Scheduling and Resource Management, edited by André M. van Tilborg and Gary M. Koob
ISBN: 0-7923-9166-7

PREFACE

Dependability has long been a central concern in the design of space-based and military systems, where survivability for the prescribed mission duration is an essential requirement, and is becoming an increasingly important attribute of government and commercial systems where reduced availability may have severe financial consequences or even lead to loss of life. Historically, research in the field of dependable computing has focused on the theory and techniques for preventing hardware and environmentally induced faults through increasing the intrinsic reliability of components and systems (fault avoidance), or surviving such faults through massive redundancy at the hardware level (fault tolerance).

Recent advances in hardware, software, and measurement technology coupled with new insights into the nature, scope, and fundamental principles of dependable computing, however, contributed to the creation of a challenging new research agenda in the late eighties aimed at dramatically increasing the power, effectiveness, and efficiency of approaches to ensuring dependability in critical systems

At the core of this new agenda was a paradigm shift spurred by the recognition that dependability is fundamentally an attribute of applications and services—not platforms. Research should therefore focus on (1) developing a scientific understanding of the manifestations of faults at the application level in terms of their ultimate impact on the correctness and survivability of the application; (2) innovative, application-sensitive approaches to detecting and mitigating this impact; and (3) hierarchical system support for these new approaches.

Such a paradigm shift necessarily entailed a concomitant shift in emphasis away from inefficient, inflexible, hardware-based approaches toward higher level, more efficient and flexible software-based solutions. Consequently, the role of hardware-based mechanisms was redefined to that of providing and implementing the abstractions required to support the higher level software-based mechanisms in an integrated, hierarchical approach to ultradependable system design. This shift was furthermore compatible with an expanded view of "dependability," which had evolved to mean "the ability of the system to deliver the specified (or expected) service." Such a definition encompasses not only survival of traditional single hardware faults and environmental disturbances but more complex and less-well understood phenomena, as well: Byzantine faults, correlated errors, timing faults, software design and process interaction errors, and—most significantly—the unique issues encountered in real-

time systems in which faults and transient overload conditions must be detected and handled under hard deadline and resource constraints.

As sources of service disruption multiplied and focus shifted to their ultimate effects, traditional frameworks for reasoning about dependability had to be rethought. The classical fault/error/failure model, in which underlying anomalies *(faults)* give rise to incorrect values *(errors)*, which may ultimately cause incorrect behavior at the output *(failures)*, required extension to capture timing and performance issues. Graceful degradation, a long standing principle codifying performance/dependability trade-offs must be more carefully applied in real-time systems, where individual task requirements supercede general throughput optimization in any assessment. Indeed, *embedded* real-time systems—often characterized by interaction with physical sensors and actuators—may possess an inherent ability to tolerate brief periods of incorrect interaction, either in the values exchanged or the timing of those exchanges. Thus, a technical failure of the embedded *computer* does not necessarily imply a *system* failure. The challenge of capturing and modeling dependability for such potentially complex requirements is matched by the challenge of successfully exploiting them to devise more intelligent and efficient—as well as more complete—dependability mechanisms.

The evolution to a hierarchical, software-dominated approach would not have been possible without several enabling advances in hardware and software technology over the past decade:

(1) Advances in VLSI technology and RISC architectures have produced components with more chip real estate available for incorporation of efficient concurrent error detection mechanisms and more on-chip resources permitting software management of fine-grain redundancy;

(2) The emergence of practical parallel and distributed computing platforms possessing inherent coarse-grain redundancy of processing and communications resources—also amenable to efficient software-based management by either the system or the application;

(3) Advances in algorithms and languages for parallel and distributed computing leading to new insights in and paradigms for problem decomposition, module encapsulation, and module interaction, potentially exploitable in refining redundancy requirements and isolating faults;

(4) Advances in distributed operating systems allowing more efficient interprocess communication and more intelligent resource management;

(5) Advances in compiler technology that permit efficient, automatic instrumentation or restructuring of application code, program decomposition, and coarse and fine-grain resource management; and

(6) The emergence of fault-injection technology for conducting controlled experiments to determine the system and application-level manifestations of faults and evaluating the effectiveness or performance of fault-tolerance methods.

In response to this challenging, new vision for dependable computing research, the advent of the technological opportunities for realizing it, and its potential for addressing critical dependability needs of Naval, Defense, and commercial systems, the Office of Naval Research launched a five-year basic research initiative in 1990 in *Ultradependable Multicomputers and Electronic Systems* to accelerate and integrate progress in this important discipline. The objective of the initiative is to establish the fundamental principles as well as practical approaches for efficiently incorporating dependability into critical applications running on modern platforms. More specifically, the initiative sought increased effectiveness and efficiency through (1) Intelligent exploitation of the inherent redundancy available in modern parallel and distributed computers and VLSI components; (2) More precise characterization of the sources and manifestations of errors; (3) Exploitation of application semantics at all levels—code, task, algorithm, and domain—to allow optimization of fault-tolerance mechanisms to both application requirements and resource limitations; (4) Hierarchical, integrated software/hardware approaches; and (5) Development of scientific methods for evaluating and comparing candidate approaches.

Implementation of this broad mandate as a coherent research program necessitated focusing on a small cross-section of promising application-sensitive paradigms (including language, algorithm, and coordination-based approaches), their required hardware, compiler, and system support, and a few selected modeling and evaluation projects. In scope, the initiative emphasizes dependability primarily with respect to an expanded class of hardware and environment (both physical and operational) faults. Many of the efforts furthermore explicitly address issues of dependability unique to the domain of embedded real-time systems.

The success of the initiative and the significance of the research is demonstrated by the ongoing associations that many of our principal investigators have forged with a variety of military, Government, and commercial projects whose critical needs are leading to the rapid assimilation of concepts, approaches, and expertise arising from this initiative. Activities influenced to date include the FAA's Advanced Automation System for air traffic control, the Navy's AX project and Next Generation Computing Resources standards program, the Air Force's Center for Dependable Systems, the OSF/1 project, the space station Freedom, the Strategic

Defense Initiative, and research projects at GE, DEC, Tandem, the Naval Surface Warfare Center, and MITRE Corporation.

This book series is a compendium of papers summarizing the major results and accomplishments attained under the auspices of the ONR initiative in its first three years. Rather than providing a comprehensive text on dependable computing, the series is intended to capture the breadth, depth, and impact of recent advances in the field, as reflected through the specific research efforts represented, in the context of the vision articulated here. Each chapter does, however, incorporate appropriate background material and references. In view of the increasing importance and pervasiveness of real-time concerns in critical systems that impact our daily lives—ranging from multimedia communications to manufacturing to medical instrumentation—the real-time material is woven throughout the series rather than isolated in a single section or volume.

The series is partitioned into three volumes, corresponding to the three principal avenues of research identified at the beginning of this preface. While many of the chapters actually address issues at multiple levels, reflecting the comprehensive nature of the associated research project, they have been organized into these volumes on the basis of the primary conceptual contribution of the work. Agha and Sturman, for example, describe a framework (reflective architectures), a paradigm (replicated actors), and a prototype implementation (the Screed language and Broadway runtime system). But because the salient attribute of this work is the use of reflection to dynamically adapt an application to its environment, it is included in the *Frameworks* volume.

Volume I, *Models and Frameworks for Dependable Systems*, presents two comprehensive frameworks for reasoning about system dependability, thereby establishing a context for understanding the roles played by specific approaches presented throughout the series. This volume then explores the range of models and analysis methods necessary to design, validate, and analyze dependable systems.

Volume II, *Paradigms for Dependable Applications,* presents a variety of specific approaches to achieving dependability at the application level. Driven by the higher level fault models of Volume I and buiilt on the lower level abstractions implemented in Volume III, these approaches demonstrate how dependability may be tuned to the requirements of an application, the fault environment, and the characteristics of the target platform. Three classes of paradigms are considered: protocol-based paradigms for distributed applications, algorithm-based paradigms for parallel applications, and approaches to exploiting application semantics in embedded real-time control systems.

Volume III, *System Implementation,* explores the system infrastructure needed to support the various paradigms of Volume II. Approaches to implementing

suppport mechanisms and to incorporating additional appropriate levels of fault detection and fault tolerance at the processor, network, and operating system level are presented. A primary concern at these levels is balancing cost and performance against coverage and overall dependability. As these chapters demonstrate, low overhead, practical solutions are attainable and not necessarily incompatible with performance considerations. The section on innovative compiler support, in particular, demonstrates how the benefits of application specificity may be obtained while reducing hardware cost and run-time overhead.

This second volume of the series builds on the modeling foundation established in Volume I by exploring specific paradigms for managing redundancy and faults at the application level through specialized algorithms or protocols. Consistent with the layered view of dependability that characterizes this series, these software-oriented approaches rely not only on the underlying models of Volume I for their soundness, but on the abstractions of Volume III for their practicality.

In distributed systems, general-purpose dependability is often achieved through process replication managed through protocols. The three approaches described in Section 1 vary in purpose and degree of insulation from the application. Bianchini and Stahl explore the nuances of adapting distributed diagnosis algorithms to a real-time environment. Whereas the diagnosis paradigm is largely independent of the application, the authors demonstrate how consideration of the fault environment and scheduling constraints can lead to unanticipated modes of interaction. Marzullo, et al, present the refinement mapping approach for deriving customized dependable protocols for specific applications. The approach is illustrated through an air traffic control example. Finally, in an instantiation of Agha's concept of reflection (Vol. I), Schlichting, et al, consider two classes of language extensions to support enhanced application-specific control over redundancy and recovery management.

Parallel systems are characterized by larger degrees and finer granularity of concurrency than distributed systems. In such large-scale systems with frequent interprocess communication, conventional replication approaches are too costly, inefficient, and potentially detrimental to performance. Fortunately, unlike distributed applications which are typically decomposed by function, parallel scientific algorithms often employ data decomposition to assign each processor (running substantially the same program) a sub-domain corresponding, e.g., to a distinct region of physical space. The regular structure of these computations may be exploited through algorithmic transformations to provide low overhead error detection and recovery. Two such approaches are described in Section 2. Yajnik and Jha focus on the data by presenting a graph-theoretic methodology for generating check operations used to detect and locate faults. Kanellakis and Shvartsman exploit the homogenity of typical parallel tasks by allowing work to be dynamically redistributed in the event of failures.

Although real-time issues are addressed throughout this series, the tight coupling of embedded real-time systems to applications such as process control and the semantics of those applications—characterized by continuously changing physical variables—suggest an opportunity to explore highly effective and efficient dependability mechanisms that recognize potentially relaxed constraints derived from the additional latitude in error sensitivity typical of these applications. In Section 3, Liu, et al, present one such approach for managing redundancy and supporting rapid recovery under hard real-time constraints by trading off result quality for computation time. Bodson, et al, present a paradigm for software fault tolerance based on the concept of analytical redundancy in which the behavior of a complex control algorithm of uncertain integrity is monitored by a simpler, robust algorithm of similar but less refined functionality.

Gary M. Koob
Mathematical, Computer and Information Sciences Division
Office of Naval Research

Clifford G. Lau
Electronics Division
Office of Naval Research

ACKNOWLEDGEMENTS

The editors regret that, due to circumstances beyond their control, two planned contributions to this series could not be included in the final publications: "Compiler Generated Self-Monitoring Programs for Concurrent Detection of Run-Time Errors," by J.P. Shen and "The Hybrid Fault Effects Model for Dependable Systems," by C.J. Walter, M.M. Hugue, and N. Suri. Both represent significant, innovative contributions to the theory and practice of dependable computing and their omission diminishes the overall quality and completeness of these volumes.

The editors would also like to gratefully acknowledge the invaluable contributions of the following individuals to the success of the Office of Naval Research initiative in *Ultradependable Multicomputers and Electronic Systems* and this book series: Joe Chiara, George Gilley, Walt Heimerdinger, Robert Holland, Michelle Hugue, Miroslaw Malek, Tim Monaghan, Richard Scalzo, Jim Smith, André van Tilborg, and Chuck Weinstock.

CONTENTS

SECTION 1

PROTOCOL-BASED PARADIGMS FOR DISTRIBUTED APPLICATIONS

SECTION 1.1

Adaptive System-Level Diagnosis in Real-Time[1]

Mark E. Stahl[2]
Ronald P. Bianchini, Jr.[3]

Distributed real-time systems are subject to stricter fault-tolerance requirements than non-real time systems. This work presents an application of system-level diagnosis to a real-time distributed system as a first step in providing fault-tolerance. An existing algorithm for distributed system-level diagnosis, Adaptive_DSD, is converted to a real-time framework, establishing a deadline for the end-to-end diagnosis latency. Rate monotonic analysis is chosen as the framework for achieving real-time performance. The ADSD algorithm is converted into a set of independent periodic tasks running at each node, and a systematic procedure is used to assign priorities and deadlines to minimize the hard deadline of the diagnosis function. The resulting algorithm, Real-Time Adaptive Distributed System-Level Diagnosis (RT-ADSD), is fully compatible with a real-time environment, where both the processors and the network support fixed-priority scheduling. The RT-ADSD algorithm provides a useful first step in adding fault-tolerance to distributed real-time systems by quickly and reliably diagnosis node failures. The key results presented here include a framework for specifying real-time distributed algorithms and a scheduling model for analyzing them that accounts for many requirements of distributed systems, including network I/O, task jitter, and critical sections caused by shared resources.

---

1. This research is supported in party by the Office of Naval Research under Grant N00014-91-J-1304 and under a National Science Foundation Graduate Research Fellowship. Any opinions, findings, conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the Office of Naval Research or the National Science Foundation.

---

2. Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213.

---

3. Associate Professor, Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213.