

Preface

Fault tolerance, reliability, and availability are becoming major design issues nowadays in massively parallel distributed computing systems. Examples of systems in which fault tolerance is needed include mission-critical, computation-intensive, transactions (such as banking), and mobile/wireless computing systems/networks. High performance, measured in terms of speed and computing power, is essentially used as major design objective for such systems. It is however conceivable that great loss of crucial transactions can take place due to a small system/component error. The emergence of new paradigms, such as mobile/wireless computing, requires the introduction of new techniques for fault tolerance. It is therefore prudent that the issue of fault tolerance become among the set of design objectives of current and future computing systems.

A number of pressing challenges need to be faced in designing high performance computing systems. One such challenge is minimizing the performance loss due to the incorporation of fault tolerance techniques. Another challenge is to maximize the mean-time between failures (MTBF) of transaction-related systems, e.g., banking and airfare reservation systems. Yet another more critical challenge is to eliminate system interruptions due to fault(s) during the mission time of mission-critical systems, such as space crafts. These and other challenges require new innovative techniques for achieving fault tolerance in computing and networking systems.

My interest and commitment to the area of fault tolerance and reliability analysis started at the time when I was enrolled, as a doctoral candidate at the University of Toronto in Canada, in a graduate course entitled “Fault-Tolerant Computing”. The course was taught by my PhD thesis advisor Professor Zvonko Vranesic; a well known name in the area of Fault Tolerance and Testability Analysis. It was then that I came to

realize the importance of reliability and fault tolerance not only in computing but also in day to day life. Just imagine the devastation resulting from a situation such as an error prone elevator control circuitry in a 50-story building, or an error prone circuitry in a space craft sent to space with humans onboard, or even a small error in the control circuitry of an atomic reactor. It is clear that such possible situations make design for fault tolerance and reliability as important as design for optimization and/or performance gains.

The material presented in this book is meant to guide the design of computer systems with fault tolerance and reliability as basic design objectives. The book is an attempt to complement earlier efforts made by other colleagues on the same subject. It is comprehensive in its coverage and covers recent developments in the design of fault-tolerant and reliable systems, including mobile/wireless networking and distributed systems. The material presented in the book represents the outcome of more than twenty years of teaching undergraduate and graduate courses on Fault-Tolerant Computing and Reliability Engineering at institutions in Canada, Saudi Arabia, and Kuwait. During this period, I have developed my own class notes and improved it over the years based on the feedback obtained from my students and fellow colleagues. The material presented in the book has been class-tested. It evolved out of the class notes for the University of Saskatchewan's (Canada) CMPT 874, the King Fahd University of Petroleum and Minerals, Saudi Arabia (KFUPM) COE 421, COE 523 & CSE 641, and the Kuwait University (KU) ECE 564. The experiences gained in these courses have been incorporated into this book.

The book is divided into six parts. Each part consists of a number of chapters. Part I deals with the fundamental concepts involved in fault analysis. This part consists of three chapters: Fundamental Concepts, Fault Modeling, Simulation & Diagnosis, and Error Correcting Codes. The main objective in this part is to introduce the reader to the basic concepts in fault detection and correction. A number of fundamental concepts, simple reliability, and fault tolerance mathematical relations that will be used throughout the book are introduced in Part I of the book.

While Part I of the book deals primarily with a system consisting of a single component, Part II is devoted to a discussion of the fundamental issues related to fault tolerance and reliability analysis of computer systems consisting of more than one component. This part consists of three chapters: Fault Tolerance in Multi-Processor Systems, Fault-Tolerant Routing in Multi-Computer Systems, and Fault Tolerance and Reliability in Hierarchical Interconnection Networks. The discussion in these chapters spans a number of issues including the fault tolerance aspects of bus structures, fault tolerance and reliability features in simple multi-computer architectures, such as the hypercube and the mesh, and finally the achievement of various degrees of fault tolerance using hierarchical networks.

The reliability and fault tolerance of computer networks/distributed systems is the main theme in Part III of the book. These systems are characterized by having a number of processing units (nodes) collaborate together and execute a given task while being connected over a number of geographically distributed locations. Mobile/Wireless systems are good examples of distributed systems. Part III consists of four chapters: Fault Tolerance and Reliability of Computer Networks, Fault Tolerance in High Speed Switching Networks, Fault Tolerance in Distributed Computing Systems, and Fault Tolerance in Mobile Networks. With the present advances in VLSI/WSI technology, it is now economically feasible to implement massive parallel processor arrays consisting of a large number of fine-grained regular processing elements (PEs). The advantages offered by PEs include low assembly cost, high reliability, and high performance. Part IV of the book is dedicated to the discussion of the reliability and yield enhancement of array processors. This part consists of three chapters: Reliability and Yield Enhancement of VLSI/WSI Circuits, Design of Fault-Tolerant Processor Arrays, and Algorithm-Based Fault Tolerance Techniques.

Mission-critical systems incorporate massive parallel complex computing systems in solving real-time and computation-intensive problems. A small mistake in such systems can lead to devastating losses of human life. It is prudent in designing such systems to incorporate a sizeable degree of fault tolerance and reliability by incorporating

redundant processing elements. These redundant elements can be used to replace faulty ones. In performing such replacement, the system needs to be switched periodically to what is known as the *diagnostic mode*. During this mode, detection of faulty processing element(s) is performed. As the complexity of the diagnosed systems increases, the amount of data that needs to be analyzed also increases. Part V of the book is concerned with System Level Diagnosis. Part VI of the book is used to present two fault-tolerant and highly reliable/available practical systems. These are the Redundant Array of Independent Disks (RAID) systems and the TANDEM highly available systems, including a brief discussion on achieving high availability in Client/Server Computing paradigm.

Students in Computer Engineering, Information/Computer Sciences, and Electrical Engineering should find this book useful for their program of study. The majority of these programs include undergraduate as well as graduate courses in Fault-Tolerant Computing and Reliability Analysis of Computer Systems. Selected chapters can be used to offer core/elective courses in these disciplines. We offer the following suggestions: A one-semester course on “Introduction to Fault-Tolerant Computing” may cover Part I (three chapters), Chapter 4 (part of Part II), Chapter 7 (part of Part III) and Chapter 12 (part of Part IV). Another one-semester course on “Fault Tolerance and Reliability Analysis of Computer Networks” may cover Chapters 1 and 2 (part of Part I), Chapter 4 (part of Part II) and Part III (four chapters). Yet a third one-semester course on “Reliability and Fault Tolerance of Multi-Processor Systems” may cover Chapters 1 and 2 (part of Part I), Part II (three chapters), Chapter 8 (part of Part III) and Part IV (three chapters). In addition, selected individual chapters can be used to offer special topic courses with different emphasis. Above all, the book can also be used as a comprehensive reference for designers of fault-tolerant and reliable computer systems.

The book assumes that students studying fault tolerance must have had exposure to a basic course in logic design and a course on college-level probability/statistics. No other academic background is assumed of the readers in order to grasp, understand, and digest the material presented in the book and be able to apply what is presented in real life

designs. Every possible effort was made to make the book self-contained. Any feedback/comments are welcome on this aspect or any other related aspects. Comments can be sent to me through the publisher or directly to my emails: mostafa@cfw.kuniv.edu or mostafa@ccse.kfupm.edu.sa.

Mostafa Abd-El-Barr
KFUPM (Saudi Arabia) 2003 &
CFW, Kuwait University 2004-2006