



# Contents

|                        |              |
|------------------------|--------------|
| <b>Foreword</b>        | <b>ix</b>    |
| <b>Preface</b>         | <b>xv</b>    |
| <b>Acknowledgments</b> | <b>xxiii</b> |

|                  |  |    |
|------------------|--|----|
| <b>Chapter 1</b> | <b>Fundamentals of Computer Design</b>                     |    |
| 1.1              | Introduction   | 2  |
| 1.2              | Classes of Computers                                       | 4  |
| 1.3              | Defining Computer Architecture                             | 8  |
| 1.4              | Trends in Technology                                       | 14 |
| 1.5              | Trends in Power in Integrated Circuits                     | 17 |
| 1.6              | Trends in Cost   | 19 |
| 1.7              | Dependability  | 25 |
| 1.8              | Measuring, Reporting, and Summarizing Performance          | 28 |
| 1.9              | Quantitative Principles of Computer Design                 | 37 |
| 1.10             | Putting It All Together: Performance and Price-Performance | 44 |
| 1.11             | Fallacies and Pitfalls                                     | 48 |
| 1.12             | Concluding Remarks   | 52 |
| 1.13             | Historical Perspectives and References                     | 54 |
|                  | Case Studies with Exercises by Diana Franklin              | 55 |

|                  |   |     |
|------------------|---|-----|
| <b>Chapter 2</b> | <b>Instruction-Level Parallelism and Its Exploitation</b> |     |
| 2.1              | Instruction-Level Parallelism: Concepts and Challenges    | 66  |
| 2.2              | Basic Compiler Techniques for Exposing ILP                | 74  |
| 2.3              | Reducing Branch Costs with Prediction                     | 80  |
| 2.4              | Overcoming Data Hazards with Dynamic Scheduling           | 89  |
| 2.5              | Dynamic Scheduling: Examples and the Algorithm            | 97  |
| 2.6              | Hardware-Based Speculation                                | 104 |
| 2.7              | Exploiting ILP Using Multiple Issue and Static Scheduling | 114 |

|                  |   |     |
|------------------|---|-----|
| 2.8              | Exploiting ILP Using Dynamic Scheduling, Multiple Issue, and Speculation                  | 118 |
| 2.9              | Advanced Techniques for Instruction Delivery and Speculation                              | 121 |
| 2.10             | Putting It All Together: The Intel Pentium 4  | 131 |
| 2.11             | Fallacies and Pitfalls  | 138 |
| 2.12             | Concluding Remarks  | 140 |
| 2.13             | Historical Perspective and References   | 141 |
|                  | Case Studies with Exercises by Robert P. Colwell  | 142 |
| <b>Chapter 3</b> | <b>Limits on Instruction-Level Parallelism</b>  |     |
| 3.1              | Introduction  | 154 |
| 3.2              | Studies of the Limitations of ILP   | 154 |
| 3.3              | Limitations on ILP for Realizable Processors  | 165 |
| 3.4              | Crosscutting Issues: Hardware versus Software Speculation                                 | 170 |
| 3.5              | Multithreading: Using ILP Support to Exploit Thread-Level Parallelism                     | 172 |
| 3.6              | Putting It All Together: Performance and Efficiency in Advanced Multiple-Issue Processors | 179 |
| 3.7              | Fallacies and Pitfalls  | 183 |
| 3.8              | Concluding Remarks  | 184 |
| 3.9              | Historical Perspective and References   | 185 |
|                  | Case Study with Exercises by Wen-mei W. Hwu and John W. Sias                              | 185 |
| <b>Chapter 4</b> | <b>Multiprocessors and Thread-Level Parallelism</b>                                       |     |
| 4.1              | Introduction  | 196 |
| 4.2              | Symmetric Shared-Memory Architectures   | 205 |
| 4.3              | Performance of Symmetric Shared-Memory Multiprocessors                                    | 218 |
| 4.4              | Distributed Shared Memory and Directory-Based Coherence                                   | 230 |
| 4.5              | Synchronization: The Basics   | 237 |
| 4.6              | Models of Memory Consistency: An Introduction   | 243 |
| 4.7              | Crosscutting Issues   | 246 |
| 4.8              | Putting It All Together: The Sun T1 Multiprocessor  | 249 |
| 4.9              | Fallacies and Pitfalls  | 257 |
| 4.10             | Concluding Remarks  | 262 |
| 4.11             | Historical Perspective and References   | 264 |
|                  | Case Studies with Exercises by David A. Wood  | 264 |
| <b>Chapter 5</b> | <b>Memory Hierarchy Design</b>  |     |
| 5.1              | Introduction  | 288 |
| 5.2              | Eleven Advanced Optimizations of Cache Performance  | 293 |
| 5.3              | Memory Technology and Optimizations   | 310 |

|     |   |     |
|-----|---|-----|
| 5.4 | Protection: Virtual Memory and Virtual Machines       | 315 |
| 5.5 | Crosscutting Issues: The Design of Memory Hierarchies | 324 |
| 5.6 | Putting It All Together: AMD Opteron Memory Hierarchy | 326 |
| 5.7 | Fallacies and Pitfalls                                | 335 |
| 5.8 | Concluding Remarks                                    | 341 |
| 5.9 | Historical Perspective and References                 | 342 |
|     | Case Studies with Exercises by Norman P. Jouppi       | 342 |

## Chapter 6 **Storage Systems**

|      |   |     |
|------|---|-----|
| 6.1  | Introduction  | 358 |
| 6.2  | Advanced Topics in Disk Storage   | 358 |
| 6.3  | Definition and Examples of Real Faults and Failures                                 | 366 |
| 6.4  | I/O Performance, Reliability Measures, and Benchmarks                               | 371 |
| 6.5  | A Little Queuing Theory   | 379 |
| 6.6  | Crosscutting Issues   | 390 |
| 6.7  | Designing and Evaluating an I/O System—The Internet Archive Cluster                 | 392 |
| 6.8  | Putting It All Together: NetApp FAS6000 Filer                                       | 397 |
| 6.9  | Fallacies and Pitfalls  | 399 |
| 6.10 | Concluding Remarks  | 403 |
| 6.11 | Historical Perspective and References   | 404 |
|      | Case Studies with Exercises by Andrea C. Arpaci-Dusseau and Remzi H. Arpaci-Dusseau | 404 |

## Appendix A **Pipelining: Basic and Intermediate Concepts**

|      |   |      |
|------|---|------|
| A.1  | Introduction  | A-2  |
| A.2  | The Major Hurdle of Pipelining—Pipeline Hazards             | A-11 |
| A.3  | How Is Pipelining Implemented?                              | A-26 |
| A.4  | What Makes Pipelining Hard to Implement?                    | A-37 |
| A.5  | Extending the MIPS Pipeline to Handle Multicycle Operations | A-47 |
| A.6  | Putting It All Together: The MIPS R4000 Pipeline            | A-56 |
| A.7  | Crosscutting Issues   | A-65 |
| A.8  | Fallacies and Pitfalls                                      | A-75 |
| A.9  | Concluding Remarks  | A-76 |
| A.10 | Historical Perspective and References                       | A-77 |

## Appendix B **Instruction Set Principles and Examples**

|     |   |      |
|-----|---|------|
| B.1 | Introduction                              | B-2  |
| B.2 | Classifying Instruction Set Architectures | B-3  |
| B.3 | Memory Addressing                         | B-7  |
| B.4 | Type and Size of Operands                 | B-13 |
| B.5 | Operations in the Instruction Set         | B-14 |

|      |  |      |
|------|--|------|
| B.6  | Instructions for Control Flow                  | B-16 |
| B.7  | Encoding an Instruction Set                    | B-21 |
| B.8  | Crosscutting Issues: The Role of Compilers     | B-24 |
| B.9  | Putting It All Together: The MIPS Architecture | B-32 |
| B.10 | Fallacies and Pitfalls                         | B-39 |
| B.11 | Concluding Remarks                             | B-45 |
| B.12 | Historical Perspective and References          | B-47 |

## Appendix C    **Review of Memory Hierarchy**

|     |   |      |
|-----|---|------|
| C.1 | Introduction                              | C-2  |
| C.2 | Cache Performance                         | C-15 |
| C.3 | Six Basic Cache Optimizations             | C-22 |
| C.4 | Virtual Memory                            | C-38 |
| C.5 | Protection and Examples of Virtual Memory | C-47 |
| C.6 | Fallacies and Pitfalls                    | C-56 |
| C.7 | Concluding Remarks                        | C-57 |
| C.8 | Historical Perspective and References     | C-58 |

## **Companion CD Appendices**

### Appendix D    **Embedded Systems**

*Updated by Thomas M. Conte*

### Appendix E    **Interconnection Networks**

*Revised by Timothy M. Pinkston and José Duato*

### Appendix F    **Vector Processors**

*Revised by Krste Asanovic*

### Appendix G    **Hardware and Software for VLIW and EPIC**

### Appendix H    **Large-Scale Multiprocessors and Scientific Applications**

### Appendix I    **Computer Arithmetic**

*by David Goldberg*

### Appendix J    **Survey of Instruction Set Architectures**

### Appendix K    **Historical Perspectives and References**

**Online Appendix** ([textbooks.elsevier.com/0123704901](http://textbooks.elsevier.com/0123704901))

### Appendix L    **Solutions to Case Study Exercises**

**References** R-1

**Index** I-1