

Contents

Part I Architectures and Systems

| | |
|---|-----|
| 1 High Performance Network Architectures for Data Intensive Computing | 3 |
| Geng Lin and Eileen Liu | |
| 2 Architecting Data-Intensive Software Systems | 25 |
| Chris A. Mattmann, Daniel J. Crichton, Andrew F. Hart, Cameron Goodale, J. Steven Hughes, Sean Kelly, Luca Cinquini, Thomas H. Painter, Joseph Lazio, Duane Waliser, Nenad Medvidovic, Jinwon Kim, and Peter Lean | |
| 3 ECL/HPCC: A Unified Approach to Big Data | 59 |
| Anthony M. Middleton, David Alan Bayliss, and Gavin Halliday | |
| 4 Scalable Storage for Data-Intensive Computing | 109 |
| Abhishek Verma, Shivaram Venkataraman, Matthew Caesar, and Roy H. Campbell | |
| 5 Computation and Storage Trade-Off for Cost-Effectively Storing Scientific Datasets in the Cloud | 129 |
| Dong Yuan, Yun Yang, Xiao Liu, and Jinjun Chen | |

Part II Technologies and Techniques

| | |
|---|-----|
| 6 A Survey of Load Balancing Techniques for Data Intensive Computing | 157 |
| Zhiqian Sui and Shrideep Pallickara | |

| | | |
|--------------------------|--|-----|
| 7 | Resource Management for Data Intensive Clouds Through Dynamic Federation: A Game Theoretic Approach | 169 |
| | Mohammad Mehedi Hassan and Eui-Nam Huh | |
| 8 | Salt: Scalable Automated Linking Technology for Data-Intensive Computing | 189 |
| | Anthony M. Middleton and David Alan Bayliss | |
| 9 | Parallel Processing, Multiprocessors and Virtualization in Data-Intensive Computing | 235 |
| | Jonathan Burger, Richard Chapman, and Flavio Villanustre | |
| 10 | Challenges in Data Intensive Analysis at Scientific Experimental User Facilities | 249 |
| | Kerstin Kleese van Dam, Dongsheng Li, Stephen D. Miller, John W. Cobb, Mark L. Green, and Catherine L. Ruby | |
| 11 | Large-Scale Data Analytics Using Ensemble Clustering | 285 |
| | Martin Hahmann, Dirk Habich, and Wolfgang Lehner | |
| 12 | Specification of Data Intensive Applications with Data Dependency and Abstract Clocks | 323 |
| | Abdoulaye Gamatié | |
| 13 | Ensemble Feature Ranking Methods for Data Intensive Computing Applications | 349 |
| | Wilker Altidor, Taghi M. Khoshgoftaar, Jason Van Hulse, and Amri Napolitano | |
| 14 | Record Linkage Methodology and Applications | 377 |
| | Ling Qin Zhang | |
| 15 | Semantic Wrapper: Concise Semantic Querying of Legacy Relational Databases | 415 |
| | Naphtali Rische, Borko Furht, Malek Adjouadi, Armando Barreto, Debra Davis, Ouri Wolfson, Yelena Yesha, and Yaacov Yesha | |
| Part III Security | | |
| 16 | Security in Data Intensive Computing Systems | 447 |
| | Eduardo B. Fernandez | |
| 17 | Data Security and Privacy in Data-Intensive Computing Clusters ... | 467 |
| | Flavio Villanustre and Jarvis Robinson | |
| 18 | Information Security in Large Scale Distributed Systems | 485 |
| | Salvatore Distefano and Antonio Puliafito | |

19 Privacy and Security Requirements of Data Intensive Computing in Clouds 501
 Arash Nourian and Muthucumar Maheswaran

Part IV Applications

20 On the Processing of Extreme Scale Datasets in the Geosciences 521
 Sangmi Lee Pallickara, Matthew Malensek, and Shrideep Pallickara

21 Parallel Earthquake Simulations on Large-Scale Multicore Supercomputers 539
 Xingfu Wu, Benchun Duan, and Valerie Taylor

22 Data Intensive Computing: A Biomedical Case Study in Gene Selection and Filtering 563
 Michael Slavik, Xingquan Zhu, Imad Mahgoub, Taghi Khoshgoftar, and Ramaswamy Narayanan

23 Design Space Exploration for Efficient Data Intensive Computing on SoCs 581
 Rosilde Corvino, Abdoulaye Gamatié, and Pierre Boulet

24 Information Quality and Relevance in Large-Scale Social Information Systems 617
 Munmun De Choudhury

25 Geospatial Data Management with Terraflly 637
 Naphtali Rishe, Borko Furht, Malek Adjouadi, Armando Barreto, Evgenia Cheremisina, Debra Davis, Ori Wolfson, Nabil Adam, Yelena Yesha, and Yaacov Yesha

26 An Application for Processing Large and Non-Uniform Media Objects on MapReduce-Based Clusters 667
 Rainer Schmidt and Matthias Rella

27 Feature Selection Algorithms for Mining High Dimensional DNA Microarray Data 685
 David J. Dittman, Taghi M. Khoshgoftar, Randall Wald, and Jason Van Hulse

28 Application of Random Matrix Theory to Analyze Biological Data 711
 Feng Luo, Pradip K. Srimani, and Jizhong Zhou

29 Keyword Search on Large-Scale Structured, Semi-Structured, and Unstructured Data 733
Bin Zhou

30 A Distributed Publish/Subscribe System for Large Scale Sensor Networks 753
Masato Yamanouchi, Ryota Miyagi, Satoshi Matsuura,
Satoru Noguchi, Kazutoshi Fujikawa, and Hideki Sunahara

Index 777



<http://www.springer.com/978-1-4614-1414-8>

Handbook of Data Intensive Computing

Furht, B.; Escalante, A. (Eds.)

2011, XVIII, 794 p., Hardcover

ISBN: 978-1-4614-1414-8