

Table of Contents

Preface	xiii
----------------------	------

Section 1 **New Paradigms in Data Intensive Computing**

Chapter 1

Data-Aware Distributed Computing	1
<i>Esma Yildirim, State University of New York at Buffalo (SUNY), USA</i>	
<i>Mehmet Balman, Lawrence Berkeley National Laboratory, USA</i>	
<i>Tevfik Kosar, State University of New York at Buffalo (SUNY), USA</i>	

Chapter 2

Towards Data Intensive Many-Task Computing	28
<i>Ioan Raicu, Illinois Institute of Technology, USA & Argonne National Laboratory, USA</i>	
<i>Ian Foster, University of Chicago, USA & Argonne National Laboratory, USA</i>	
<i>Yong Zhao, University of Electronic Science and Technology of China, China</i>	
<i>Alex Szalay, Johns Hopkins University, USA</i>	
<i>Philip Little, University of Notre Dame, USA</i>	
<i>Christopher M. Moretti, University of Notre Dame, USA</i>	
<i>Amitabh Chaudhary, University of Notre Dame, USA</i>	
<i>Douglas Thain, University of Notre Dame, USA</i>	

Chapter 3

Micro-Services: A Service-Oriented Paradigm for Scalable, Distributed Data Management	74
<i>Arcot Rajasekar, University of North Carolina at Chapel Hill, USA</i>	
<i>Mike Wan, University of California at San Diego, USA</i>	
<i>Reagan Moore, University of North Carolina at Chapel Hill, USA</i>	
<i>Wayne Schroeder, University of California at San Diego, USA</i>	

Section 2

Distributed Storage

Chapter 4

Distributed Storage Systems for Data Intensive Computing	95
<i>Sudharshan S. Vazhkudai, Oak Ridge National Laboratory, USA</i>	
<i>Ali R. Butt, Virginia Polytechnic Institute and State University, USA</i>	
<i>Xiaosong Ma, North Carolina State University, USA</i>	

Chapter 5

Metadata Management in PetaShare Distributed Storage Network	118
<i>Ismail Akturk, Bilkent University, Turkey</i>	
<i>Xinqi Wang, Louisiana State University, USA</i>	
<i>Tevfik Kosar, State University of New York at Buffalo (SUNY), USA</i>	

Chapter 6

Data Intensive Computing with Clustered Chirp Servers.....	140
<i>Douglas Thain, University of Notre Dame, USA</i>	
<i>Michael Albrecht, University of Notre Dame, USA</i>	
<i>Hoang Bui, University of Notre Dame, USA</i>	
<i>Peter Bui, University of Notre Dame, USA</i>	
<i>Rory Carmichael, University of Notre Dame, USA</i>	
<i>Scott Emrich, University of Notre Dame, USA</i>	
<i>Patrick Flynn, University of Notre Dame, USA</i>	

Section 3

Data & Workflow Management

Chapter 7

A Survey of Scheduling and Management Techniques for Data-Intensive Application Workflows	156
<i>Suraj Pandey, The Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia</i>	
<i>Rajkumar Buyya, The University of Melbourne, Australia</i>	

Chapter 8

Data Management in Scientific Workflows	177
<i>Ewa Deelman, University of Southern California, USA</i>	
<i>Ann Chervenak, University of Southern California, USA</i>	

Chapter 9

Replica Management in Data Intensive Distributed Science Applications	188
<i>Ann L. Chervenak, University of Southern California, USA</i>	
<i>Robert Schuler, University of Southern California, USA</i>	

Section 4 Data Discovery & Visualization

Chapter 10

Data Intensive Computing for Bioinformatics.....	207
--	-----

Judy Qiu, Indiana University - Bloomington, USA
Jaliya Ekanayake, Indiana University - Bloomington, USA
Thilina Gunarathne, Indiana University - Bloomington, USA
Jong Youl Choi, Indiana University - Bloomington, USA
Seung-Hee Bae, Indiana University - Bloomington, USA
Yang Ruan, Indiana University - Bloomington, USA
Saliya Ekanayake, Indiana University - Bloomington, USA
Stephen Wu, Indiana University - Bloomington, USA
Scott Beason, Computer Sciences Corporation, USA
Geoffrey Fox, Indiana University - Bloomington, USA
Mina Rho, Indiana University - Bloomington, USA
Haixu Tang, Indiana University - Bloomington, USA

Chapter 11

Visualization of Large-Scale Distributed Data	242
---	-----

Jason Leigh, University of Illinois at Chicago, USA
Andrew Johnson, University of Illinois at Chicago, USA
Luc Renambot, University of Illinois at Chicago, USA
Venkatram Vishwanath, University of Illinois at Chicago, USA & Argonne National Laboratory, USA
Tom Peterka, Argonne National Laboratory, USA
Nicholas Schwarz, Northwestern University, USA

Chapter 12

On-Demand Visualization on Scalable Shared Infrastructure	275
---	-----

Huadong Liu, University of Tennessee, USA
Jinzhu Gao, University of The Pacific, USA
Jian Huang, University of Tennessee, USA
Micah Beck, University of Tennessee, USA
Terry Moore, University of Tennessee, USA

Compilation of References	291
---------------------------------	-----

About the Contributors	319
------------------------------	-----

Index.....	331
------------	-----