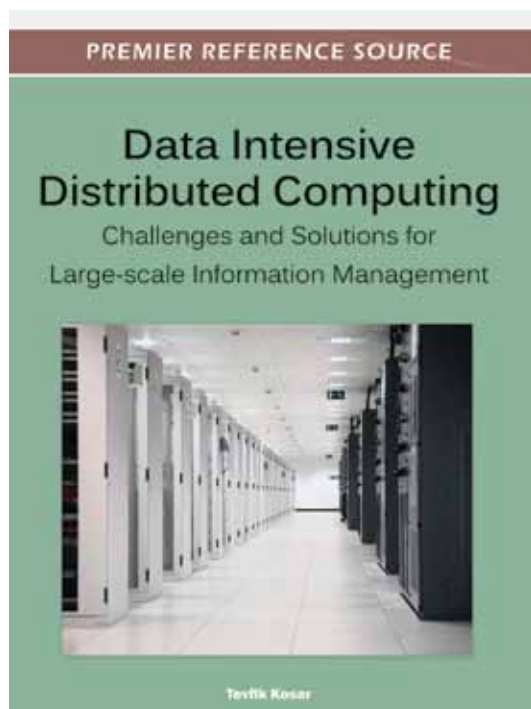


# An Excellent Addition to Your Library!

Released: January 2012

## Data Intensive Distributed Computing: Challenges and Solutions for Large-scale Information Management



Tevfik Kosar (University at Buffalo, USA)

The trend in scientific, as well as commercial, applications from a diverse range of fields has been towards being more and more data-intensive over time.

**Data Intensive Distributed Computing: Challenges and Solutions for Large-scale Information Management** focuses on the challenges of distributed systems imposed by data intensive applications and on the different state-of-the-art solutions proposed to overcome such challenges. Providing hints on how to manage low-level data handling issues when performing data intensive distributed computing, this publication is ideal for scientists, researchers, engineers, and application developers, alike. With the knowledge of the correct data management techniques for their applications, readers will be able to focus on their primary goal, assured that their data management needs are handled reliably and efficiently.

### Topics Covered:

- Bulk Data Movement
- Data and Workflow Management
- Data Staging and Replication
- Digital Libraries
- Hierarchical Storage Systems
- High Speed Data Streaming
- Metadata and Semantic Web
- Network-Aware Storage
- Parallel and Global File Systems
- Remote Access to Data
- Virtual Data Systems
- Visualization of Large Scale Data

ISBN: 9781615209712; © 2012; 352 pp.

Print: US \$180.00 | Perpetual: US \$255.00 | Print + Perpetual: US \$360.00

**Market:** This premier publication is essential for all academic and research library reference collections. It is a crucial tool for academicians, researchers, and practitioners and is ideal for classroom use.

**Tevfik Kosar** is an Associate Professor of Computer Science and Engineering at University at Buffalo (SUNY). He holds a Ph.D. degree in Computer Science from University of Wisconsin-Madison under the guidance of Prof. Miron Livny. Dr. Kosar's main research interests lie in the cross-section of petascale distributed systems, eScience, Grids, Clouds, and collaborative computing with a focus on large-scale data-intensive distributed applications. He is the primary designer and developer of the Stork distributed data scheduling system which has been adopted by many national and international institutions, and the lead investigator of the state-wide PetaShare distributed storage network in Louisiana. He has published more than fifty academic papers in leading journals and conferences. Some of the awards received by Dr. Kosar include NSF CAREER Award (for his work on "data-aware distributed computing"), LSU Rainmaker Award, LSU Flagship Faculty Award, Baton Rouge Business Report's Top 40 Under 40 Award, 1012 Corridor's Young Scientist Award, College of Basic Science's Research Award, and CCT Faculty of the Year Award. Dr. Kosar's work on data intensive computing has been funded by NSF, DOE, ONR, DoEd, SURA, and Louisiana Board of Regents.



www.igi-global.com

Publishing Academic Excellence  
at the Pace of Technology Since 1988

## Section 1: New Paradigms in Data Intensive Computing

### Chapter 1

#### *Data-Aware Distributed Computing*

Yildirim Esmat (State University of New York at Buffalo (SUNY), USA)  
Balman Mehmet (Lawrence Berkeley National Laboratory, USA)  
Kosar Tevfik (State University of New York at Buffalo (SUNY), USA)

### Chapter 2

#### *Towards Data Intensive Many-Task Computing*

Raicu Ioan (Illinois Institute of Technology, USA & Argonne National Laboratory, USA)  
Foster Ian (University of Chicago, USA & Argonne National Laboratory, USA)  
Zhao Yong (University of Electronic Science and Technology of China, China)  
Szalay Alex (Johns Hopkins University, USA)  
Little Philip (University of Notre Dame, USA)  
Moretti Christopher M. (University of Notre Dame, USA)  
Chaudhary Amitabh (University of Notre Dame, USA)  
Thain Douglas (University of Notre Dame, USA)

### Chapter 3

#### *Micro-Services:*

Rajasekar Arcot (University of North Carolina at Chapel Hill, USA)  
Wan Mike (University of California at San Diego, USA)  
Moore Reagan (University of North Carolina at Chapel Hill, USA)  
Schroeder Wayne (University of California at San Diego, USA)

## Section 2: Distributed Storage

### Chapter 4

#### *Distributed Storage Systems for Data Intensive Computing*

Vazhkudai Sudharshan S. (Oak Ridge National Laboratory, USA)  
Butt Ali R. (Virginia Polytechnic Institute and State University, USA)  
Ma Xiaosong (North Carolina State University, USA)

### Chapter 5

#### *Metadata Management in PetaShare Distributed Storage Network*

Akturk Ismail (Bilkent University, Turkey)  
Wang Xinqi (Louisiana State University, USA)  
Kosar Tevfik (State University of New York at Buffalo (SUNY), USA)

### Chapter 6

#### *Data Intensive Computing with Clustered Chirp Servers*

Thain Douglas (University of Notre Dame, USA)  
Albrecht Michael (University of Notre Dame, USA)  
Bui Hoang (University of Notre Dame, USA)  
Bui Peter (University of Notre Dame, USA)  
Carmichael Rory (University of Notre Dame, USA)  
Emrich Scott (University of Notre Dame, USA)  
Flynn Patrick (University of Notre Dame, USA)

## Section 3: Data & Workflow Management

### Chapter 7

#### *A Survey of Scheduling and Management Techniques for Data-Intensive Application Workflows*

Pandey Suraj (The Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia)  
Buyya Rajkumar (The University of Melbourne, Australia)

### Chapter 8

#### *Data Management in Scientific Workflows*

Deelman Ewa (University of Southern California, USA)  
Chervenak Ann (University of Southern California, USA)

### Chapter 9

#### *Replica Management in Data Intensive Distributed Science Applications*

Chervenak Ann L. (University of Southern California, USA)  
Schuler Robert (University of Southern California, USA)

## Section 4: Data Discovery & Visualization

### Chapter 10

#### *Data Intensive Computing for Bioinformatics*

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

### Chapter 11

#### *Visualization of Large-Scale Distributed Data*

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

### Chapter 12

#### *On-Demand Visualization on Scalable Shared Infrastructure*

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

## Order Your Copy Today!

Name: \_\_\_\_\_

Organization: \_\_\_\_\_

Address: \_\_\_\_\_

City, State, Zip: \_\_\_\_\_

Country: \_\_\_\_\_

Tel: \_\_\_\_\_

Fax: \_\_\_\_\_

E-mail: \_\_\_\_\_

Enclosed is check payable to IGI Global in  
US Dollars, drawn on a US-based bank

Credit Card  Mastercard  Visa  Am. Express

3 or 4 Digit Security Code: \_\_\_\_\_

Name on Card: \_\_\_\_\_

Account #: \_\_\_\_\_

Expiration Date: \_\_\_\_\_