2016 Seventh International Workshop on Data-Intensive Computing in the Clouds (DataCloud 2016)

Salt Lake City, Utah, USA
14 November 2016

IEEE Catalog Number: CFP16A52-POD
ISBN: 978-1-5090-6159-4

Copyright © 2016 by the Institute of Electrical and Electronics Engineers, Inc
All Rights Reserved

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number: CFP16A52-POD
ISBN (Print-On-Demand): 978-1-5090-6159-4
ISBN (Online): 978-1-5090-6158-7

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

2016 Seventh International Workshop on Data-Intensive Computing in the Clouds

DataCloud 2016

Table of Contents

Workshop Papers

Asterism: Pegasus and Dispel4py Hybrid Workflows for Data-Intensive Science 1
Rosa Filgueira, Rafael Ferreira da Silva, Amrey Krause, Ewa Deelman, and Malcolm Atkinson

An Ensemble-Based Recommendation Engine for Scientific Data Transfers 9
William Agnew, Michael Fischer, Ian Foster, and Kyle Chard

Pecos: A Scalable Solution for Analyzing and Managing Qualitative Data 17
Ritu Arora, Trung Nguyen Ba, and Tiffany A. Connors

Data-Intensive Supercomputing in the Cloud: Global Analytics for Satellite Imagery 24
Michael S. Warren, Samuel W. Skillman, Rick Chartrand, Tim Kelton, Ryan Keisler, David Raleigh, and Matthew Turk

Model Driven Advanced Hybrid Cloud Services for Big Data: Paradigm and Practice 32
Xi Yang and Tom Lehman

Improved Data-Aware Task Dispatching for Batch Queuing Systems 37
Xieming Li and Osamu Tatebe

A Multi-tenant Fair Share Approach to Full-text Search Engine 45
Zong Peng and Beth Plale

An Efficient Parallel Implementation of a Light-weight Data Privacy Method for Mobile Cloud Users 51
Mehdi Bahrami, Dong Li, Mukesh Singhal, and Ashish Kundu

Author Index 59