

2014 5th International Workshop on Data-Intensive Computing in the Clouds

(DataCloud 2014)

**New Orleans, Louisiana, USA
21 November 2014**



**IEEE Catalog Number: CFP14A52-POD
ISBN: 978-1-4799-7156-5**

2014 5th International Workshop on Data-Intensive Computing in the Clouds

DataCloud 2014

Table of Contents

Introduction	iv
Workshop Committee	v

Technical Papers

Towards Scalable Distributed Graph Database Engine for Hybrid Clouds	1
<i>Miyuru Dayarathna and Toyotaro Suzumura</i>	
To Overlap or Not to Overlap: Optimizing Incremental MapReduce Computations for On-Demand Data Upload	9
<i>Stefan Ene, Bogdan Nicolae, Alexandru Costan, and Gabriel Antoniu</i>	
Locality and Network-Aware Reduce Task Scheduling for Data-Intensive Applications	17
<i>Engin Arslan, Mrigank Shekhar, and Tefvik Kosar</i>	
Skyport - Container-Based Execution Environment Management for Multi-cloud Scientific Workflows	25
<i>Wolfgang Gerlach, Wei Tang, Kevin Keegan, Travis Harrison, Andreas Wilke, Jared Bischof, Mark D'Souza, Scott Devoid, Daniel Murphy-Olson, Narayan Desai, and Folker Meyer</i>	
Integrating Pig with Harp to Support Iterative Applications with Fast Cache and Customized Communication	33
<i>Tak-Lon Wu, Abhilash Koppula, and Judy Qiu</i>	
Author Index	40