# IEEE IPDPS HPDIC2012 Workshop Program

**FRIDAY, 25 May 2012**


Regal Shanghai East Asia Hotel

Shanghai, China

## Program at a Glance

<table>
<thead>
<tr>
<th>Time</th>
<th>Sessions</th>
<th>Titles/Papers</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00-08:30 AM</td>
<td>Keynote Speech 1</td>
<td>Computing: A Technology Challenger and A Computing Paradigm Changer</td>
<td>TBD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prof. Shyamasundar, Tata Institute of Fundamental Research Data Intensive</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>IEEE Fellow, ACM Fellow</td>
<td></td>
</tr>
<tr>
<td>08:30--09:00 AM</td>
<td>Keynote Speech 2</td>
<td>Toward High Efficiency Data-Intensive Scalable Computing</td>
<td>TBD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prof. Weisong Shi, Wayne State University</td>
<td></td>
</tr>
<tr>
<td>09:00--10:00 AM</td>
<td>Session 1 GPU/GPGPU Applications for Data Intensive Computing</td>
<td>HPDIC07 HPDIC20 HPDIC21 HPDIC26</td>
<td>TBD</td>
</tr>
<tr>
<td>10:00--10:15 AM</td>
<td>Coffee Break</td>
<td></td>
<td>TBD</td>
</tr>
<tr>
<td>10:15--12:15 AM</td>
<td>Session 2 MapReduce and Hadoop</td>
<td>HPDIC03 HPDIC06 HPDIC13 HPDIC14 HPDIC15 HPDIC22 HPDIC23 HPDIC30</td>
<td>TBD</td>
</tr>
<tr>
<td>12:15--13:30 PM</td>
<td>Lunch</td>
<td></td>
<td>TBD</td>
</tr>
<tr>
<td>13:30--15:00 PM</td>
<td>Session 3 Algorithms</td>
<td>HPDIC04 HPDIC10 HPDIC17 HPDIC24 HPDIC28 HPDIC29</td>
<td>TBD</td>
</tr>
<tr>
<td>15:00--15:30 PM</td>
<td>Coffee Break</td>
<td></td>
<td>TBD</td>
</tr>
<tr>
<td>15:30--18:30 PM</td>
<td>Session 4 Cloud,Grid,Virtualization, and Miscellaneous</td>
<td>HPDIC01 HPDIC12 HPDIC08 HPDIC09 HPDIC19 HPDIC25 HPDIC31 HPDIC32 HPDIC02 HPDIC05 HPDIC16</td>
<td>TBD</td>
</tr>
</tbody>
</table>
IEEE HPDIC 2012 Program Details

08:00-08:30 AM
Keynote Speech 1
Title: Data Intensive Computing: A Technology Challenger and A Computing Paradigm Changer
Speaker: Prof. Shyamasundar, Tata Institute of Fundamental ResearchData Intensive IEEE Fellow, ACM Fellow

Abstract: In the era of data “deluge”, data intensive computing, is an exciting research area with a spectrum of challenges to reach out to scalable science discoveries and enhance the quality of life. Further, it shall lead to computing paradigms that shall use innate human abilities that are beyond the comprehension of present-day computers. In this talk, we shall first walk through the technical challenges in realizing scalable solutions using cluster computing paradigms, innovative memory architectures including self aware systems, realizing a cloud for HPC. Subsequently, we shall address how various human capabilities are being harnessed through “interaction” that enables reaching beyond today’s computers.

Speaker Bio: Prof. R.K. Shyamasundar is with the Tata Institute of Fundamental Research, Mumbai, India where he is currently a Senior Professor and JC Bose National Fellow. He was the Founder Dean of the School of Technology and Computer Science. He took his B.E. (Electrical Engineering,) from University of Mysore, M.E. (Electrical Engineering) from Indian Institute of Science, Bangalore, and Ph.D.(Computer Science and Automation), from Indian Institute of Science, Bangalore. His principle areas of research are: Specification, Design and Verification of reactive and real-time systems, Programming Languages, Logics of programs, Formal methods, Computer and Network & Information Security. He has published widely and has more than 200 publications in refereed journals, conference proceedings, two monograph, several edited books, and holds several patents in US and India. He has given several invited talks at various conferences and has guided more than 35 Ph.D. students who are occupying leading positions in academia and industry. He is a Fellow of IEEE, Fellow of ACM and serves on IEEE Esterel Std Committee. He is a Fellow of the Indian Academy of Sciences, Fellow of the Indian National Science Academy, Fellow of the National Academy of Sciences, India, Fellow of the Indian National Academy of Engineering and Fellow of TWAS (Academy of Sciences for the Developing World, Trieste.)

http://www.tcs.tifr.res.in/~shyam/
Abstract: Our world is awash in data. Recent breakthroughs in science and technology generate more and more useful data rich in both quality and quantity. This opens many great opportunities for scientists to discover and advance to an unforeseen level of knowledge. However, it also introduces new challenges in the preservation, movement, access and analysis of massive datasets data beyond traditional high-performance computing. This talk will discuss about two fundamental challenges in data-intensive scalable computing (DISC) and propose potential solutions. The first part is about failure, which should be treated as an inherent feature in large-scale system. We propose the waste-free computing vision and an OPEn ReputAtion (OPERA) framework that tackles this problem. The second part is about the data moving cost. Instead of focusing on computation, we argue that data movement cost will dominate the total cost in the near future as moving large amount of data might take more time and resources than the computation itself. To reduce the moving cost, we propose two solutions dealing with the internal and external data movement of data centers.

Speaker Bio: Dr. Weisong Shi is an Associate Professor of Computer Science at Wayne State University. There he directs and co-directs the Mobile and Internet SysTemS Laboratory (MIST), Laboratory for Sustainable Computing (LaST), and Wayne Wireless Health Initiative, investigating performance, trust, power- and energy-efficiency issues of computer systems and applications. He received his B. E. from Xidian University in 1995, and Ph.D. degree from the Chinese Academy of Sciences in 2000, both in Computer Engineering. He authored one book, edited one book, published over 100 publications cited by 1800+ times, received research support from and consulted for a variety of governmental and industrial organizations, such as National Science Foundation, Department of Veteran Affairs, Air Force Research Laboratory, Swedish Research Council, Michigan Life Science Corridor, Huawei, Chrysler and so on. He has chaired several conferences and workshops, and served on technical program committees of numerous international conferences. He is currently serving on the editorial board of Elsevier Sustainable Computing, Journal of Computer Science and Technology (JCST) and International Journal of Sensor Networks. He served as a guest co-editor of several journals, including IEEE Internet Computing Magazine and Journal of Parallel and Distributed Computing. He is a recipient of Microsoft Fellowship (1999), the President Outstanding award of the Chinese Academy of Sciences (2000), the National Outstanding Ph.D. dissertation award of China (2002), the Faculty Research award of Wayne State University in 2004 and 2005, the CLAS teaching award in 2007 and 2011, the Best Paper award of ICWE’04 and IEEE IPDPS’05. Dr. Shi is a recipient of the NSF CAREER award (2007) and Wayne State University Career Development Chair award (2009). Dr. Shi’s H-index at Google Scholar is 21.
09:00--10:00 AM
Session 1 GPU/GPGPU Applications for Data Intensive Computing

HPDIC07
Chenxi Wang, Kang Kang, Maohua Zhu and Yangdong Deng
A Polyhedral Modeling Based Source-to-Source Code Optimization Framework for GPGPU

HPDIC20
Dan Chen and Lizhe Wang
A Massively Parallel Approach for Nonlinear Interdependency Analysis of Multivariate Signals with GPGPU

HPDIC21
Xianggao Cai and Xiaola Lin
Forecasting High Dimensional Volatility Using Conditional Restricted Boltzmann Machine on GPU

HPDIC26
Maria Malik, Lubomir Riha, Colin Shea and Tarek El-Ghazawi
Task Scheduling for GPU Accelerated Hybrid OLAP Systems with Multi-core Support and Text-to-Integer Translation

10:00--10:15 AM Coffee Break

10:15--12:15 AM
Session 2 MapReduce and Hadoop

HPDIC03
Tong Liu, Kathrini Wolfsgruber and Wensheng Guo
A Large-Scale Graph Learning Framework of Technological Gatekeepers by MapReduce

HPDIC06
Lizhe Wang and Dan Chen
MapReduce Across Distributed Clusters

HPDIC13
Zhuo Tang, Junqing Zhou, Kenli Li and Ruixuan Li
MTSD: A task scheduling algorithm for MapReduce base on deadline constraints

HPDIC14
Zengfeng Zeng, Bin Wu and Tiantian Zhang
A multi-source message passing model to improve the parallelism efficiency of graph mining on MapReduce

HPDIC15
Hiroya Matsuura, Masaru Ganse and Toyotaro Suzumura
A Highly Efficient Consolidated Platform for Stream Computing and Hadoop
Liang Chen, Li Kuang and Jian Wu
MapReduce based skyline services selection for QoS-aware composition

Hailong Yang, Zhongzhi Luan, Wenjun Li and Depei Qian
Statistics-based Workload Modeling for MapReduce

Hongwei Huo, Shuai Lin, Qiang Yu, Yipu Zhang and Vojislav Stojkovic
A MapReduce-based Algorithm for Motif Search

12:15--13:30 PM Lunch

Yuhang Liu, Mingfa Zhu, Jue Wang, Limin Xiao and Gong Tao
Xtorus: An Extended Torus Topology for On-Chip Massive Data Communication

Zhiming Zhao, Cosmin Dumitru, Paola Grosso, Cees de Laat
Network resource control for data intensive applications in heterogeneous infrastructures

Qian Cao and Min Zuo
A Scheduling Strategy Supporting OpenMP Task on Heterogeneous Multicore

Yang Shuangquan
Online scheduling with migration cost

Yunqin Zhong, Jizhong Han, Tieying Zhang, Zhenhua Li, Jinyun Fang, Guihai Chen
Towards Parallel Spatial Query Processing for Big Spatial Data

Huaiming Song, Hui Jin, Jun He, Xian-He Sun, Rajeev Thakur
A Server-Level Adaptive Data Layout Strategy for Parallel File Systems

15:00--15:30 PM Coffee Break

15:30--18:30 PM Session 4 Cloud, Grid, Virtualization, and Miscellaneous
HPDIC01
Xingwei Wang, Jiajia Sun, Huang Min and Xueyi Wang
A reverse auction based allocation mechanism in the cloud computing environment

HPDIC12
Shanshan Tu and Shaozhang Niu
Fine-grained Access Control and Revocation for Sharing Data on Clouds

HPDIC08
Javid Taheri and Albert Zomaya
A Pareto Frontier for Optimizing Data Transfer and Job Execution in Grids

HPDIC09
Jian Wan, Fei Pan and Congfeng Jiang
Placement Strategy of Virtual Machines Based on Workload Characteristics

HPDIC19
Li Zhou, Yicheng Wang, Jian Wan
Optimize Block-Level Cloud Storage System With Load-Balance Strategy

HPDIC25
Zhenyu Liu, Changjun Hu, Yang Li and Jingyuan Hu
DSDC: A Domain Scientific Data Cloud Based on Virtual Dataspaces

HPDIC31
Cheng Chen, Xiaomeng Huang, Haohuang Fu and Guangwen Yang
The Chunk-Locality Index: An Efficient Query Method for Climate Datasets

HPDIC32
Jian Wan, Dechuan Deng and Congfeng Jiang
Non-Cooperative Gaming and Bidding Model Based Resource Allocation in Virtual Machine Environment

HPDIC02
Yi Xie, Guoming Tang, Daifei Wang, Weidong Xiao, Daquan Tang and Jiuyang Tang
A Fault Tolerance Target-Tracking Strategy Based on Unreliable Sensing in Wireless Sensor Networks

HPDIC05
Hanfei Lv
One Double-Reduct Approach to Get Key Rules and the Experiment in Prison Computer Information Security

HPDIC16
Chuanzhi Liu, Chunping Ouyang and Yongbin Liu
Component Interface Testing in Virtual Experiment for Visualization of Material Property Data