

2016 International Conference on Parallel Architecture and Compilation Techniques (PACT 2016)

**Haifa, Israel
11-15 September 2016**



**IEEE Catalog Number: CFP16073-POD
ISBN: 978-1-5090-5308-7**

**Copyright © 2016, Association for Computing Machinery (ACM)
All Rights Reserved**

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

IEEE Catalog Number:	CFP16073-POD
ISBN (Print-On-Demand):	978-1-5090-5308-7
ISBN (Online):	978-1-4503-4121-9
ISSN:	1089-795X

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

CURRAN ASSOCIATES INC.
proceedings
.com

Table of Contents

PACT 2016 Conference Organization	ix
PACT 2016 Sponsors & Supporters	xii
Session 1: Keynote	
• Big Data Analytics on Flash Storage with Accelerators	1
Arvind (<i>Massachusetts Institute of Technology</i>)	
Session 2A: GPU - Architectures	
• Combating the Reliability Challenge of GPU Register File at Low Supply Voltage	3
Jingweijia Tan (<i>University of Houston</i>), Shuaiwen Leon Song (<i>Pacific Northwest National Lab</i>), Kaige Yan, Xin Fu (<i>University of Houston</i>), Andres Marquez, Darren Kerbyson (<i>Pacific Northwest National Lab</i>)	
• µC-States: Fine-grained GPU Datapath Power Management	17
Onur Kayiran (<i>Advanced Micro Devices, Inc.</i>), Adwait Jog (<i>College of William & Mary</i>), Ashutosh Pattnaik (<i>Pennsylvania State University</i>), Rachata Ausavarungnirun (<i>Carnegie Mellon University</i>), Xulong Tang, Mahmut T. Kandemir (<i>Pennsylvania State University</i>), Gabriel H. Loh (<i>Advanced Micro Devices, Inc.</i>), Onur Mutlu (<i>Carnegie Mellon University</i>), Chita R. Das (<i>Pennsylvania State University</i>)	
• Scheduling Techniques for GPU Architectures with Processing-In-Memory Capabilities	31
Ashutosh Pattnaik, Xulong Tang (<i>Pennsylvania State University</i>), Adwait Jog (<i>The College of William and Mary</i>), Onur Kayiran (<i>Advanced Micro Devices Inc.</i>), Asit K. Mishra (<i>Intel Corporation</i>), Mahmut T. Kandemir (<i>Pennsylvania State University</i>), Onur Mutlu (<i>Carnegie Mellon University</i>), Chita R. Das (<i>Pennsylvania State University</i>)	
• OAWS: Memory Occlusion Aware Warp Scheduling	45
Bin Wang (<i>Auburn University</i>), Yue Zhu, Weikuan Yu (<i>Florida State University</i>)	
Session 2B: Performance Optimizations	
• Integrating Algorithmic Parameters into Benchmarking and Design Space Exploration in 3D Scene Understanding	57
Bruno Bodin (<i>University of Edinburgh</i>), Luigi Nardi, M. Zeeshan Zia (<i>Imperial College London</i>), Harry Wagstaff, Govind Sreekar Shenoy (<i>University of Edinburgh</i>), Murali Emani (<i>Lawrence Livermore National Laboratory</i>), John Mawer, Christos Kotselidis, Andy Nisbet, Mikel Lujan (<i>University of Manchester</i>), Björn Franke (<i>University of Edinburgh</i>), Paul H. J. Kelly (<i>Imperial College London</i>), Michael O'Boyle (<i>University of Edinburgh</i>)	
• Fusion of Parallel Array Operations	71
Mads R. B. Kristensen, Simon A. F. Lund, Troels Blum, James Avery (<i>University of Copenhagen</i>)	
• Reduction Drawing: Language Constructs and Polyhedral Compilation for Reductions on GPUs	87
Chandan Reddy, Michael Kruse, Albert Cohen (<i>INRIA & École Normale Supérieure</i>)	
• Resource Conscious Reuse-Driven Tiling for GPUs	99
Prashant Singh Rawat, Changwan Hong (<i>The Ohio State University</i>), Mahesh Ravishankar Vinod Grover (<i>NVIDIA Corporation</i>), Louis-Noel Pouchet, Atanas Rountev, P. Sadayappan (<i>The Ohio State University</i>)	
Session 3: Best Paper	
• Accelerating Linked-list Traversal Through Near-Data Processing	113
Byungchul Hong, Gwangsun Kim (<i>Korea Advanced Institute of Science and Technology</i>), Jung Ho Ahn (<i>Seoul National University</i>), Yongkee Kwon, Hongsik Kim (<i>SK Hynix</i>), John Kim (<i>Korea Advanced Institute of Science and Technology</i>)	

- **Scalable Task Parallelism for NUMA: A Uniform Abstraction for Coordinated Scheduling and Memory Management**..... 125
Andi Drebes, Antoniu Pop (*The University of Manchester*), Karine Heydemann (*Sorbonne Universités*), Albert Cohen (*INRIA & École Normale Supérieure*), Nathalie Drach (*Sorbonne Universités*)
- **A Static Cut-off for Task Parallel Programs**..... 139
Shintaro Iwasaki, Kenjiro Taura (*The University of Tokyo*)

Session 4: Keynote

- **Greater Performance and Better Efficiency: Predicated Execution has Shown Us the Way** 151
Yale N. Patt (*The University of Texas at Austin*)

Session 5A: System Optimization I

- **WearCore: A Core for Wearable Workloads**..... 153
Sanyam Mehta, Josep Torrellas (*University of Illinois at Urbana-Champaign*)
- **Energy Aware Persistence: Reducing Energy Overheads of Memory-based Persistence in NVMs**..... 165
Sudarsun Kannan, Moinuddin Qureshi, Ada Gavrilovska, Karsten Schwan (*Georgia Institute of Technology*)
- **Power Tuning HPC Jobs on Power-Constrained Systems**..... 179
Neha Gholkar, Frank Mueller (*North Carolina State University*), Barry Rountree (*Lawrence Livermore National Laboratory*)
- **Online Scalability Characterization of Data-Parallel Programs on Many Cores** 191
Younghyun Cho, Surim Oh, Bernhard Egger (*Seoul National University*)

Session 5B: Parallel Software Optimization

- **Speculatively Exploiting Cross-Invocation Parallelism**..... 207
Jialu Huang, Prakash Prabhu (*Google Inc.*), Thomas B. Jablin (*University of Illinois at Urbana-Champaign*), Soumyadeep Ghosh, Sotiris Apostolakis (*Princeton University*), Jae W. Lee (*Sungkyunkwan University*), David I. August (*Princeton University*)
- **MicroSpec: Speculation-Centric Fine-Grained Parallelization for FSM Computations**..... 221
Junqiao Qiu, Zhijia Zhao (*University of California, Riverside*), Bin Ren (*The College of William and Mary*)
- **Hash Map Inlining** 235
Dibakar Gope, Mikko H. Lipasti (*University of Wisconsin - Madison*)
- **Sparso: Context-driven Optimizations of Sparse Linear Algebra** 247
Hongbo Rong, Jongsoo Park, Lingxiang Xiang, Todd A. Anderson, Mikhail Smelyanskiy (*Intel Corporation*)

Session 6A: Cache Coherence

- **Tardis 2.0: Optimized Time Traveling Coherence for Relaxed Consistency Models** 261
Xiangyao Yu (*Massachusetts Institute of Technology*), Hongzhe Liu (*Algonquin Regional High School*), Ethan Zou (*Lexington High School*), Srinivas Devadas (*Massachusetts Institute of Technology*)
- **Reducing Cache Coherence Traffic with Hierarchical Directory Cache and NUMA-Aware Runtime Scheduling**..... 275
Paul Caheny, Marc Casas, Miquel Moretó (*Universitat Politècnica de Catalunya*), Hervé Gloaguen, Maxime Saintes (*Bull Atos Technologies*), Eduard Ayguadé, Jesús Labarta, Mateo Valero (*Universitat Politècnica de Catalunya*)

Session 6B: Memory Access Efficiency

- **Characterizing and Optimizing the Performance of Multithreaded Programs Under Interference**..... 287
Yong Zhao, Jia Rao (*University of Texas at Arlington*), Qing Yi (*University of Colorado Colorado Springs*)
- **Optimizing Indirect Memory References with Milk** 299
Vladimir Kiriansky, Yunming Zhang, Saman Amarasinghe (*Massachusetts Institute of Technology*)

Session 7: Keynote

- **Scaling Data Analytics with Moore’s Law** 313
Kunle Olukotun (*Stanford University*)

Session 8A: System Acceleration

- **Bridging the Semantic Gaps of GPU Acceleration for Scale-out CNN-based Big Data Processing: Think Big, See Small** 315
Mingcong Song, Yang Hu (*University of Florida*), Yunlong Xu (*Xi’an Jiaotong University*),
Chao Li (*Shanghai Jiao Tong University*), Huixiang Chen (*University of Florida*),
Jingling Yuan (*Wuhan University of Technology*), Tao Li (*University of Florida*)
- **A DSL Compiler for Accelerating Image Processing Pipelines on FPGAs**..... 327
Nitín Chugh (*International Institute of Information Technology*),
Vinay Vasista (*Indian Institute of Science*), Suresh Purini (*International Institute of Information Technology*),
Uday Bondhugula (*Indian Institute of Science*)
- **Automatically Exploiting Implicit Pipeline Parallelism from Multiple Dependent Kernels for GPUs** 339
Gwangsun Kim, Jiyun Jeong, John Kim (*Korea Advanced Institute of Science and Technology*),
Mark Stephenson (*NVIDIA*)
- **CAF: Core to Core Communication Acceleration Framework**..... 351
Yipeng Wang (*North Carolina State University*), Ren Wang, Andrew Herdrich, James Tsai (*Intel Corporation*),
Yan Solihin (*North Carolina State University*)

Session 8B: System Optimization II

- **Vectorization of Multibyte Floating Point Data Formats** 363
Andrew Anderson, David Gregg (*Trinity College*)
- **Rinnegan: Efficient Resource Use in Heterogeneous Architectures** 373
Sankaralingam Panneerselvam, Michael Swift (*University of Wisconsin, Madison*)
- **Auto-tuning Spark Big Data Workloads on POWER8: Prediction-Based Dynamic SMT Threading**..... 387
Zhen Jia (*Chinese Academy of Sciences*), Chao Xue, Guancheng Chen (*IBM Research-China*),
Jianfeng Zhan, Lixin Zhang (*Chinese Academy of Sciences*),
Yonghua Lin (*IBM Research-China*), Peter Hofstee (*IBM Research-Austin*)
- **EXCITE-VM: Extending the Virtual Memory System to Support Snapshot Isolation Transactions** 401
Heiner Litz, Benjamin Braun, David Cheriton (*Stanford University*)

Poster Presentations

- **POSTER: Fly-Over: A Light-Weight Distributed Power-Gating Mechanism for Energy-Efficient Networks-on-Chip** 413
Rahul Boyapati, Jiayi Huang (*Texas A&M University*), Ningyuan Wang (*Google Inc.*),
Kyung Hoon Kim, Ki Hwan Yum, Eun Jung Kim (*Texas A&M University*)
- **POSTER: Exploiting Asymmetric Multi-Core Processors with Flexible System Software** 415
Kallia Chronaki, Miquel Moretó, Marc Casas (*Barcelona Supercomputing Center*), Alejandro Rico (*ARM*),
Rosa M. Badia, Eduard Ayguadé, Jesus Labarta, Mateo Valero (*Barcelona Supercomputing Center*)
- **POSTER: Easy PRAM-based High-performance Parallel Programming with ICE**..... 419
Fady Ghanim, Rajeev Barua, Uzi Vishkin (*University of Maryland*)
- **POSTER: Fault-tolerant Execution on COTS Multi-core Processors with Hardware Transactional Memory Support**..... 421
Florian Haas, Sebastian Weis, Theo Ungerer (*University of Augsburg*),
Gilles Pokam, Youfeng Wu (*Intel Corporation*)
- **POSTER: Collective Dynamic Parallelism for Directive Based GPU Programming Languages and Compilers** 423
Guray Ozen, Eduard Ayguade, Jesus Labarta (*Universitat Politecnica de Catalunya*)

- **POSTER – Firestorm: Operating Systems for Power-Constrained Architectures**425
Sankaralingam Pannecerselvam, Michael Swift (*University of Wisconsin, Madison*)
- **POSTER: ξ -TAO: A Cache-centric Execution Model and Runtime for Deep Parallel Multicore Topologies**.....429
Miquel Pericàs (*Chalmers University of Technology*)
- **POSTER: Efficient Self-Invalidation/Self-Downgrade for Critical Sections with Relaxed Semantics**433
Alberto Ros (*Universidad de Murcia*), Carl Leonardsson, Christos Sakalis, Stefanos Kaxiras (*Uppsala Universitet*)
- **POSTER: SILC-FM: Subblocked InterLeaved Cache-Like Flat Memory Organization**435
Jee Ho Ryoo (*The University of Texas at Austin*), Mitesh R. Meswani (*AMD*), Reena Panda, Lizy K. John (*The University of Texas at Austin*)
- **POSTER: Hybrid Data Dependence Analysis for Loop Transformations**439
Diogo Sampaio, Alain Ketterlin (*Institut National de Recherche en Informatique*), Louis-Noël Pouchet (*The Ohio State University*), Fabrice Rastello (*Institut National de Recherche en Informatique*)
- **POSTER: An Optimization of Dataflow Architectures for Scientific Applications**441
Xiaowei Shen, Xiaochun Ye, Xu Tan, Da Wang, Zhimin Zhang, Dongrui Fan, Zhimin Tang (*Chinese Academy of Sciences & University of Chinese Academy of Sciences*)
- **POSTER: hVISC: A Portable Abstraction for Heterogeneous Parallel Systems**.....443
Prakalp Srivastava, Maria Kotsifakou, Matthew D. Sinclair (*University of Illinois at Urbana-Champaign*), Rakesh Komuravelli (*Qualcomm Technologies Inc.*), Vikram Adve, Sarita Adve (*University of Illinois at Urbana-Champaign*)
- **POSTER: An Integrated Vector-Scalar Design on an In-order ARM Core**.....447
Milan Stanic (*Barcelona Supercomputing Center*), Oscar Palomar (*University of Manchester*), Timothy Hayes, Ivan Ratkovic, Osman Unsal, Arian Cristal, Mateo Valero (*Barcelona Supercomputing Center*)
- **POSTER: Pagoda: A Runtime System to Maximize GPU Utilization in Data Parallel Tasks with Limited Parallelism**.....449
Tsung Tai Yeh, Amit Sabne, Putt Sakdhnagool, Rudolf Eigenmann, Timothy G. Rogers (*Purdue University*)

Student Research Poster Presentations

- **Student Research Poster: Slack-Aware Shared Bandwidth Management in GPUs**451
Saumay Dublsh (*University of Edinburgh*)
- **Student Research Poster: From Processing-in-Memory to Processing-in-Storage**.....453
Roman Kaplan (*Technion*)
- **Student Research Poster: Network Controller Emulation on a Sidecore for Unmodified Virtual Machines**.....454
Arthur Kiyanovski (*Technion*)
- **Student Research Poster: A Low Complexity Cache Sharing Mechanism to Address System Fairness**.....455
Vicent Selfa, Julio Sahuquillo, Salvador Petit, María E. Gómez (*Universitat Politècnica de València*)
- **Student Research Poster: A Scalable General Purpose System for Large-Scale Graph Processing**.....456
Jiawen Sun (*High Performance and Distributed Computing*)
- **Student Research Poster: Compiling Boolean Circuits to Non-deterministic Branching Programs to be Implemented by Light Switching Circuits**.....457
Vladislav Tartakovsky (*Haiifa University*)
- **Student Research Poster: Software Out-of-Order Execution for In-Order Architectures**458
Kim-Anh Tran (*Uppsala University*)

- Author Index**.....459