FPGA 2017

Conference Program of the
25ᵗʰ ACM/SIGDA International Symposium on
Field-Programmable Gate Arrays

February 22 - 24, 2017

Monterey Marriott Hotel, Monterey, California

FPGA 2017 Sponsors and Corporate Patrons

Sponsors

Gold Corporate Patrons

Silver Corporate Patrons

Logistics Support:

Wednesday February 22 (All Technical Sessions in San Carlos 2-4)

Int'l Workshop on Overlay Architectures for FPGAs (OLAF)
Chair: Hayden So, The University of Hong Kong; Co-Chair: John Wawryznek, UC Berkeley

9:00	Welcome and Opening Remarks *John Wawrzynek (UC Berkeley)*
9:10	Paper Presentations (http://olaf.eecs.berkeley.edu/program)
12:00	**Lunch (Ferrantes Room, 10th Floor)**

Afternoon Special Session: The Role of FPGAs in Machine Learning
Chair: Andrew Ling, Intel

1:30	Deep Learning -- Tutorial and Recent Trends *Song Han (Stanford and DeePhi)*
2:30	**Can FPGAs Beat GPUs in Accelerating Next-Generation Deep Neural Networks?** *Eriko Nurvitadhi, Ganesh Venkatesh, Jaewoong Sim, Debbie Marr, Randy Huang, Jason Gee Hock Ong, Yeong Tat Liew, Srivatsan Krishnan, Duncan Moss, Suchit Subhaschandra, Guy Boudoukh*, Intel

3:00 Break

3:30	**Accelerating Binarized Convolutional Neural Networks with Software-Programmable FPGAs** *Ritchie Zhao[1], Weinan Song[1], Wentao Zhang[1], Tianwei Xing[2], Jeng-Hau Lin[3], Mani Srivastava[2], Rajesh Gupta[3], Zhiru Zhang[1]* [1]Cornell University, [2]UCLA, [3]UCSD
3:55	**Improving the Performance of OpenCL-based FPGA Accelerator for Convolutional Neural Network** *Jialiang Zhang and Jing Li,* UW-Madison
4:20	**Frequency Domain Acceleration of Convolutional Neural Networks on CPU-FPGA Shared Memory System** *Chi Zhang and Viktor Prasanna,* USC
4:45	**Optimizing Loop Operation and Dataflow in FPGA Acceleration of Deep Convolutional Neural Networks** *Yufei Ma, Yu Cao, Sarma Vrudhula, Jae-sun Seo*, Arizona State University
7:00	**Opening Reception (Ferrantes Room, 10th Floor)**

Thursday February 23 (All Technical Sessions in San Carlos 2-4)

8:00   **Continental Breakfast**

8:45   Welcome and Opening Remarks
Jason Anderson (University of Toronto), Jonathan Greene (Microsemi)

Machine Learning
Chair: Jason Cong, UCLA

9:00   **An OpenCL Deep Learning Accelerator on Arria 10**
Utku Aydonat, Shane O'Connell, Davor Capalija, Andrew Ling, Gordon Chiu, Intel

9:25   **FINN: A Framework for Fast, Scalable Binarized Neural Network Inference**
Yaman Umuroglu[1], Nicholas J. Fraser[2], Giulio Gambardella[3], Michaela Blott[3], Philip Leong[4], Magnus Jahre[5], Kees Vissers[3]
[1]Xilinx and Norwegian University of Science and Technology, [2]Xilinx and University of Sydney, [3]Xilinx, [4]University of Sydney, [5]Norwegian University of Science and Technology

9:50   **ESE: Efficient Speech Recognition Engine with Compressed LSTM on FPGA**
Song Han[1], Junlong Kang[2], Huizi Mao[1], Yiming Hu[3], Xin Li[2], Yubin Li[2], Dongliang Xie[2], Hong Luo[2], Song Yao[2], Yu Wang[3], Huazhong Yang[3], Bill Dally[1]
[1]Stanford University, [2]DeePhi, [3]Tsinghua University

10:15   **Poster Session 1 and Break (San Carlos 1)**

Interconnect and Routing
Chair: Sinan Kaptanoglu, Microsemi

11:15   **Quality-Time Tradeoffs in Component-Specific Mapping**
Hans Giesen[1], Raphael Rubin[1], Benjamin Gojman[2], Andre DeHon[1]
[1]University of Pennsylvania, [2]Google

11:40   **Synchronization Constraints for Interconnect Synthesis**
Alex Rodionov and Jonathan Rose, University of Toronto

12:05   **Corolla: GPU-Accelerated FPGA Routing Based on Subgraph Dynamic Expansion**
Minghua Shen and Guojie Luo, Peking University

12:30   **Lunch (Ferrantes Room, 10th Floor)**

Architecture
Chair: Steve Wilton, University of British Columbia

2:00   **Don't Forget the Memory: Automatic Block RAM Modelling, Optimization, and Architecture Exploration**
Sadegh Yazdanshenas, Kosuke Tatsumura, Vaughn Betz, University of Toronto

2:25   **Automatic Construction of Program-Optimized FPGA Memory Networks**
Hsin-Jung Yang[1], Kermin Fleming[2], Felix Winterstein[3], Annie Chen[1], Michael Adler[4], Joel Emer[1]
[1]MIT, [2]Intel, [3]Imperial College London, [4]Intel Corporation

2:50	**NAND-NOR: A Compact, Fast, and Delay Balanced FPGA Logic Element** *Zhihong Huang[1], Xing Wei[1], Grace Zgheib[2], Wei Li[1], Yu Lin[1], Zhenghong Jiang[1], Kaihui Tu[1], Paolo Ienne[2], Haigang Yang[1]* [1]Chinese Academy of Sciences, [2]EPFL
2:55	**120-core microAptiv MIPS Overlay for the Terasic DE5-NET FPGA board** *Nachiket Kapre[1], Prashanth Ravi[2], Gourav Modi[3], Chethan Kumar H B[2]* [1]University of Waterloo, [2]Nanyang Technological University, [3]NTU
3:00	**Poster Session 2 and Break (San Carlos 1)**

CAD Tools
Chair: Lesley Shannon, Simon Fraser University

4:00	**A Parallelized Iterative Improvement Approach to Area Optimization for LUT-Based Technology Mapping** *Gai Liu and Zhiru Zhang*, Cornell University
4:25	**A Parallel Bandit-Based Approach for Autotuning FPGA Compilation** *Chang Xu[1], Gai Liu[2], Ritchie Zhao[2], Stephen Yang[3], Guojie Luo[1], Zhiru Zhang[2]* [1]Peking University, [2]Cornell University, [3]Xilinx
6:30	**Banquet (San Carlos 2-4)**
7:45	**Banquet Panel: FPGAs in the Cloud** Chair: George Constantinides, Imperial College London Panelists: Andrew Putnam (Microsoft, USA), Wei Qi (Baidu, China), Gaurav Singh (Xilinx, USA), Mark Shand (Waymo, USA), Ling Shao (IBM Research, China), Richard Veitch (Maxeler, USA)

Friday February 24 (All Technical Sessions in San Carlos 2-4)

8:00	**Continental Breakfast**

High-Level Synthesis -- Tools and Applications
Chair: Stephen Neuendorffer, Xilinx

9:00	**Hardware Synthesis of Weakly Consistent C Concurrency** *Nadesh Ramanathan, Shane Fleming, John Wickerson, George Constantinides,* Imperial College London
9:25	**A New Approach to Automatic Memory Banking using Trace-Based Address Mining** *Yuan Zhou, Khalid Al-Hawaj, Zhiru Zhang*, Cornell University
9:50	**Dynamic Hazard Resolution for Pipelining Irregular Loops in High-Level Synthesis** *Steve Dai[1], Ritchie Zhao[1], Gai Liu[1], Shreesha Srinath[1], Udit Gupta[2], Christopher Batten[1], Zhiru Zhang[1]* [1]Cornell University, [2]Harvard University

Accelerating Face Detection on Programmable SoC Using C-Based Synthesis

9:55 *Nitish Srivastava[1], Steve Dai[1], Rajit Manohar[2], Zhiru Zhang[1]*
[1]Cornell University, [2]Cornell NYC Tech

Packet Matching on FPGAs Using HMC Memory: Towards One Million Rules

10:00 *Daniel Rozhko, Geoffrey Elliott, Daniel Ly-Ma, Paul Chow, Hans-Arno Jacobsen*
University of Toronto

10:05 **Poster Session 3 and Break (San Carlos 1)**

Graph Processing Applications
Chair: Nachiket Kapre, University of Waterloo

Boosting the Performance of FPGA-based Graph Processor using Hybrid using Hybrid Memory Cube: A Case for Breadth First Search

11:00 *Jialiang Zhang, Soroosh Khoram, Jing Li*, UW-Madison

ForeGraph: Exploring Large-scale Graph Processing on Multi-FPGA Architecture

11:25 *Guohao Dai[1], Tianhao Huang[1], Yuze Chi[2], Ningyi Xu[3], Yu Wang[1], Huazhong Yang[1]*
[1]Tsinghua University, [2]UCLA, [3]Microsoft Research Asia

FPGA-Accelerated Transactional Execution of Graph Workloads

11:50 *Xiaoyu Ma[1], Dan Zhang[2], Derek Chiou[3]*
[1]UT-Austin, [2]University of Texas at Austin, [3]Microsoft/UT Austin

12:15 **Lunch (Ferrantes Room, 10th Floor)**

Virtualization and Applications
Chair: John Lockwood, Algo-Logic Systems

Enabling Flexible Network FPGA Clusters in a Heterogenous Cloud Data Center

2:00 *Naif Tarafdar, Thomas Lin, Eric Fukuda, Hadi Bannazadeh, Alberto Leon-Garcia, Paul Chow*, University of Toronto

Energy Efficient Scientific Computing on FPGAs using OpenCL

2:25 *Dennis Weller[1], Fabian Oboril[1], Dimitar Lukarski[2], Juergen Becker[1], Mehdi Tahoori[1]*
[1]Karlsruhe Instistute of Technology, [2]PARALUTION Labs

Secure Function Evaluation using an FPGA Overlay Architecture

2:50 *Xin Fang, Stratis Ioannidis, Miriam Leeser*, Northeastern University

3:15 **Break**

Applications
Chair: Miriam Leeser, Northeastern University

3:45 **FPGA Acceleration for Computational Glass-Free Displays**
Zhuolun He and Guojie Luo, Peking University

4:10 **Hardware Acceleration of the Pair-HMM Algorithm for DNA Variant Calling**
Sitao Huang[1], Gowthami Jayashri Manikandan[2], Anand Ramachandran[2], Kyle Rupnow[3], Wen-mei W. Hwu[2], Deming Chen[2]
[1]University of Illinois at Urbana-Champaign, [2]UIUC, [3]Advanced Digital Sciences Center

4:35 **Conference Closing and Best Paper Award**
Jason Anderson (University of Toronto), Jonathan Greene (Microsemi)

Poster Session 1

Measuring the Power-Constrained Performance and Energy Gap between FPGAs and Processors
Andy Ye[1] and Karthik Ganesan[2]
[1]Ryerson University, [2]University of Toronto

A Mixed-Signal Data-Centric Reconfigurable Architecture enabled by RRAM Technology
Yue Zha[1], Jialiang Zhang[1], Zhiqiang Wei[2], Jing Li[1]
[1]UW-Madison, [2]Panasonic

A Framework for Iterative Stencil Algorithm Synthesis on FPGAs from OpenCL Programming Model
Shuo Wang and Yun Liang, Peking University

Scala Based FPGA Design Flow
Yanqiang Liu[1], Yao Li[1], Weilun Xiong[1], Meng Lai[1], Cheng Chen[2], Zhengwei Qi[1], Haibing Guan[1]
[1]Shanghai JiaoTong University, [2]Morgan Stanley

Thermal Flattening in 3D FPGAs using Embedded Cooling
Girish Deshpande and Dinesh Bhatia, UT-Dallas

A Machine Learning Framework for FPGA Placement
Gary Grewal, Shawki Areibi, Matthew Westrik, Ziad Abuowaimer, Betty Zhao,
University of Guelph

Precise Coincidence Detection on FPGAs: Three Case Studies
Ralf Salomon and Ralf Joost, University of Rostock

DTP: Enabling Exhaustive Exploration of FPGA Temporal Partitions for Streaming HPC Applications
Mostafa Koraei[1], Magnus Jahre[2], S.Omid Fatemi[1]
[1]Tehran University, [2]Norwegian University of Science and Technology

Accurate and Efficient Hyperbolic Tangent Activation Function on FPGA using the DCT Interpolation Filter

Ahmed Abdelsalam, Pierre Langlois, Farida Cheriet, École Polytechnique de Montréal

An FPGA Overlay Architecture for Cost Effective Regular Expression Search

Thomas Luinaud, J.M. Pierre Langlois, Yvon Savaria, École Polytechnique de Montréal

Storage-Efficient Batching for Minimizing Bandwidth of Fully-Connected Neural Network Layers

Yongming Shen, Michael Ferdman, Peter Milder, Stony Brook University

Poster Session 2

Using Vivado-HLS for Structural Design: a NoC Case Study

Zhipeng Zhao and James C. Hoe, CMU

Automatic Generation of Hardware Sandboxes for Trojan Mitigation in Systems on Chip

Christophe Bobda[1], Taylor Whitaker[1], Charles Kamhoua[2], Kevin Kwiat[2], Laurent Njilla[2]
[1]University of Arkansas, [2]Air Force Research Lab

Accelerating Financial Market Server through Hybrid List Design

Haohuan Fu[1], Conghui He[1], Huabin Ruan[1], Itay Greenspon[2], Wayne Luk[3], Yongkang Zheng[4], Junfeng Liao[1], Qing Zhang[4], Guangwen Yang[1]
[1]Tsinghua University, [2]Maxeler Technologies, [3]Imperial College London, [4]China Financial Futures Exchange

Joint Modulo Scheduling and Memory Partitioning with Multi-Bank Memory for High-Level Synthesis

Tianyi Lu, Shouyi Yin, Xianqing Yao, Zhicong Xie, Leibo Liu, Shaojun Wei
Tsinghua University

A Batch Normalization Free Binarized Convolutional Deep Neural Network on an FPGA

Hiroki Nakahara[1], Haruyoshi Yonekawa[1], Hisashi Iwamoto[2], Masato Motomura[3]
[1]Tokyo Institute of Technology, [2]Poco a Poco Networks, [3]Hokkaido University

A 7.663-TOPS 8.2-W Energy-efficient FPGA Accelerator for Binary Convolutional Neural Networks

Yixing Li[1], Zichuan Liu[2], Kai Xu[1], Fengbo Ren[1], Hao Yu[2]
[1]Arizona State University, [2]Nanyang Technological University

CPU-FPGA Co-Optimization for Big Data Applications: A Case Study of In-Memory Samtool Sorting

Jason Cong[1], Zhenman Fang[1], Muhuan Huang[2], Libo Wang[1], Di Wu[1]
[1]UCLA, [2]University of California, Los Angeles

Stochastic-Based Multi-stage Streaming Realization of a Deep Convolutional Neural Network

Mingjie Lin and Mohammed Alawad, University of Central Florida

fpgaConvNet: Automated Mapping of Convolutional Neural Networks on FPGAs

Stylianos Venieris and Christos Bouganis, Imperial College London

Poster Session 3

FPGA-based Hardware Accelerator for Image Reconstruction in Magnetic Resonance Imaging
Emanuele Pezzotti[1], Alex Iacobucci[1], Gregory Nash[2], Umer Cheema[1], Paolo Vinella[1], Rashid Ansari[1]
[1]University of Illinois at Chicago, [2]University of Illinois at Chicago, Altera

ASAP: Accelerated Short Read Alignment on Programmable Hardware
Subho Banerjee, Mohamed El Hadedy, Jong Bin Lim, Daniel Chen, Zbigniew T. Kalbarczyk, Deming Chen, Ravishankar K. Iyer, UIUC

RxRE: Throughput Optimization for High-Level Synthesis using Resource-Aware Regularity Extraction
Atieh Lotfi and Rajesh Gupta, UCSD

GRT 2.0: An FPGA-based SDR Platform for Cognitive Radio Networks
Haoyang Wu[1], Tao Wang[1], Zhiwei Li[1], Boyan Ding[1], Xiaoguang Li[1], Tianfu Jiang[1], Jun Liu[1], Songwu Lu[2]
[1]Peking University, [2]UCLA

FPGA Implementation of Non-Uniform DFT for Accelerating Wireless Channel Simulations
Srinivas Siripurapu[1], Aman Gayasen[2], Nitin Chandrachoodan[1], Padmini Gopalakrishnan[2]
[1]IIT Madras, [2]Xilinx

Learning Convolutional Neural Networks for Data-Flow Graph Mapping on Spatial Programmable Architectures
Shouyi Yin, Dajiang Liu, Lifeng Sun, Xinhan Lin, Leibo Liu, Shaojun Wei
Tsinghua University

Cache Timing Attacks from The SoCFPGA Coherency Port
Sumanta Chaudhuri, Telecom ParisTech

Dynamic Partitioning for Library based Placement on Heterogeneous FPGAs
Fubing Mao[1], Wei Zhang[2], Bingsheng He[3], Siew Kei Lam[1]
[1]Nanyang Technological University, [2]Hong Kong University of Science and Technology, [3]National University of Singapore

An Energy-Efficient Design-Time Scheduler for FPGAs Leveraging Dynamic Frequency Scaling Emulation
Wei Ting Loke[1] and Chin Yang Koay[2]
[1]National University of Singapore, [2]Xilinx Asia Pacific