Abhishek Kumar Jain

Senior Member of Technical Staff, Architecture Group
AMD-Xilinx, San Jose, USA

San Jose, CA, 95124
☎ *925.495.9525*
✉ *jain.abhishekjain123@gmail.com*
http://abhishekkumarjain.github.io/

Education

2012–2016 **Ph.D., Computer Engineering**, *Nanyang Technological University (NTU)*, Singapore.
Architecture Centric Coarse-Grained FPGA Overlays. Advisors: Douglas L. Maskell and Suhaib A. Fahmy
NTU Research Scholarship (2012–2016)
Courses: Advance Image Processing, Advance Computer Architecture, Real Time DSP Design, Advance Topic in Convex Optimization, Computer-Aided VLSI System Design, Principles of Embedded Systems

2008–2012 **B.Tech., Electronics & Communication Engineering (ECE)**, *Indian Institute of Information Technology (IIIT)*, Allahabad, India.
University Silver Medalist: 2^{nd} *Rank in the ECE Department (CGPA: 9.44/10)*
Dissertation work carried out at STMicroelectronics, India
Relevant Courses: Data and File Structures, Image Processing, Computer Architecture, Digital Signal Processing, Embedded Systems, Digital VLSI Design, Microprocessor Programming, Digital Electronics

Research and Work Experience

2022–Present **AMD-Xilinx**, *Senior Member of Technical Staff (Architecture Group)*.
- Accelerators for Sparse Linear Algebra workloads. **[US Patent Apps. 17/679,887, 18/076,602]**
- Design and Development of SpMV Engine on HBM-FPGAs. **[FCCM 2023]**
- Design and Development of Scalable Multi-port Memories. **[US Patents. 11,720,255, 12,079,484]**
- Analysis and mapping of Networking workloads on FPGAs.
- Evaluation of FPGA fabric and AI Engine CGRA for MLPerf Tiny Autoencoder. **[FastML 2024]**

2018–2022 **Xilinx, USA**, *Staff Engineer (Architecture Group)*.
- Design, modelling and evaluation of Domain-specific Architectures. **[FPL 2020]**
- Analysis of workloads in the domain of HPC and Graph Analytics.
- System-level performance benchmarking and architecture evaluation.
- Mapping Sparse MLP Networks on FPGAs (Graph Challenge 2021 Innovation Award). **[HPEC 2021]**
- **Open source contribution: https://github.com/Xilinx/hydra**

2017–2018 **Lawrence Livermore National Laboratory**, *Postdoctoral Research Staff Member*.
- Emulation of memory hierarchy and near-memory accelerators. **[FCCM 2018]**
- Leveraging Zynq Ultrascale+ device for developing LiME-ZU+ (advanced version of LiME).
- Evaluation of near-memory associative indexing using LiME-ZU+ (LLNL LDRD Project).
- Linux support for LiME-ZU+ to allow execution of multithreaded applications running on multi-cores.
- Non-intrusive capture of application memory traces for Exascale Computing Project (ECP).
- **Open source contribution: https://github.com/LLNL/lime**

Fall 2016 **School of Computer Science & Engineering, NTU, Singapore**, *Postdoctoral Research Fellow*.
- **Software-Hardware Communication on Xilinx Zynq**: Developed AXI compatible Scratch-pad Memory Architecture around the FPGA overlay (including Linux Drivers) and integrated within Xilinx Zynq.
- **Just in time OpenCL compilation framework for FPGA overlays**: Developed mechanism for compiling OpenCL kernels at runtime for FPGA overlays on Xilinx Zynq.

2012–2016 **Hardware & Embedded Systems Lab, NTU, Singapore**, *Research Scholar*.
- **FPGA Overlay Architectures**: Research and development of FPGA overlays. Designed and implemented a family of overlays that maximizes application throughput through the use of an array of DSP blocks.
- **Tool-chain development for compiling C description onto Overlay architectures**: Software tool development, specifically for application mapping onto coarse-grained overlay architectures.

Spring 2012 **STMicroelectronics, Greater Noida, India**, *Graduate Intern*.
- Project Title: Firmware library development for ARM Processor based STM32 Microcontroller.
- Developer for STM32 Microcontroller based Image Processing system (Optical Barcode Reader).

Summer 2011 **University of Alberta, Edmonton, Canada**, *Summer Intern, MITACS Globalink Research Award*.
- Design and Simulation of MEMS based variable Capacitor using COMSOL simulation tool.

Skills

System Emulation	Design, modelling, simulation/emulation and evaluation of heterogeneous computing systems.
Architecture	Comprehensive understanding of reconfigurable device architectures (Xilinx FPGAs and SoCs).
Design Tools	Experience with FPGA and SoC design tools (Xilinx ISE, XPS & SDK, Vivado, Vivado HLS).
Programming	Fluent in C, C++/STL, SystemC, RTL Design and Verfication using Verilog HDL.
Scripting	Proficient in Python and Shell scripting for Design Automation.
Compiler	Familiar with OpenCL programming model, LLVM frontend (clang) and optimization passes.

Conference Talks

September 2021	**IEEE High Performance Extreme Computing Conference (HPEC)**, *Sparse Deep Neural Network Acceleration on HBM-Enabled FPGA Platform*, Virtual Conference.
November 2020	**Computer Aided Design (ICCAD)**, *Role of on-chip networks in building domain-specific architectures (DSAs)*, Workshop on System-Level Interconnect Problems, Virtual Conference.
August 2020	**Field-Programmable Logic and Applications (FPL)**, *A Domain-Specific Architecture for Accelerating Sparse Matrix Vector Multiplication on FPGAs*, Virtual Conference.
November 2017	**Supercomputing Conference (SC)**, *LiME: An Open Source Memory System Emulation Platform*, Workshop on Open Source Supercomputing (OpenSuCo), Denver, CO, USA.
November 2017	**Supercomputing Conference (SC)**, *The Role of FPGA Overlay Architectures in Exascale Computing*, Birds-of-a-Feather session on Reconfigurable Computing in Exascale, Denver, CO, USA.
May 2016	**Symposium on Field Programmable Custom Computing Machines (FCCM)**, *DeCO: A DSP Block Based FPGA Accelerator Overlay With Low Overhead Interconnect*, Washington DC, USA.
May 2015	**Symposium on Field Programmable Custom Computing Machines (FCCM)**, *Efficient Overlay Architecture Based on DSP Blocks*, Vancouver, Canada.

Invited Talks

October 2024	**FastML Conference**, *Combining FPGA and AI Engine for Accelerating Embedded AI Applications*.
October 2024	**UT Austin Computer Architecture Seminar Series**, *FPGA Device and Floorplan-aware Accelerator Implementation via Domain-specific Tooling*.
July 2023	**Stanford Agile Hardware Group**, *Domain-specific Architectures using FPGAs: Sparse Linear Algebra Applications*.
September 2021	**IEEE CASS Seasonal School**, *Domain-specific Accelerator Architectures for efficient execution of sparse workloads on FPGA platforms*.
May 2018	**Xilinx**, *Architecture Centric Coarse-Grained FPGA Overlays*.
August 2016	**LLNL**, *Architecture Centric Coarse-Grained FPGA Overlays for High Performance Computing*.
May 2016	**Algo-Logic Systems**, *Virtualizing FPGAs in Heterogeneous Computing Platforms*.
May 2016	**Xilinx**, *DeCO: A DSP Block Based FPGA Accelerator Overlay With Low Overhead Interconnect*.

Patents

August 2023	**US Patent 11,720,255**, *Random reads using multi-port memory and on-chip memory blocks*.
August 2023	**US Patent App. 17/679,887**, *Sparse matrix dense vector multiplication circuitry*.
June 2024	**US Patent App. 18/076,602**, *Compression of sparse matrices for vector processing*.
September 2024	**US Patent 12,079,484**, *Random reads using multi-port memory and on-chip memory blocks*.
October 2024	**US Patent App. 18/134,994**, *3D stacked device having improved data flow*.

Publications

Peer Reviewed Conference & Journal Papers

[1] Abhishek Kumar Jain, Chirag Ravishankar, Hossein Omidian, Sharan Kumar, Maithilee Kulkarni, Aashish Tripathi, and Dinesh Gaitonde. Modular and lean architecture with elasticity for sparse matrix vector multiplication on fpgas. In *FCCM*, 2023.

[2] Abhishek Kumar Jain and Sharan Kumar. Compressing pruned weight matrices for efficient vectorized processing on aie array. In *AMD Global Technical Authors Conference (GTAC)*, 2022.

[3] Abhishek Kumar Jain, Sharan Kumar, Aashish Tripathi, and Dinesh Gaitonde. Sparse deep neural network acceleration on hbm-enabled fpga platform. In *HPEC*, 2021.

[4] Abhishek Kumar Jain, Douglas L Maskell, and Suhaib A Fahmy. Coarse grained fpga overlay for rapid just-in-time accelerator compilation. *IEEE TPDS*, 33(6):1478–1490, 2021.

[5] Stephen Neuendorffer, Alireza Khodamoradi Khodamoradi, Kristof Denolf, Abhishek Kumar Jain, and Samuel Bayliss. The evolution of domain-specific computing for deep learning. *IEEE CAS*, 2021.

[6] Abhishek Kumar Jain, Hossein Omidian, Henri Fraisse, Mansimran Benipal, Lisa Liu, and Dinesh Gaitonde. A domain-specific architecture for accelerating sparse matrix vector multiplication on fpgas. In *FPL*, 2020.

[7] Xiangwei Li, Kizheppatt Vipin, Douglas L Maskell, Suhaib A Fahmy, and Abhishek Kumar Jain. High throughput accelerator interface framework for a linear time-multiplexed fpga overlay. In *ISCAS*, 2020.

[8] Abhishek Kumar Jain, Scott Lloyd, and Maya Gokhale. Microscope on memory: MPSoC-enabled computer memory system assessments. In *FCCM*, 2018.

[9] Abhishek Kumar Jain, Scott Lloyd, and Maya Gokhale. Performance assessment of emerging memories through FPGA emulation. *IEEE Micro*, 39(1):8–16, 2018.

[10] Xiangwei Li, Abhishek Kumar Jain, Douglas L Maskell, and Suhaib A Fahmy. A time-multiplexed FPGA overlay with linear interconnect. In *DATE*, 2018.

[11] Abhishek Kumar Jain, Xiangwei Li, Pranjul Singhai, Douglas L Maskell, and Suhaib A Fahmy. DeCO: a DSP block based FPGA accelerator overlay with low overhead interconnect. In *FCCM*, 2016.

[12] Abhishek Kumar Jain, Douglas L Maskell, and Suhaib A Fahmy. Are coarse-grained overlays ready for general purpose application acceleration on FPGAs? In *IEEE PiCom*, 2016.

[13] Abhishek Kumar Jain, Douglas L Maskell, and Suhaib A Fahmy. Throughput oriented FPGA overlays using DSP blocks. In *DATE*, 2016.

[14] Abhishek Kumar Jain, Xiangwei Li, Suhaib A Fahmy, and Douglas L Maskell. Adapting the DySER architecture with DSP blocks as an overlay for the Xilinx Zynq. *ACM SIGARCH CAN*, September 2016.

[15] Abhishek Kumar Jain, Suhaib A Fahmy, and Douglas L Maskell. Efficient overlay architecture based on DSP blocks. In *FCCM*, 2015.

[16] Abhishek Kumar Jain, Khoa Dang Pham, Jin Cui, Suhaib A Fahmy, and Douglas L Maskell. Virtualized execution and management of hardware tasks on a hybrid ARM-FPGA platform. *JSPS*, October 2014.

[17] Khoa Dang Pham, Abhishek Kumar Jain, Jin Cui, Suhaib A Fahmy, and Douglas L Maskell. Microkernel hypervisor for a hybrid ARM-FPGA platform. In *ASAP*, 2013.

Workshop Papers

[18] Aman Gupta, Sagheer Ahmed, Abhishek Kumar Jain, Ygal Arbel, Abbas Morshed, and David Schultz. Run-time reconfiguration of noc in xilinx acap architecture. In *NoCArc Workshop, at MICRO*, 2020.

[19] Abhishek Kumar Jain. Role of on-chip networks in building domain-specific architectures (dsas) for sparse computations. In *Workshop on System-Level Interconnect*, 2020.

[20] Abhishek Kumar Jain, Scott Lloyd, and Maya Gokhale. LiME: an open source memory system emulation platform. In *Workshop on Open Source Supercomputing (OpenSuCo), at SC*, 2017.

[21] Abhishek Kumar Jain, Douglas L Maskell, and Suhaib A Fahmy. Resource-aware just-in-time OpenCL compiler for coarse-grained FPGA overlays. In *Workshop on Overlays (OLAF), at FPGA*, 2017.

[22] Abhishek Kumar Jain, Douglas L Maskell, and Suhaib A Fahmy. Coarse-grained FPGA overlays for on-demand acceleration of data center workloads. In *H^2RC'16 at SC*, 2016.

[23] Xiangwei Li, Abhishek Jain, Douglas Maskell, and Suhaib A Fahmy. An area-efficient FPGA overlay using DSP block based time-multiplexed functional units. In *Workshop on Overlays (OLAF), at FPGA*, 2016.