Siddhartha

Hardware/Software Engineer

CONTACT INFORMATION

✉ sidmontu@gmail.com 📞 +65 9857-4171

🌐 https://sidmontu.github.io ⓞ github.com/sidmontu

INTERESTS
Hardware Engineering (FPGAs/ASICs), Machine Learning Engineering, HPC, Communications

PROFESSIONAL SUMMARY
Experienced mid-career software engineer with a FPGA-focused research background. Interested in projects at the crossroads of hardware design and modern deep learning applications.

Work Experience

Research Engineer, AMD Research, Singapore — Jan 2022 to Present

- Working on research projects focusing on topics such as hardware design, computer networks, and machine learning.
- Published **demo proposal** at EuroP4'23 on high-speed DNN-based anomaly detector in the network data plane.
- Lead projects and mentor interns throughout the year.

Machine Learning Engineer (Contract), SAP Singapore — Jun 2021 to Dec 2021

- Building enterprise search solutions for internal stakeholders using Elasticsearch/Opensearch stacks.
- Implementing DNN-based semantic-similarity algorithms for information retrieval and ranking of results.
- MLOps tooling: setting up labeling, data cleaning, and evaluation pipelines for efficient data-centric ML implementations.
- Tools/frameworks used (non-exhaustive): [Elastic|Open]search, Docker, FastAPI, Kubernetes, Poetry, Kibana, etc.

CTO/Founder, Inpact Technologies (inPact.ai), Singapore — May 2020 to Jun 2021

- As CTO, planned and executed in-house product development across all aspects of machine learning, frontend, backend, and deployment. Hired freelancers and interns to achieve alpha and **beta** product release milestones within 8 months.
- Frontend using **React-Redux** + **SASS** stack; backend using **serverless framework**; deployment on AWS (EC2, DynamoDB, Cognito, S3, etc); ML models trained using PyTorch and **spaCy**, REST API deployment through **cortex**.
- Trained/fine-tuned DNNs (e.g. residual CNNs, BERT) for NLP tasks like text classification and named entity recognition.
- Raised S$75K from Entrepreneur First, an international VC firm funded by some of the top investors in the world.

Posdoctoral Research Associate, University of Sydney, Australia — Oct 2017 to Dec 2019

- Research & development on a "High-Speed Machine Learning for RF Communications" project commissioned by the Australian Defense Agency, leading to a peer-reviewed publication at **MILCOM 2018**. PI: **Prof. Philip Leong**.
- Contribution to other FPGA-based research projects in the lab on topics such as **on-chip training**, **logic cell architecture design**, and **automatic modulation classification**.
- Extensive hands-on experience with technologies such as Vivado[HLS], RFSoC, PYNQ, Tensorflow, etc during stint at the lab. Undertook sysadmin duties to manage lab resources, and advised students on their final-year research projects.

Education & Certifications

Nanyang Technological University (NTU), Singapore   Jan 2013 to Feb 2019

PhD, Computer Science & Engineering, Advisor: **Prof. Nachiket Kapre**
Dissertation: **Dataflow Optimized Overlays for FPGAs**
- Developed a 16x16 mesh token dataflow overlay architecture finely-tuned for efficient mapping on Arria 10 FPGAs.
- Exploited statically-extracted instruction-criticality information to enable out-of-order execution inside each processing element; improved performance by up to $2.4\times$ over existing in-order processor architectures.
- Verilog for hardware implementation; verilator for testbenches and cycle-accurate simulations; C++ compiler to optimize and map dataflow graphs to instructions inside each PE; python/R/shell scripting for data wrangling tasks.

Teaching Assistant: Programmable System on Chip (CE4054), Embedded Software Development (CE4052)

Imperial College London, United Kingdom   Oct 2009 to Jun 2012

Bachelors of Engineering (BEng), Electrical & Electronics Engineering

Machine Learning Engineering for Production Specialization, DeepLearning.AI, Coursera   Apr 2022

Data Science Specialization, John Hopkins University, Coursera   Apr 2016

Peer-Reviewed Publications

Full Papers (Conferences/Journals)

1. Seyedramin Rasoulinezhad, **Siddhartha**, Hao Zhou, Lingli Wang, David Boland, and Philip Leong "LUXOR: An FPGA Logic Cell Architecture for Efficient Compressor Tree Implementations" *Proceedings of the 2020 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays (FPGA)*, February 2020
[DOI: 10.1145/3373087.3375303]

2. **Siddhartha**, and Nachiket Kapre "DaCO: A High-Performance Token Dataflow Coprocessor Overlay for FPGAs" *International Conference on Field-Programmable Technology*, December 2018
[DOI: 10.1109/FPT.2018.00032]

3. **Siddhartha**, Yee Hui Lee, Duncan Moss, Julian Faraone, Perry Blackmore, Daniel Salmond, David Boland, and Philip Leong "Long Short-Term Memory for Radio Frequency Spectral Prediction and its Real-Time FPGA Implementation" *IEEE Military Communications Conference (MILCOM)*, October 2018
[DOI: 10.1109/MILCOM.2018.8599833]

4. **Siddhartha**, Nachiket Kapre "Hoplite-Q: Priority-Aware Routing in FPGA Overlay NoCs" *IEEE 26th Annual International Symposium on Field-Programmable Custom Computing Machines*, May 2018
[DOI: 10.1109/FCCM.2018.00012]

5. Gopalakrishna Hegde, **Siddhartha**, Nachiket Kapre "CaffePresso: Accelerating Convolutional Networks on Embedded SoCs" *ACM Transactions on Embedded Computing Systems (TECS)*, January 2018
[DOI: 10.1145/3105925]

6. **Siddhartha**, Nachiket Kapre "eBSP: Managing NoC traffic for BSP workloads on the 16-core Adapteva Epiphany-III Processor." *Design, Automation, and Test in Europe*, March 2017
[DOI: 10.23919/DATE.2017.7926961]

7. Gopalakrishna Hegde, **Siddhartha**, Nachiappan Ramasamy, Nachiket Kapre
"CaffePresso: An Optimized Library for Deep Learning on Embedded Accelerator-based platforms." *International Conference on Compilers, Architecture, and Synthesis for Embedded Systems*, October 2016 (Best Paper Award)
[DOI: 10.1145/2968455.2968511]

8. Pradeep Moorthy, **Siddhartha**, and Nachiket Kapre "A Case for Embedded FPGA-based SoCs for Energy-Efficient Acceleration of Graph Problems."
Supercomputing Frontiers 2015, March 2015
[DOI: 10.14529/jsfi150307]

Short Papers/Posters/Workshops

1. **Siddhartha**, Justin Tan, Rajesh Bansal, Huang Chee Cheun, Yuta Tokusashi, Chong Yew Kwan, Haris Javaid, and Mario Baldi "Demo: Enabling DNN Inference in the Network Data Plane" *Proceedings of the 6th European P4 Workshop (co-located with CoNEXT 2023)*, Dec 2023
[DOI: 10.1145/3630047.3630191]

2. Stephen Tridgell, David Boland, Philip Leong, Ryan Kastner, Alireza Khodamoradi, and **Siddhartha** "Real-time Automatic Modulation Classification using RFSoC" *27th Reconfigurable Architectures Workshop (co-located with IPDPS 2020)*, May 2020
[DOI: 10.1109/IPDPSW50202.2020.00021]

3. Stephen Tridgell, David Boland, Philip Leong, and **Siddhartha** "Real-time Automatic Modulation Classification" *International Conference on Field-Programmable Technology*, December 2019 (Poster)
[DOI: 10.1109/ICFPT47387.2019.00052]

4. **Siddhartha**, David Boland, Steve Wilton, Barry Flower, Perry Blackmore, and Philip Leong "Simultaneous Inference and Training using On-FPGA Weight Perturbation Techniques" *International Conference on Field-Programmable Technology*, December 2018 (Poster)
[DOI: 10.1109/FPT.2018.00060]

5. **Siddhartha**, Nachiket Kapre "Out-of-Order Dataflow Scheduling for FPGA Overlays." *Overlay Architectures for FPGAs Workshop (co-located with FPGA 2017)*, February 2017 (Position Paper)
[DOI: arXiv:1705.02734]

6. Sidharth Maheshwari, Gourav Modi, **Siddhartha**, Nachiket Kapre "Vector FPGA Acceleration of 1-D DWT Computations using Sparse Matrix Skeletons." *26th IEEE International Conference on Field-Programmable Logic and Applications*, August 2016 (Poster)
[DOI: 10.1109/FPL.2016.7577361]

7. **Siddhartha**, Nachiket Kapre "Communication Optimization for the 16-core Epiphany Floating-Point Processor Array." *24th IEEE International Symposium on Field-Programmable Custom Computing Machines*, May 2016 (Short Paper)
[DOI: 10.1109/FCCM.2016.15]

8. Gopalakrishna Hegde, **Siddhartha**, Nachiappan Ramasamy, Vamsi Buddha, Nachiket Kapre "Evaluating Embedded FPGA Accelerators for Deep Learning Applications." *24th IEEE International Symposium on Field-Programmable Custom Computing Machines*, May 2016 (Short Paper)
[DOI: 10.1109/FCCM.2016.14]

9. Nachiket Kapre, Han Jianglei, Andrew Bean, Pradeep Moorthy, and **Siddhartha** "GraphMMU: Memory Management Unit for Sparse Graph Accelerators." *22nd Reconfigurable Architectures Workshop (co-located with IPDPS)*, May 2015
[DOI: 10.1109/IPDPSW.2015.101]

10. **Siddhartha**, Nachiket Kapre "FPGA Acceleration of Irregular Iterative Computations using Criticality-Aware Dataflow Optimizations." *International Symposium on Field-Programmable Gate Arrays*, February 2015 (Short Paper)
[DOI: 10.1145/2684746.2689110]

11. **Siddhartha**, Nachiket Kapre "Fanout Decomposition Dataflow Optimizations for FPGA-based Sparse LU Factorization." *International Conference on Field-Programmable Technology*, December 2014 (Short Paper)
[DOI: 10.1109/FPT.2014.7082787]

12. **Siddhartha**, Nachiket Kapre "Heterogeneous Dataflow Architectures for FPGA-based Sparse LU Factorization." *The International Conference on Field Programmable Logic and Applications*, September 2014 (Short Paper)
[DOI: 10.1109/FPL.2014.6927401]

13. Nachiket Kapre, **Siddhartha** "Limits of Statically-Scheduled Token Dataflow Processing." *International workshop on Data-Flow Models (DFM) for Extreme Scale Computing (co-located with PACT 2014)*, August 2014
[DOI: 10.1109/DFM.2014.21]

14. **Siddhartha**, Nachiket Kapre "Breaking Sequential Dependencies in FPGA-based Sparse LU Factorization." *International Symposium on Field Programmable Custom Computing Machines*, May 2014 (Short Paper)
[DOI: 10.1109/FCCM.2014.26]