UNIVERSITY OF SOUTHERN CALIFORNIA

COMPUTER ENGINEERING

SCREENING EXAMINATION

EE 677

Accelerated Computing using Field Programmable Gate Arrays

SUGGESTED READING

1. Data Parallel C++: Mastering DPC++ for Programming of Heterogeneous Systems using C++ and SYCL by James Reinders, Ben Ashbaugh, James Brodman, Michael Kinsner, John Pennycook, Xinmin Tian, 2020.
2. Reconfigurable Computing: Accelerating Computation with Field-Programmable Gate Arrays, Maya B. Gokhale and Paul S. Graham, Springer, 2005. (Chapters 1-6)
3. Trimberger, S. M. S. (2018). Three ages of fpgas: a retrospective on the first thirty years of fpga technology: this paper reflects on how Moore's law has driven the design of fpgas through three epochs: the age of invention, the age of expansion, and the age of accumulation. IEEE Solid-State Circuits Magazine, 10(2), 16-29.
4. Cong, J., Lau, J., Liu, G., Neuendorffer, S., Pan, P., Vissers, K., & Zhang, Z. (2022). FPGA HLS today: Successes, challenges, and opportunities. ACM Transactions on Reconfigurable Technology and Systems (TRETS), 15(4), 1-42.
5. Korolija, D., Roscoe, T., & Alonso, G. (2020). Do OS abstractions make sense on FPGAs. In 14th USENIX Symposium on Operating Systems Design and Implementation (OSDI 20) (pp. 991-1010).
6. Bobda, C., Mbongue, J. M., Chow, P., Ewais, M., Tarafdar, N., Vega, J. C., ... & Tessier, R. (2022). The future of FPGA acceleration in datacenters and the cloud. ACM Transactions on Reconfigurable Technology and Systems (TRETS), 15(3), 1-42.
7. Wu, Y. W., Wang, Q. G., Zheng, L., Liao, X. F., Jin, H., Jiang, W. B., ... & Hu, K. (2021). FDGLib: a communication library for efficient large-scale graph processing in FPGA-accelerated data centers. Journal of Computer Science and Technology, 36, 1051-1070.
8. Zhang, B., & Prasanna, V. (2023). Dynasparse: Accelerating GNN Inference through Dynamic Sparsity Exploitation. IPDPS 2023.
9. Wijeratne, S., Wang, T. Y., Kannan, R., & Prasanna, V. (2023, February). Accelerating Sparse MTTKRP for Tensor Decomposition on FPGA. In

Proceedings of the 2023 ACM/SIGDA International Symposium on Field Programmable Gate Arrays (pp. 259-269).
10. Meng, Y., Kannan, R., & Prasanna, V. (2023, February). A Framework for Monte-Carlo Tree Search on CPU-FPGA Heterogeneous Platform via on-chip Dynamic Tree Management. In Proceedings of the 2023 ACM/SIGDA International Symposium on Field Programmable Gate Arrays (pp. 235-245).
11. Lin, Y. C., & Prasanna, V. (2023). HyScale-GNN: A Scalable Hybrid GNN Training System on Single-Node Heterogeneous Architecture. IPDPS 2023. arXiv preprint arXiv:2303.00158.
12. Zhang, P; Kannan, R; Prasanna, V., Phases, Modalities, Temporal and Spatial Locality: Domain Specific ML Prefetcher for Accelerating Graph Analytics, The International Conference for High Performance Computing, Networking, Storage, and Analysis (SC '23), 2023
13. de Fine Licht, J., Kwasniewski, G., & Hoefler, T. (2020, February). Flexible communication avoiding matrix multiplication on FPGA with high-level synthesis. In Proceedings of the 2020 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays (pp. 244-254).
14. Besta, M., Gerstenberger, R., Fischer, M., Podstawski, M., Müller, J., Blach, N., ... & Hoefler, T. (2023). High-Performance Graph Databases That Are Portable, Programmable, and Scale to Hundreds of Thousands of Cores. arXiv preprint arXiv:2305.11162.
15. Srivastava, N., Jin, H., Smith, S., Rong, H., Albonesi, D., & Zhang, Z. (2020s). Tensaurus: A versatile accelerator for mixed sparse-dense tensor computations. In 2020 IEEE International Symposium on High Performance Computer Architecture (HPCA) (pp. 689-702). IEEE.
16. Zhang, C., Meng, Y., & Prasanna, V. (2023). A Framework for Mapping DRL Algorithms With Prioritized Replay Buffer Onto Heterogeneous Platforms. IEEE Transactions on Parallel and Distributed Systems.
17. Zhang, B., Zeng, H., & Prasanna, V. (2023). GraphAGILE: An FPGA-based Overlay Accelerator for Low-latency GNN Inference. IEEE Transactions on Parallel and Distributed Systems.
18. Brian, G., Dinesh, G., Chirag, R., & Trevor, B. (2019). Xilinx adaptive compute acceleration platform: Versal™ architecture. In Proceedings of the ACM/SIGDA International Symposium on Field-programmable Gate Arrays (pp. 84-93).
19. Asiatici, M., & Ienne, P. (2021, June). Large-scale graph processing on FPGAs with caches for thousands of simultaneous misses. In 2021 ACM/IEEE 48th Annual International Symposium on Computer Architecture (ISCA) (pp. 609-622). IEEE.
20. Intel FPGA. (2019). AgilexTM FPGAs Deliver a Game-Changing Combination of Flexibility and Agility for the Data-Centric World. Intel White Paper.
21. Xilinx FPGA. (2021). AI Engines and Their Applications. Xilinx White Paper.
22. Chen, P., Manjunath, P., Wijeratne S., Zhang, B., Prasanna, V., (2023) Exploiting On-chip Heterogeneity of Versal Architecture for GNN Inference Acceleration. 2023 33rd International Conference on Field-Programmable Logic and Applications (FPL).

23. Yang, Y., Kuppannagari, S. R., Kannan, R., & Prasanna, V. K. (2022, December). Bandwidth Efficient Homomorphic Encrypted Matrix Vector Multiplication Accelerator on FPGA. In 2022 International Conference on Field-Programmable Technology (ICFPT) (pp. 1-9). IEEE.

****

Please be aware that these references are for guidance in BASIC knowledge. Ph.D. candidates are screened on the basis of their talent, course knowledge, independent reading and experience.

****

Updated August 2023