R244: Paper Review Presentation Assignment

2022/10/19 Session 2: Data flow programming

tl559 (Theo) 6.2. D. Murray, F. McSherry, R. Isaacs, M. Isard, et al.: Naiad: A Timely Dataflow System, SOSP, 2013.
You can also read 6.1 and 6.3 if you can.

aib36 (Antonia) 8. M. Abadi et al. Tensorflow: A system for large-scale machine learning. OSDI, 2016.
You can also read: 8.2. M. Abadi, M. Isard and D. Murray: A Computational Model for TensorFlow - An Introduction, MAPL, 2017.

kc642 (Kian) 11. R. Nishihara, P. Moritz, et al.: Ray: A Distributed Framework for Emerging AI Applications, OSDI, 2018.

mbg34 (Martin) 12. M. Schaarschmidt, S. Mika, et al.: RLgraph: Flexible Computation Graphs for Deep Reinforcement Learning, SysML, 2019.

gob24 (George) 16 . P. Barham, et al.: Pathways: Asynchronous Distributed Dataflow for ML, MLSys, 2022.

2022/10/26 Session 3: Large-scale graph data processing

****tr500 (Teodora) 1.** G. Malewicz, M. Austern, A. Bik, J. Dehnert, I. Horn, N. Leiser, and G. Czajkowski: Pregel: A System for Large-Scale Graph Processing, SIGMOD, 2010.

zl525 (Zejian) 4. J. Gonzalez, Y. Low, H. Gu, D. Bickson, and C. Guestrin: Powergraph: distributed graph-parallel computation on natural graphs. OSDI, 2012.

smxz2 (Sarah) 5. J. Shun and G. Blelloch: Ligra: A Lightweight Graph Processing Framework for Shared Memory, PPoPP, 2013.

sa2140 (Sharan) 9. A. Roy, I. Mihailovic, W. Zwaenepoel: X-Stream: Edge-Centric Graph Processing using Streaming Partitions, SOSP, 2013.

rm2084 (Ridwan) 19. Z. Jia, Y. Kwon, G. Shipman, P. McCormick, M. Erez, A. Aiken: A Distributed Multi-GPU System for Fast Graph Processing, VLDB, 2018.

at2008 (Anna) 14. S. Hong, H. Chafi, E. Sedlar, f K.Olukotun: Green-Marl: A DSL for Easy and Efficient Graph Analysis, ASPLOS, 2012.

2022/11/09 Session 5: Probabilistic Programming

sl955 (Shuntian) 14 . J. Shao, et al.: Tensor Program Optimization with Probabilistic Programs, ArXiv, 2022.

jh2422 (Jocob) 7. V. Dalibard, M. Schaarschmidt, and E. Yoneki: BOAT: Building Auto-Tuners with Structured Bayesian Optimization, WWW, 2017.

****rm2084 (Ridwan) 15 .** X. Wan, et al.: Bayesian Generational Population-Based Training, ICLR, 2022.

2022/11/16 Session 6: Optimisation in Computer Systems I

tr500 (Teodora) 6. O. Alipourfard et al.: CherryPick: Adaptively Unearthing the Best Cloud Configurations for Big Data Analytics, NSDI, 2017.

kc642 (Kian) 16. T. Chen, T. Moreau, Z. Jiang, L. Zheng, S. Jiao, E. Yan, H. Shen, M. Cowan, L. Wang, Y. Hu, L. Ceze, C. Guestrin, and A. Krishnamurthy: TVM: An Automated End-to-End Optimizing Compiler for Deep Learning, OSDI, 2018. …plus if you can skim the following paper (3 pages poster), that will be good.
17. T. Chen, T. Moreau, Z. Jiang, L. Zheng, S. Jiao, E. Yan, H. Shen, M. Cowan, L. Wang, Y. Hu, L. Ceze, C. Guestrin, and A. Krishnamurthy: TVM: End-to-End Compilation Stack for Deep Learning, SysML, 2017.

mbg34 (Martin) 29. Z. Jia, O. Padon, J. Thomas, T. Warszawski, M. Zaharia, A. Aiken: TASO: Optimizing Deep Learning Computation with Automated Generation of Graph Substitutions: SOSP, 2019.

****zl525 (Zejian) 32.** KH. Wang, J. Zhai, M. Gao, Z. Ma, S. Tang, L. Zheng, Y. Li, K. Rong, Y. Chen, and Z. Jia: PET: Optimizing Tensor Programs with Partially Equivalent Transformations and Automated Corrections, ODSI, 2021.

smxz2 (Sarah) 35. Y. Yang, et al.: Equality Saturation for Tensor Graph Superoptimization, MLSys, 2021.

tl559 (Theo) 37. Z. Jia, et al.: Unity: Accelerating DNN Training Through Joint Optimization of Algebraic Transformations and Parallelization, OSDI, 2022.

2022/11/23 Session 7: Optimisation in Computer Systems II

sa2140 (Sharan) 1. A. Mirhoseini et al.: Device Placement Optimization with Reinforcement Learning**, ICML, 2017.**

…plus **1.2.** A. Mirhoseini, A. Goldie et al.: A Hierarchical Mode for Device Placement, ICLR, 2018.

****at2008 (Anna) 19.** D. Ha and J. Schmidhuber: World Models, arXiv, 2018 (https://worldmodels.github.io).

sl955 (Shuntian) 28. A. Paliwal et al.: REGAL: Transfer Learning For Fast Optimization of Computation Graphs, arxiv, 2019.

jh2422 (Jocob) 12. D. Aken , D. Yang, S. Brillard, A. Fiorino, B. Zhang, C. Bilien, and A. Pavlo: An Inquiry into Machine Learning-based Automatic Configuration Tuning Services on Real-World Database Management Systems, VLDB, 2021.

The original OtterTune paper is following:

5. D Van Aken, A Pavlo, GJ Gordon, and B Zhang: Automatic database management system tuning through large-scale machine learning, SIGMOD, 2017.

aib36 (Antonia) 9. G. Li, X. Zhou, S. Li, and B. Gao: Qtune: RL for DB query optimisation, VLDB, 2019.

gob24 (George) 20. R. Marcus, P. Negi, H. Mao, N. Tatbul, M. Alizadeh, and T. Kraska: Bao: Learning to Steer Query Optimizers, VLDB, 2020.