Hanyu Zhao

5-year Ph.D. Candidate, Peking University

158-1003-6083 (Tel./WeChat)
zhaohanyu1994@gmail.com (Email)
https://zhypku.github.io/ (Homepage)

Research Interests

Distributed systems, systems for machine learning, cloud computing, with a recent focus on *cluster-scale* machine learning systems, in particular resource scheduling and data storage.

Education

- **Peking University**   Ph.D. Candidate   2016.9 – 2021.7 (Expected)
  Computer Architecture, School of EECS, *Supervisor: Prof. Yafei Dai*

- **Wuhan University**   B.Eng.   2012.9 – 2016.6
  Simulation Science and Technology, Computer School

Internship

- **Microsoft Research Asia**   Full-time Research Intern   2017.11 – 2020.7 (33 months)
  Systems Research Group, *Mentor: Dr. Quanlu Zhang*

Projects (Selected)

- **HiveD: Multi-Tenant GPU Cluster Scheduling**   *OSDI '20 (first-author)*   2018.6 – 2020.7

  HiveD is the first scheduling framework that explicitly reserves GPU *affinity* for tenants, to eliminate unexpected anomalies in traditional systems that only reserve GPU *number* (quota), thereby providing guaranteed resource sharing experiences. https://github.com/microsoft/hivedscheduler

  – Analysis of Microsoft clusters showing the anomaly of quota-based systems: worse performance (affinity) in the shared cluster than in a private cluster, damaging the incentive of sharing resources

  – New resource abstraction Cell for defining the affinity share of a tenant, and Buddy Cell Allocation algorithm for dynamic allocation with guaranteed access to everyone's own affinity

  – Job priorities, compatibility with state-of-the-art scheduling policies, bad hardware tolerance

  – Open-source K8s-based implementation, and real deployment at Microsoft, managing 1000+ GPUs

- **Training Data Scheduling for Deep Learning**   *Paper in prep. (first-author)*   2019.12 – 2020.7

  Decoupled compute and storage is a common practice in cloud computing. We treat the provisioning of training data from remote storage as a new dimension of cluster scheduling, by exploiting the unique characteristics of DL workloads, to improve job performance and cluster utilization.

  – New caching policy Uniform Caching leveraging DL's uniform data access to minimize cache miss

  – Resource-performance model derived from Uniform Caching and DL's execution pattern, to estimate job performance given available cache space and remote-local bandwidth

  – Dynamic cache-bandwidth joint allocation to jobs in a cluster, and data-job co-scheduling, by exploiting the performance model, to optimize performance and utilization

- **SDPaxos: Semi-Decentralized State Machine Replication**   *SoCC '18 (first-author)*
  2016.10 – 2017.5

  SDPaxos is a new replication protocol that adopts a hybrid architecture which decentralizes operation replication, and centralizes operation ordering, to overcome the limitations and inherit the advantages of existing purely centralized or decentralized protocols. https://github.com/zhypku/SDPaxos

- Semi-decentralized replication scheme, which separates replication and ordering into two Paxos phases
- Techniques to coordinate the two Paxos phases to achieve the optimal 1-round-trip latency under realistic configurations, with proven consistency and linearizability
- Up to $6\times$ throughput of centralized protocols, and $1.7\times$ throughput of decentralized protocols

Publications

[1] **HiveD: Sharing a GPU Cluster for Deep Learning with Guarantees**
  Hanyu Zhao, Zhenhua Han, Zhi Yang, Quanlu Zhang, Fan Yang, Lidong Zhou, Mao Yang, Francis C.M. Lau, Yuqi Wang, Yifan Xiong, Bin Wang
  14th USENIX Symposium on Operating Systems Design and Implementation (**OSDI '20**)

[2] **SDPaxos: Building Efficient Semi-Decentralized Geo-replicated State Machines**
  Hanyu Zhao, Quanlu Zhang, Zhi Yang, Ming Wu, Yafei Dai
  ACM Symposium on Cloud Computing 2018 (**SoCC '18**)

[3] **Don't Miss Any Piece of Knowledge: In-Network Mutual Learning with Sketch Side Branches**
  Yunteng Luan, **Hanyu Zhao**, Zhi Yang, Yafei Dai
  arXiv preprint (1911.09418)

[4] **ScheD2: Scheduling Deep Learning Training via Deep Reinforcement Learning**
  Yunteng Luan, Xukun Chen, **Hanyu Zhao**, Zhi Yang, Yafei Dai
  IEEE Global Communications Conference 2019 (**GlobeCom '19**)

[5] **Gandiva: Introspective Cluster Scheduling for Deep Learning**
  Wencong Xiao, Romil Bhardwaj, Ramachandran Ramjee, Muthian Sivathanu, Nipun Kwatra, Zhenhua Han, Pratyush Patel, Xuan Peng, **Hanyu Zhao**, Quanlu Zhang, Fan Yang, Lidong Zhou
  13th USENIX Symposium on Operating Systems Design and Implementation (**OSDI '18**)

[6] **Scheduling CPU for GPU-based Deep Learning Jobs** (Poster)
  Wencong Xiao, Zhenhua Han, **Hanyu Zhao**, Xuan Peng, Quanlu Zhang, Fan Yang
  ACM Symposium on Cloud Computing 2018 (**SoCC '18**)

[7] **Building Efficient and Available Distributed Transaction with Paxos-based Coding Consensus**
  Shenglong Li, Quanlu Zhang, Zhi Yang, **Hanyu Zhao**, Yafei Dai
  IEEE INFOCOM WKSHPS DCPerf 2018

Awards & Activities (Selected)

- **Award for Scientific Research, Peking University** 2019.12
- **Tianwang-Miaozhen Scholarship, Peking University** 2018.12
- **SoCC '18 Student Scholarship** 2018.10
- **Outstanding Graduate, Wuhan University** 2016.6
- **Merit Student, Wuhan University** 2014.11, 2013.11
- **Member of the Debate Team of Wuhan University** $2013.11 - 2016.6$
- **Captain of the Debate Team of Computer School, Wuhan University** $2013.11 - 2014.11$

Professional Skills

- **Languages**: C, Go, Python, C++, Java, LaTeX, Shell, Markdown
- **Systems**: Linux, TensorFlow, Kubernetes, Docker, PyTorch, Hadoop
- **Knowledge**: Distributed Systems, Machine Learning, Scheduling, Consensus Protocols
- **Skills**: Git, GitHub Open-source Collaboration, English Writing