

# Jahanzeb Maqbool Hashmi

+1 (614) 867 4532 | [jahanzeb.maqbool@gmail.com](mailto:jahanzeb.maqbool@gmail.com)

<https://hashmij.github.io>

## OBJECTIVE

---

I am interested to work with collaborative research and development teams and design large-scale parallel and distributed systems to scale state-of-the-art computational and deep learning workloads on HPC and Cloud systems.

## EDUCATION

---

Ph.D. in Computer Science and Engineering

[The Ohio State University](#), Columbus, Ohio, USA

2015 - 2020

**Thesis:** Designing High Performance Shared-Address-Space and Adaptive Communication Middlewares for Next-Generation HPC Systems

**Advisor:** [Dhabaleswar K. \(DK\) Panda](#)

M.S. in Computer Engineering

[Ajou University](#), Suwon, South Korea

2012 - 2014

**Thesis:** Exploring Performance and Energy Efficiency of ARM Multicore Cluster for High Performance Scientific Computing

**Advisor:** [Sangyoon Oh](#)

B.S. Information Technology

[National University of Science and Technology](#), Islamabad, Pakistan

2007 - 2011

**Thesis:** Implementation and Evaluation of Scientific Simulations on HPC Architectures

**Advisor:** [Aamir Shafi](#)

## PROFESSIONAL EXPERIENCE

---

I design novel algorithms and develop system software that help distributed communication middlewares like MPI achieve high-performance on highly parallel architectures. My work is released to the community as part of the MVAPICH2 (<http://mvapich.cse.ohio-state.edu/>) library and empowers world's top supercomputers (e.g., TACC Frontera, SDSC Expanse) and HPC Clouds (e.g., [Microsoft Azure HB](#), Amazon AWS). My research is also adopted by the leading high-performance communication stacks (e.g., MPICH, UCX).

- **Senior Research Associate**, [The Ohio State University](#), USA 2020 - Present
  - Working on the design and development of high-performance MPI library for next-generation HPC and Cloud systems with many-core CPUs, GPUs (NVIDIA CUDA and AMD ROCm), and accelerators.
  - Working on semantic and performance optimizations of MPI runtimes for scaling distributed deep learning on native and cloud HPC systems.
  - Working on design and development of generalized hierarchical MPI collective communications framework for modern CPU and GPU systems.
- **Graduate Research Associate**, [Network Based Computing Laboratory](#), OSU, USA 2015 - 2020
  - Designed an adaptive and topology aware algorithm for mapping of MPI processes to hardware cores by capturing the communication-patterns of AI and HPC applications. The proposed designs particularly helped mitigate the performance penalties incurred by arbitrary vNUMA to pNUMA mappings of Microsoft Azure HBv1 hypervisor by transparently generating optimal placements of MPI processes while considering hardware and applications' characteristics – (IPDPS'20)
  - Worked collaboratively on efficient parallelization of large-scale distributed DNN training (data and model parallel) on CPU and GPU systems – (SC'20, PPoPP'17)
  - Designed and developed a truly zero-copy based inter-process (IPC) communication called shared address space communication backend for manycore architectures. We designed a truly zero-copy Allreduce algorithm that helped accelerate the CPU based distributed DL training and brought MVAPICH2 at the forefront of the first-choice MPI library for scientific and DL applications – (IPDPS'18, Cluster'18, CCGRID'19)
  - Designed a novel algorithm to cache data layouts to help mitigate the performance costs of layout translation of MPI derived datatypes. I further designed a zero-copy communication framework for non-contiguous datatypes focusing sparse data layouts on CPU and GPU resident data – (IPDPS'19 Best Paper Finalist, JPDC, HiPC'19)

- Worked on PGAS libraries e.g., OpenSHMEM, UPC++ and task-based programming models e.g., Kokkos with MPI backend – (HPCC’16, SC’20)

## TECHNICAL SKILLS

---

- Parallel Programming Systems – MPI, OpenMP, CUDA, HIP, OpenSHMEM, UPC++
- Deep Learning – Tensorflow, PyTorch, Horovod
- Programming Languages – C, C++, Java
- Scripting Languages – Bash, Python
- Tools – Git, LaTeX, GDB, PerfAPI (PAPI), mpiP, Valgrind
- Familiarity with developing Linux Kernel modules
- Strong programming, debugging, and problem solving skills.
- Experienced with large-scale software design, development, and release life-cycle.
- Experienced with performance benchmarking of parallel hardware with focus on scientific and AI applications
- Strong communication and presentation skills.

## SELECT PUBLICATIONS

---

For complete list of publications, please refer to my [Google Scholar](#).

1. A. Jain, A. Awan, A. Aljuhani, **J. Hashmi**, Q. Anthony, H. Subramoni, D. Panda, R. Machiraju, A. Parwani. “GEMS: GPU Enabled Memory Aware Model Parallelism System for Distributed DNN Training“, in proceeding of *IEEE/ACM International Conference for High Performance Computing, Networking, Storage, and Analysis (SC ’20)* , Nov 2020
2. [HiPC ’20] A. Shafi, **J. Hashmi**, H. Subramoni, and D. K. Panda. “Blink: Towards Efficient RDMA-based Communication Coroutines for Parallel Python Applications“, in proceeding of *27th IEEE International Conference on High Performance Computing, Data, Analytics and Data Science*, Dec. 2020.
3. **J. Hashmi**, C. Chu, S. Chakraborty, M. Bayatpour, H. Subramoni, and DK Panda. “FALCON-X: Zero-copy MPI Derived Datatype Processing on Modern CPU and GPU Architectures“, *Journal of Parallel and Distributed Computing (JPDC)*, Volume 144, October 2020, Pages 1-13, doi.org/10.1016/j.jpdc.2020.05.008
4. **J. Hashmi**, S. Xu, B. Ramesh, M. Bayatpour, H. Subramoni, and D. K. Panda. “Machine-agnostic and Communication-aware Designs for MPI on Emerging Architectures“, in proceeding of *34th IEEE International Parallel and Distributed Processing Symposium (IPDPS ’20)* , May 2020
5. **J. Hashmi**, S. Chakraborty, M. Bayatpour, H. Subramoni, D. K. Panda. “FALCON: Efficient Designs for Zero-copy MPI Datatype Processing on Emerging Architectures“, in proceeding of *33rd IEEE International Parallel and Distributed Processing Symposium (IPDPS ’19)* , May 2019 [\[Best Paper Finalist\]](#)
6. **J. Hashmi**, S. Chakraborty, M. Bayatpour, H. Subramoni, D. K. Panda. “Designing Efficient Shared Address Space Reduction Collectives for Multi-/Many-cores“, in proceeding of *32nd IEEE International Parallel and Distributed Processing Symposium (IPDPS ’18)* , May 2018
7. S. Chakraborty, M. Bayatpour, **J. Hashmi**, H. Subramoni, D. K. Panda. “Cooperative Rendezvous Protocols for Improved Performance and Overlap“, in proceeding of *IEEE/ACM International Conference for High Performance Computing, Networking, Storage, and Analysis (SC ’18)* , Nov 2018 [\[Best Paper Finalist\]](#)
8. M. Bayatpour, **J. Hashmi**, S. Chakraborty, H. Subramoni, P. Kousha, D. K. Panda. “SALaR: Scalable and Adaptive Designs for Large Message Reduction Collectives“, in proceeding of *IEEE International Conference on Cluster Computing (CLUSTER 2018)* , Sep 2018 [\[Best Paper Award\]](#)
9. A. Awan, K. Hamidouche, **J. Hashmi**, and D. K. Panda. “S-Caffe: Co-designing MPI Runtimes and Caffe for Scalable Deep Learning on Modern GPU Clusters“, in proceeding of *22nd ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming (PPoPP ’17)* , February 2017