

Sourav Chakraborty

8 years of experience in building high-performance and scalable software for large HPC clusters with cutting-edge CPUs and GPUs

Austin, TX, USA
+1 (614) 500-3409
mail@souravc.com
<https://souravc.com>

EXPERIENCE

NVIDIA, Staff HPC Middleware Developer (2021 - Present)

Lead development of offloaded collective framework for NVIDIA BlueField DPUs to accelerate HPC applications and AI frameworks (PyTorch)

Mentor and guide a team of 3 developers to plan and execute development roadmap for NVIDIA SmartNICs software

Design collective algorithms to maximize performance and overlap of communication and computation with limited system resources

Collaborate with inside sales teams and external customers to identify, prioritize, and solve customer requirements

AMD, Senior Software Engineer (2019-2021)

Lead UCX and MPI development efforts for AMD GPUs

Improve performance of point-to-point (2x) and collective operations (8x) on AMD GPUs on RDMA capable networks (InfiniBand and RoCE)

Improve functionality and performance QA process for UCX and MPI

Identify and contribute support for AMD GPUs to strategically relevant open source software ecosystems (RDMA, MPI Benchmarks, etc.)

Helped secure \$1M+ deal by analyzing and solving customer need

The Ohio State University, HPC Researcher (2014-2019)

Develop scalable and adaptive communication runtimes for large scale HPC systems with thousands of nodes

Design intra-node and inter-node communication protocols and collective algorithms for MPI and OpenSHMEM

Develop optimized MPI libraries for cloud environments (AWS and Azure)

Develop scalable fault-tolerant MPI library with SLURM

Yahoo!, Software Development Engineer (2011-2014)

Develop web-service APIs for serving ~1Billion user profiles to high-traffic pages including Yahoo! Frontpage

Build APIs to import ~500M FB profiles to Yahoo! NoSQL database

EDUCATION

The Ohio State University, M.S., Ph.D., Computer Science - 2019

Jadavpur University, B.E., Information Technology - 2011

TECHNOLOGIES

RDMA, MPI, UCX, LibFabric, GPUDirect, InfiniBand, RoCE, Omni-Path, SHMEM, SHARP, RCCL, NCCL, SLURM

PUBLICATIONS

25+ peer-reviewed papers in journals and conferences (see website for details)

Patent: Peer-to-peer architecture for web traffic management (US2014,0012906A1)

AWARDS

Best Graduate Student Research Award, OSU-CSE 2018

ACM Student Research Award at SuperComputing 2016

Several best paper/poster awards and nominations at SC, IPDPS, ISC, and Cluster

COMMUNITY

Represented OSU and AMD in MPI Forum standardization efforts

Presented in MVAPICH, UCX, and SLURM user group meetings

Presented tutorials on InfiniBand and MPI at SC '18 and other HPC conferences