

Jaemin Choi

PhD in Computer Science

jaemin@acm.org

Updated August 18, 2022

TECHNICAL INTERESTS	Distributed Deep Learning, High Performance Computing, GPGPU, Performance Modeling, Asynchronous Task-based Runtime Systems
EDUCATION	<p><i>Doctor of Philosophy (PhD)</i>, Computer Science University of Illinois Urbana-Champaign - Urbana, Illinois, USA Advisor: Prof. Laxmikant V. Kale Aug 2016 - Aug 2022</p> <p><i>Bachelor of Science (BS)</i>, Computer Science and Engineering Seoul National University - Seoul, Republic of Korea Advisor: Prof. Jaejin Lee Mar 2010 - Feb 2016</p>
TECHNICAL SKILLS	<p><i>Programming Languages</i>: C, C++, Python <i>Parallel/Distributed Programming</i>: MPI, Charm++, CUDA, NVSHMEM, SYCL/DPC++, PyTorch <i>Containers</i>: Docker, Singularity, HPC Container Maker <i>Tools</i>: Git, GDB, NVProf, Nsight Systems, gprof <i>Job Schedulers</i>: PBS, Slurm, IBM Spectrum LSF <i>HPC systems</i>: OLCF Summit, ALCF Theta, LLNL Lassen, PSC Bridges-2, SDSC Expanse</p>
EXPERIENCES	<p><i>Senior Deep Learning Architect</i> Aug 2022 - Present NVIDIA Corporation - Santa Clara, CA</p> <ul style="list-style-type: none">Understand, analyze, profile, and optimize deep learning training workloads on state-of-the-art hardware and software platforms.Guide development of future generations of deep learning processors and computing platforms. <p><i>Research Assistant</i> Aug 2016 - Aug 2022 Parallel Programming Laboratory, University of Illinois Urbana-Champaign</p> <ul style="list-style-type: none">Optimized performance for GPU-accelerated applications on modern heterogeneous HPC platforms by developing new features in the Charm++ parallel programming system, including asynchronous completion notification and GPU-aware communication.Developed CharminG, a GPU-resident parallel programming framework built on CUDA and NVSHMEM, with the goal of performing task scheduling and communication inside the GPU devices.Improved support for NVIDIA and Intel GPUs in the NAMD molecular dynamics simulation framework. <p><i>Graduate Technical Intern (Mentor: Gengbin Zheng, Manager: Craig Belusar)</i> May - Aug 2021 Intel Corporation - Austin, TX (Virtual)</p> <ul style="list-style-type: none">Developed support for Intel GPUs in OpenMPI using Intel oneAPI Level Zero and Libfabric/OFI.Validated point-to-point and collective MPI calls on Intel GPU clusters with OSU micro-benchmark suite. <p><i>Research Intern (Mentor: Prof. Abhinav Bhatele)</i> May - Aug 2019 Lawrence Livermore National Laboratory - Livermore, CA</p>

- Created performance models using parallel discrete event simulation (PDES) and roofline model to analyze and predict the performance of GPU-accelerated proxy applications in the Exascale Computing Project (ECP), including SW4lite and MiniFE.

Technology Research Intern (Mentor: [Rasmus Tamstorf](#))

May - Aug 2018

Walt Disney Animation Studios - Burbank, CA

- Optimized memory usage in a parallel path tracing renderer via de-duplication of scene objects.

PUBLICATIONS

David J. Hardy, **Jaemin Choi**, Wei Jiang, Emad Tajkhorshid. 2022. Experiences Porting NAMD to the Data Parallel C++ Programming Model. *10th International Workshop on OpenCL and SYCL (IWOCCL'22)*.

Jaemin Choi, David F. Richards, Laxmikant V. Kale. 2022. Improving Scalability with GPU-Aware Asynchronous Tasks. *The 27th International Workshop on High-Level Parallel Programming Models and Supportive Environments (HIPS'22), in conjunction with IPDPS'22*.

Joseph Hutter, Justin Szaday, **Jaemin Choi**, Spencer Wallace, Simeng Liu, Laxmikant V. Kale, Thomas Quinn. 2022. ParaTreeT: A Fast, General Framework for Spatial Tree Traversal. *36th IEEE International Parallel and Distributed Processing Symposium (IPDPS'22)*.

Zane Fink, Simeng Liu, **Jaemin Choi**, Matthias Diener, Laxmikant V. Kale. 2021. Performance Evaluation of Python Parallel Programming Models: Charm4Py and mpi4py. *IEEE/ACM 6th International Workshop on Extreme Scale Programming Models and Middleware (ESPM2'21), in conjunction with SC'21*.

Jaemin Choi, Zane Fink, Sam White, Nitin Bhat, David F. Richards, Laxmikant V. Kale. 2021. Accelerating Communication for Parallel Programming Models on GPU Systems. *Special Issue on Topics on Heterogeneous Computing of the Elsevier International Journal on Parallel Computing (PARCO)*.

Jaemin Choi, Zane Fink, Sam White, Nitin Bhat, David F. Richards, Laxmikant V. Kale. 2021. GPU-aware Communication with UCX in Parallel Programming Models: Charm++, MPI, and Python. *Eleventh International Workshop on Accelerators and Hybrid Emerging Systems (AsHES'21), in conjunction with IPDPS'21*.

Jaemin Choi, David F. Richards, Laxmikant V. Kale. 2020. Achieving Computation-Communication Overlap with Overdecomposition on GPU Systems. *IEEE/ACM 5th International Workshop on Extreme Scale Programming Models and Middleware (ESPM2'20), in conjunction with SC'20*.

Jaemin Choi, David F. Richards, Laxmikant V. Kale, Abhinav Bhatele. 2020. End-to-end Performance Modeling of Distributed GPU Applications. *International Conference on Supercomputing (ICS'20)*.

RESEARCH POSTERS

Jaemin Choi, David F. Richards, Laxmikant V. Kale. 2021. CharminG: A Scalable GPU-resident Runtime System. *ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC'21)*.

Jaemin Choi, David F. Richards, Abhinav Bhatele. 2019. Fast Profiling-based Performance Modeling of Distributed GPU Applications. *ACM Student Research Competition at International Conference for High Performance Computing, Networking Storage and Analysis (SC'19)*.

Jaemin Choi, Laxmikant V. Kale. 2017. Runtime Support for Concurrent Execution of Overdecomposed Heterogeneous Tasks. *ACM Student Research Competition at International Conference for High Performance Computing, Networking Storage and Analysis (SC'17)*.

TALKS

Jaemin Choi. 2021. GPU-aware Communication with Charm++. *Charm++ and AMPI: Adaptive and Asynchronous Parallel Programming, Birds of a Feather at International Conference for High Performance Computing, Networking Storage and Analysis (SC'21)*.

Jaemin Choi, David Hardy. 2021. Porting NAMD to DPC++. *oneAPI DevSummit at ISC'21*.

Nitin Bhat, **Jaemin Choi**. 2020. Charm++ with UCX. *UCF Virtual Workshop 2020*.

Jaemin Choi. 2020. Improving the Performance of Overdecomposed Applications on GPU-accelerated Systems. *15th CSL Student Conference (CSLSC'20) at University of Illinois Urbana-Champaign*. **Best Presentation Award**.

Jaemin Choi. 2019. Messaging with GPU-resident Data. *Charm++ and AMPI: Adaptive and Asynchronous Parallel Programming, Birds of a Feather at International Conference for High Performance Computing, Networking Storage and Analysis (SC'19)*.

Jaemin Choi. 2019. Distributed Deep Learning: Leveraging Heterogeneity and Data-Parallelism. *17th Annual Workshop on Charm++ and Its Applications*.

Jaemin Choi. 2019. Interoperability of Shared Memory Parallel Programming Models with Charm++. *17th Annual Workshop on Charm++ and Its Applications*.

Jaemin Choi. 2018. Recent Advances in Heterogeneous Computing using Charm++. *16th Annual Workshop on Charm++ and Its Applications*.

Laxmikant Kale, Michael Robson, Ronak Buch, **Jaemin Choi**. 2017. Migratable Objects and Task-Based Parallel Programming with Charm++. *Tutorial at International Conference for High Performance Computing, Networking Storage and Analysis (SC'17)*.

AWARDS & HONORS

HPC Session Best Presentation Award Feb 2020
15th CSL Student Conference (CSLSC '20), University of Illinois Urbana-Champaign

Graduated with Honors (Cum Laude) Feb 2016
Seoul National University

National Science and Technology Scholarship Mar 2010 - Feb 2016
Korea Scholarship Foundation

ACTIVITIES

Chair Positions 2018 - 2021
Annual Workshop on Charm++ and Its Applications

Student Volunteer Nov 2017
SC '17, Denver, Colorado

SNU Tomorrow's Edge Membership (STEM) Dec 2014 - Feb 2016
Honor Society, College of Engineering, Seoul National University

Korean Augmentation to the United States Army (KATUSA) Apr 2011 - Jan 2013
Military Service, KATUSA Training Academy/NCO Academy, Camp Jackson