Tanzima Islam

Assistant Professor
Texas State University

601 University Drive
San Marcos, TX, 78666
tanzima@txstate.edu
www.tanzimaislam.com

Research Objective

My research involves modeling both compute and data movement performance for large-scale applications in HPC environments. I am interested in leveraging machine learning for comparative performance modeling, focusing on HPC co-design in the short term. My long-term research objective is to leverage data-driven decision-making to address the challenges of developing scalable systems in extremely heterogeneous computing environments.

Education

2013 **Ph.D., Computer Engineering**, *Purdue University*.
Thesis: Reliable and scalable checkpointing systems for distributed computing environments.

2006 **B.Sc., Computer Science & Engineering**, *Bangladesh University of Engineering & Technology*.

Professional Experience

2020-Present **Assistant Professor**, *Department of Computer Science*, Texas State University.

Summer 2022 **Visiting Scholar**, *Brookhaven National Laboratory*, Computational Science Initiative.

Summer 2019 **Visiting Scholar**, *Lawrence Berkeley National Laboratory*, Computation Directorate.

2017-2019 **Assistant Professor**, *Department of Computer Science*, Western Washington University.

2013-2017 **Postdoctoral Research Staff Member**, *Lawrence Livermore National Laboratory*, Center for Applied Scientific Computing.
Developed machine-learning techniques for performance analysis.

2006-2007 **Member, Research and Development**, *Commlink Info Tech Ltd.*, Bangladesh.
Developed software for a service-independent telecommunication network (Intelligent Network).

Awards & Honors

2021 **Dean's College Achievement Award for Excellence in Scholarly/Creative Activities** at Texas State University.

2019-2020 **R&D 100 award**, in collaboration with Lawrence Livermore National Laboratory for the Scalable Checkpoint/Restart Framework 2.0 (SCR).

2014 **LLNL Director's Science & Technology Award.**

2016, 2015, 2014 Best Poster Award, Lawrence Livermore National Laboratory Annual Scholars Poster Symposium.

2014 2^{nd} Place Winner, LLNL Computation Directorate Postdoctoral Poster Symposium.

2012, 2009 Best Student Paper Nominations, International Conference for High Performance Computing, Networking, Storage and Analysis (SC).

2010 2^{nd} Place Winner, ACM Student Research Competition, Grace Hopper Celebration of Women in Computing.

Travel Awards Sustainable Research Pathways Program at Lawrence Berkeley National Laboratory'18.
CRA-W Career Mentoring Workshop at Phoenix, AZ'18.
Google Computer Science Grad Forum'12.
SC'09—'11, HPDC'12.

Research and Other Funding

2022-2027 **PI**, InteLytics: An Efficient Data-Driven Decision-Making Engine for Performance In the Era of Heterogeneity, DOE Early Career Research Program (ECRP), $770K.

2022-2025 **Co-PI**, Scalable Metadata And Provenance Services for Reproducible Hybrid Workflows, Department of Energy (DOE), $300K.

2022-2025 **Sub-contract**, ICE4HPC: Towards the Intelligent Center for HPC, Lawrence Livermore National Laboratory, $450K.

2022 **PI**, Performance Characterization of Workflow Applications, DOE SRP Fellowship, $68K.

2022 **PI**, Cross-platform performance prediction and analysis using deep learning, AMD Research Gift, $50K.

2021 **Member**, REU Site: Research Experiences for Undergraduates in Edge Computing, NSF, $389K.

2021 **Co-PI**, 5-petaflops Research cluster for COVID-19 research, AMD, $400K.

2021 **PI**, PerfROCm: A study of hardware resource utilization behaviors of HPC and ML applications on AMD GPUs, AMD Research Gift, $50K.

2021 **PI**, Predicting Performance using Few Shot Learning, Research Enhancement Program (REP) at Texas State University, $8K.

Summer 2019 **PI**, Proxy Application Validation for Exascale Co-design, DOE Sustainable Research Pathway Fellowship, $40K.

2018, 2019, 2021, 2022 **PI**, Parallel Computing course, Time allocation grant from XSEDE, 100,000 core-hours.

2018 **Co-PI**, Scientific Data Visualization course development, Office of Research and Sponsored Programs at Western Washington University, $12K.

2016 **PI**, VERITAS for Understanding Performance Evolution during Code Development, Linking Exploratory Application Research to Next-gen Development at Lawrence Livermore National Laboratory, $200K.

Publications

Peer Reviewed Conference & Journal Papers

[1] Holland Schutte, Chase Phelps, Aniruddha Marathe, and **Tanzima Islam**. LIBNVCD: An extendable and user-friendly multi-gpu performance measurement tool. In *46th Annual Computers, Software, and Applications Conference (COMPSAC)*. IEEE, 2022. (Accepted). Acceptance rate: 20%.

[2] Tanzima Z Islam, Philip Wu Liang, Forest Sweeney, Cody Pragner, Jayaraman J Thiagarajan, Moushumi Sharmin, and Shameem Ahmed. College life is hard!-shedding light on stress prediction for autistic college students using data-driven analysis. In *45th Annual Computers, Software, and Applications Conference (COMPSAC)*, pages 428–437. IEEE, 2021.

[3] Nathan Pinnow, Tarek Ramadan, **Tanzima Islam**, Chase Phelps, and Jayaraman J. Thiagarajan. Comparative code structure analysis using deep learning for performance prediction. In *International Symposium on Performance Analysis of Systems and Software (ISPASS)*. IEEE, March 28-30 2021.

[4] Tapasya Patki, Jayaraman J. Thiagarajan, Alexis Ayala, and **Tanzima Islam**. Performance optimality or reproducibility: that is the question. In *International Conference for High Performance Computing, Networking, Storage and Analysis (SC)*, November 17-22 2019. Acceptance rate: 20%.

[5] Jayaraman J. Thiagarajan, Rushil Anirudh, Bhavya Kailkhura, Nikhil Jain, **Tanzima Islam**, Abhinav Bhatele, Jae-Seung Yeom, and Todd Gamblin. PADDLE: Performance Analysis using a Data-driven Learning Environment. In *IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, May 21-25 2018. Acceptance rate: 24.5%.

[6] Teng Wang, Adam Moody, Yeh Zhu, Kathryn Mohror, Kento Sato, **Tanzima Islam**, and Waikuan Yu. MetaKV: A Key-Value Store for Metadata Management of Distributed Burst Buffers. In *IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, pages 1174–1183, May 2017. Acceptance rate: 22%.

[7] **Tanzima Islam**, Kathryn Mohror, and Martin Schulz. Exploring the MPI Tool Information Interface: Features and Capabilities. In *The International Journal of High Performance Computing Applications (IJHPCA)*, volume 30, pages 212–222, 2016.

[8] Tania Banerjee, Jason Hackl, Mrugesh Shringarpure, **Tanzima Islam**, S Balachandar, Thomas Jackson, and Sanjay Ranka. CMT-Bone — A Proxy Application for Compressible Multiphase Turbulent Flows. In *IEEE 23rd International Conference on High Performance Computing (HiPC)*, pages 173–182, Dec 2016. Acceptance rate: 23%.

[9] **Tanzima Islam**, Jayaraman J. Thiagarajan, Abhinav Bhatele, Martin Schulz, and Todd Gamblin. A Machine-Learning Framework for Performance Coverage Analysis of Proxy Applications. In *International*

Conference for High Performance Computing, Networking, Storage and Analysis (SC), Salt Lake City, UT, November 13-18 2016. Acceptance rate: 23%.

[10] Lee Savoie, David K. Lowenthal, Bronis R. de Supinski, **Tanzima Islam**, Kathryn Mohror, Barry Rountree, and Martin Schulz. I/O Aware Power Shifting. In *IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, pages 740–749, May 2016. Acceptance rate: 23%.

[11] Anup Mohan, Thomas Hacker, Gregory P. Rodgers, and **Tanzima Islam**. Batchsubmit: A high-volume Batch Submission System for Earthquake Engineering Simulation. In *Concurrency and Computation: Practice and Experience*, volume 26, pages 2240–2252. Wiley Online Library, 2014.

[12] **Tanzima Islam**, Saurabh Bagchi, and Rudolf Eigenmann. Reliable and Efficient Distributed Checkpointing System for Grid Environments. In *Journal of Grid Computing (JoGC)*, volume 12, pages 593–613, Dec 2014.

[13] **Tanzima Islam**, Kathryn Mohror, Saurabh Bagchi, Adam Moody, Bronis R De Supinski, and Rudolf Eigenmann. McrEngine: A Scalable Checkpointing System Using Data-Aware Aggregation and Compression. In *Scientific Programming*, volume 21, pages 149–163. Hindawi, 2013.

* [14] **Tanzima Islam**, Kathryn Mohror, Saurabh Bagchi, Adam Moody, Bronis R. de Supinski, and Rudolf Eigenmann. McrEngine: A Scalable Checkpointing System Using Data-aware Aggregation and Compression. In *International Conference on High Performance Computing, Networking, Storage and Analysis (SC)*, pages 17:1–17:11, 2012. Acceptance rate: 20%. **Best student paper finalist.**

* [15] **Tanzima Islam**, Saurabh Bagchi, and Rudolf Eigenmann. FALCON: A System for Reliable Checkpoint Recovery in Shared Grid Environments. In *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis (SC)*, pages 1–12, 2009. **Best student paper nomination.**

[16] Hemayet Hossain, Mostofa Ahmed, Abdullah Al-Nayeem, **Tanzima Islam**, and Md Mostofa Akbar. gpNoCSim - A General Purpose Simulator for Network-On-Chip. pages 254–257, March 2007.

Workshop Papers

[17] **Tanzima Islam** and Chase Phelps. Hpc@scale: A hands-on approach for training next-gen hpc software architects. In *EduHiPC workshop at the 28th International Conference on High Performance Computing, Data, Analytics (HiPC)*. IEEE, 2021. Invited paper.

[18] Quentin Jensen, Filip Jagodzinski, and **Tanzima Islam**. Filcio: Application agnostic i/o aggregation to scale scientific workflows. In *2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)*, pages 1587–1592. IEEE, 2021.

[19] Jack Stratton, Michael Albert, Quentin Jensen, Max Ismailov, Filip Jagodzinski, and **Tanzima Islam**. Towards aggregation based i/o optimization for scaling bioinformatics applications. In *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, pages 1250–1255, 2020.

[20] **Tanzima Islam**, Alexis Ayala, Quentin Jensen, and Khaled Ibrahim. Towards A Programmable Analysis and Visualization Framework for Interactive Performance Analytics. In *Workshop on Programming and Performance Visualization Tools held in conjunction with International Conference for High Performance Computing, Networking, Storage and Analysis (SC)*, Denver, CO, November 16-23 2019. IEEE.

[21] Nicholas Majeske, Filip Jagodzinski, Brian Hutchinson, and **Tanzima Islam**. Low Rank Smoothed Sampling Methods for Identifying Impactful Pairwise Mutations. In *International Conference on Bioinformatics, Computational Biology, and Health Informatics*, pages 681–686. ACM, 2018.

[22] Aiman Fang, Ignacio Laguna, Kento Sato, **Tanzima Islam**, and Kathryn Mohror. Fault Tolerance Assistant (FTA): An Exception Handling Programming Model for MPI Applications. 2015.

[23] **Tanzima Islam**, Kathryn Mohror, and Martin Schulz. Exploring the Capabilities of the New MPI_T Interface. In *Proceedings of the 21st European MPI Users' Group Meeting*, page 91. ACM, 2014.

[24] John Tramm, Andrew Siegel, **Tanzima Islam**, and Martin Schulz. XSBench-the Development and Verification of a Performance Abstraction for Monte Carlo Reactor Analysis. *The Role of Reactor Physics toward a Sustainable Future (PHYSOR)*, 2014.

Ph.D. Dissertation

[25] **Tanzima Islam**. *Reliable and scalable checkpointing systems for distributed computing environments*. PhD thesis, Purdue University, West Lafayette, IN, May 2013.

Research Posters

[26] Alicia Guite, **Tanzima Islam**, Chris Kelly, and Wei Xu. Interactive Visual Analysis Tool for Anomaly Provenance Data. In *International Conference for High Performance Computing, Networking, Storage and Analysis (SC'22)*, Nov 2022.

[27] Mohammad Zaeed, **Tanzima Islam**, Younghyun Cho, Sherry Li, Hangrui Luo, and Yang Liu. Analysis and Visualization of Important Performance Counters To Enhance Interpretability of Autotuner Output. In *International Conference for High Performance Computing, Networking, Storage and Analysis (SC'22)*, Nov 2022.

[28] Logan Moody, Nathan Pinnow, Michael Lam, Harshitha Menon, Markus Schordan, Scott G. Lloyd, and **Tanzima Islam**. Automatic Generation of Mixed-Prevision Programs. In *International Conference for High Performance Computing, Networking, Storage and Analysis (SC'18)*, 2018.

[29] Simone Smarr, **Tanzima Islam**, and Yolanda Rankin. Modular Extensible Framework for Performance Comparative Analysis. In *ACM Richard Tapia Celebration of Diversity in Computing Conference*, 2015.

[30] Xiang Ni, **Tanzima Islam**, Kathryn Mohror, Adam Moody, and Laxmikant V. Kale. Lossy Compression for Checkpointing: Fallible or Feasible. In *International Conference for High Performance Computing, Networking, Storage and Analysis (SC'14)*, 2014.

* [31] **Tanzima Islam**, Saurabh Bagchi, and Rudolf Eigenmann. Harnessing multiple cores for efficient checkpointing in grid systems. In *ACM Student Research Competition (SRC) at Anita Borg Institute Grace Hopper Celebration of Women in Computing*, 2010. **ACM Student Research Competition 2^{nd} prize winner.**

[32] **Tanzima Islam**, Kathryn Mohror, Adam Moody, Bronis de Supinski, Saurabh Bagchi, and Rudolf Eigenmann. Data-Aware Inter-Process Checkpoint Compression. In *International Conference for High Performance Computing, Networking, Storage and Analysis (SC'10)*, 2010.

[33] Mohammad S. Hossain, **Tanzima Islam**, Saurabh Bagchi, and Vijay Raghunathan. Fast and Collaborative Interference Avoidance for Wireless Medical Devices. In *International Conference on Dependable Systems and Networks (DSN)*, 2009.

Professional Activities

Organizer Co-Chair, Performance track, Cluster'23. Chair, Performance track, SC'22. Vice/Co-Chair, Performance track, International Conference on Parallel Processing, 2019. IEEE International Workshop on Big Data Computation, Analysis & Applications in conjunction with IEEE COMPSAC, 2019-Present.

Technical Review Board IEEE Transactions on Parallel and Distributed Systems (TPDS)

Technical Program Committees High-Performance Distributed Computing (HPDC)'22, International Parallel & Distributed Processing Symposium (IPDPS)'17–, International Conference for High Performance Computing, Networking, Storage and Analysis (SC)'17– (Performance track). SC'19-'21 (Posters). SC'18 (HPC for Undergrad), International Conference in Supercomputing (ISC)'19–, Women in HPC Technical Conference'20, Euro/MPI User Forum'20, The Platform for Advanced Scientific Computing (PASC)'19, International Conference on Parallel Processing (ICPP)'18, International Symposium on Computer Architecture and High Performance Computing (SBAD-PAD)'13–'14, Symposium on Principles and Practice of Parallel Programming (PPoPP) '11, Dependable Systems and Networks (DSN)'10, '22.

Grant Review Committees DOE SBIR, NSF CISE/CSSI, NSF HBCU-UP, NSF CISE/OAC, NSF CISE/CCF, DOE SciDAC.

Journal Reviewing IEEE Transactions on Parallel and Distributed Systems (TPDS)'19–, International Conference on Parallel Computing (ParCo)'19-'21, Journal of Grid Computing (JoGC)'17–'21, International Journal of High Performance Computing (IJHPCA)'18.

Memberships IEEE, ACM

Invited Talks and Panels

September, 2022 **Bangladesh University of Engineering and Technology (BUET)**, *How to apply to Graduate School: Things I wish I knew*, Virtual.

May, 2022 **How to be a Great Mentor**, *Exascale Computing Project (ECP)*, Virtual.

April, 2022 **Scalability challenges and opportunities for I/O bound applications**, *CHEOPS Workshop at EuroSys*, Virtual, https://tinyurl.com/2xyb2wnr.

March, 2022 **Exascale Computing Project (ECP): Enabling Next-Generation of Hardware-Software Co-design using Data Science**, *LLNL*, Virtual.

February, 2022 **Sustainable Horizons Institute**, *SHI*, Virtual.

Dec, 2021 **Characterizing Performance of Workflow Applications**, *Brookhaven National Laboratory*, Virtual.

Dec, 2021 **Center of Excellence: Performance Characterization of Deep Learning Workloads**, *AMD*, Virtual.

Sep, 2021 **Enabling Next-Generation Software and Co-Design**, *AMD at Texas State*, San Marcos, TX.

Sep, 2021 **HPC Enables Digitalization Panel**, *Digital 360 Summit*, San Marcos, TX.

Nov, 2020 **Careers in HPC Panel, Taking the Leap: Changing Careers**, *International Conference for High Performance Computing, Networking, Storage and Analysis (SC)*, Virtual.

Oct, 2020 **May the HPC Force be with you!**, *Organized a panel for undergraduate students*, Richard Tapia Celebration of Diversity in Computing.

Aug, 2020 **Data-Driven Performance Modeling for HPC Co-design**, *Invited scholar's talk*, AMD.

Aug, 2020 **DASHING: An Interpretable Machine Learning Toolkit for Performance Analysis and Visualization**, *Exascale Computing Project Hackathon*, Oak Ridge National Laboratory.

May, 2019 **The NIMBioS Workshop on Scientific Collaboration Enabled by High Performance Computing**, *Scalable I/O Performance for Scientific Applications–Challenges and Potentials*, Knoxville, TN.

December, 2018 **Lawrence Berkeley National Laboratory**, *Understanding the Performance Portability Challenges and Opportunities using Machine Learning*, Berkeley, CA.

November, 2018 **HPC for Undergraduates**, *International Conference for High Performance Computing, Networking, Storage and Analysis*, Dallas, TX.

July, 2018 **The Platform for Advanced Scientific Computing Conference**, *Machine Learning Framework for Performance Coverage Analysis of Proxy Applications*, ACM and the Swiss National Supercomputing Center (CSC), Basel, Switzerland.

February, 2015 **JOWOG-34**, *Proxy Application Validation using Machine Learning (Veritas)*, Sandia National Laboratories, Albuquerque, NM.

January, 2014 **Spellman College**, *Opportunities after Graduate School*, Atlanta, GA.

Teaching Experience

2017 - Present Foundations of CS-II, Scalable Systems for Supercomputing, Compiler Construction, Parallel Computing, Data Structures, Computer Networks, Scientific Data Visualization (co-instructor).

Outreach and Mentoring Students

2014-Present **Co-founder**, *Bangladeshi Women in Computer Science and Engineering*, Dhaka, Bangladesh, First research, networking, and mentoring platform for female Computer Science and Engineering students of Bangladesh.

2017–2019 **Research advisor**, *Undergraduate (7) and graduate (3) students (Western Washington University); Undergraduate (6) and graduate (7) students*, Texas State University, High school (2).

2017-Present **Faculty advisor**, *ACM-W chapter at Western Washington University*, ACM-ICPC Programming Club, Faculty mentor.

2014–2016 **Intern supervisor**, *Undergraduate, M.S. and Ph.D. students from University of Hamburg, University of Illinois Urbana-Champaign, Spelman College, University of California San Diego*, NSF REU Advisor, Marshalls College, University of California Berkeley.

2019-Present **Mentor**, *Graduate students (3)*, SC Student Volunteer Program.

2011–2013 **Student mentor**, *Undergrad, M.S., and Ph. D. students*, Purdue University.