MICHELA TAUFER

Jack Dongarra Professor in High Performance Computing
ACM Distinguished Scientist and IEEE Senior Member
Department of Electrical Engineering and Computer and Science
University of Tennessee Knoxville
Min H. Kao Building 401G Knoxville, TN 37996
Email: taufer@utk.edu
Phone: (302) 690 7845

EDUCATION

Dec 2002 Ph.D. in Computer Science, Swiss Federal Institute of Technology Zurich (ETH), Switzerland
Dissertation title: Inverting Middleware: Performance Analysis of Layered Application Codes in High-Performance Distributed Computing.
Thesis supervisors: Thomas M. Stricker (Chair) and Daniel A. Reed

Dec 1996 MS (Laurea) in Computer Engineering, University of Padua, Italy
Dissertation title: Development of the Parallelization of the Software Package OPAL for the Simulation of Dynamic Molecules on Supercomputers.
Thesis supervisors: Gianfranco Bilardi (Chair), Walter Gander, and Geppino Pucci

RESEARCH INTERESTS

High-performance computing, cloud computing, and volunteer computing; algorithms and workflows for scientific applications; reproducibility, replicability, and transparency of scientific applications; performance analysis, modeling, and optimization of multi-scale applications; in situ and in transit data analytics.

ACADEMIC POSITIONS

Jun 2018 – present Dongarra Professor in High Performance Computing, Department of Electrical Engineering and Computer and Science, University of Tennessee Knoxville

Sep 2017 – May 2018 Professor, Department of Computer and Information Sciences, University of Delaware (Affiliated with Biomedical Engineering and Center for Bioinformatics & Computational Biology)

Sep 2012 – Aug 2017 Associate Professor, Department of Computer and Information Sciences, University of Delaware, (Affiliated with Biomedical Engineering and Center for Bioinformatics & Computational Biology)

Sep 2015 – Aug 2016 Acting Director, Center for Bioinformatics & Computational Biology, Delaware Biotechnology Institute

Jan 2013 – Aug 2016 David and Beverly J.C. Mills Career Development Chair, Department of Computer and Information Sciences, University of Delaware

Jun 2013 – May 2014	Visiting Faculty, Computer Science and Mathematics Division, Oak Ridge National Laboratory. U.S. Department of Energy Higher Education Research Experiences Faculty Program
Sep 2007 – Aug 2012	Assistant Professor, Department of Computer and Information Sciences, University of Delaware (Affiliated with Center for Bioinformatics & Computational Biology Jun 2010 – Aug 2012)
Jan 2005 – Aug 2007	Assistant Professor, Department of Computer Science, University of Texas, El Paso
Jan 2003 – Dec 2004	Postdoctoral Researcher, Center for Theoretical Biological Physics, University of California, San Diego (Affiliated with the Department of Molecular Biology, The Scripps Research Institute; San Diego Supercomputer Center; Department of Computer Science and Engineering, University of California San Diego)
Dec 1996 – Dec 2002	Research Student Assistant, Computer Systems Institute, Swiss Federal Institute of Technology Zurich (ETH)
Feb 1996 – Dec 1996	Visiting Scholar at the Swiss Center for Scientific Computing (SCSC/CSCS), Zurich

HONORS AND DISTINCTIONS

May 2023	2023 Provost Award for Research and Creative Achievement, University of Tennessee
May 2022	2022 IEEE Technical Community on Parallel Processing (TCPP) Outstanding Service and Contributions Award
April 2022	Tickle College of Engineering Research Achievement Award, University of Tennessee Knoxville
Dec 2021	IBM Faculty Award
Aug 2021	R&D 100 Award Winner in Software/Services categories with the project "Flux: Next-Generation Workload Management Software Framework." Collaboration with Lawrence Livermore National Laboratory
Jun 2021	Best Track Paper Award, International Conference on Computational Science (ICCS)
Apr 2020	IEEE Senior Member
Mar 2020	Tickle College of Engineering Faculty & Staff Award for Outstanding Service to the Discipline Category, University of Tennessee Knoxville
Feb 2019	IBM Faculty Award
Feb 2019	HPCwire's 2019 Person to Watch
Jun 2018	Dongarra Professor in High Performance Computing
May 2017	Faculty Nomination for the Excellence in Undergraduate Academic Advising and Mentoring Award, University of Delaware
Feb 2017 – May 2018	J.P. Morgan Chase Faculty Scholar, University of Delaware

2015	ACM (Association for Computing Machinery) Distinguished Scientist
2015	Winner of the 8th IEEE International Scalable Computing Challenge, Co-located with the IEEE/ACM CCGrid Conference
2014	ACM (Association for Computing Machinery) Senior Member
May 2014	Faculty Nomination for the Excellence in Undergraduate Academic Advising and Mentoring Award, University of Delaware
2012 – 2016	David and Beverly J.C. Mills Career Development Chair, University of Delaware
May 2006	UTEP Young Investigator Award, Research, and Sponsored Programs, University of Texas El Paso
Jan 2003 – Dec 2004	La Jolla Interfaces in Science (LJIS) Interdisciplinary Fellows, University of California San Diego
Feb 1996 - Dec 1996	Erasmus Fellow, European Community (EU)

RESEARCH FUNDING

In progress:

2023 *Collaborative Research: SHF: Small: Model-driven Design and Optimization of Dataflows for Scientific Applications*
Source of Support: National Science Foundation #2331152
Total Amount: $624,000 ($424,000 at University of Tennessee Knoxville), Lead PI. Collaborative research with Ewa Deelman, University of South California
Project Period: 10/01/2023 - 09/30/2025
Location of Project: University of Tennessee Knoxville
Description: This project has four main research components. First, the project defines a taxonomy of common dataflow motifs used in scientific domains, ranging from simple producer-consumer pairs to complex pipelines with multiple producers and consumers, by mapping these motifs to real scientific applications. Second, the project designs a middleware layer to handle dataflow pipelines executing on HPC, cloud, and edge resources. Third, the project develops a 2-step model for mitigating pipelines that result in data loss and inefficiencies associated with the slowdown in data production or consumption in dataflow pipelines. Finally, the project trains a broader community to utilize the taxonomy, middleware, and model to optimize real scientific applications by identifying potential bottlenecks and making necessary adjustments to maximize pipeline efficiency and accuracy, continuously monitoring and optimizing pipelines to ensure the highest quality scientific output possible.

2023 *Data-aware scheduling with the convergence of HPC and Cloud*
Source of Support: Lawrence Livermore National Laboratory
Total Amount: $100,067 Single PI
Project Period: 08/01/2023 – 07/31/2024
Location of Project: University of Tennessee Knoxville

Description: We develop data-aware scheduling methodologies for scientific workloads and enhance a tool to manage data locality across the phases of complex computational workflows via LLNL's DYAD framework. This framework handles producer-consumer data file-sharing in scientific workflows under the Flux workload management system. We augment DYAD to track performance metrics at runtime, design policies that enable data-aware scheduling for DYAD, and study approaches to data movement to enable DYAD to be used in converged computing environments. We evaluate the benefit of data-aware scheduling combined with the data-file sharing tool and explore data-sharing mechanisms in converged computing environments.

2023 *Studying the Automation of Choice-based Performance Optimization Workflows for HPC*
Source of Support: Lawrence Livermore National Laboratory
Total Amount: $100,000, Single PI
Project Period: 08/01/2023 – 07/31/2024
Location of Project: University of Tennessee Knoxville
Description: We automate choice-based performance optimization workflows for HPC systems relevant to LLNL by building a model to predict which compilers among a suite used at LLNL generate faster code on a mixed workload, extending the work from multiple compilers to multiple platforms and libraries, and validating the model with diverse applications at LLNL.

2023 *Study Performance Portability of the Vector Particle-In-Cell Project (VPIC) across architectures (Stage 3)*
Source of Support: Los Alamos National Laboratory
Total Amount: $120,000, Single PI
Project Period: 08/01/2023 – 07/31/2024
Location of Project: University of Tennessee Knoxville
Description: We collaborate with the Vector Particle-In-Cell Project (VPIC) group on optimizing and analyzing code performance across multiple architectures, emphasizing the costs and benefits of using the performance portability framework Kokkos. Testing platforms involve a variety of host architectures and accelerators, including Intel Xeon CPUs, AMD EPYC CPUs and Radeon Instinct GPUs (El Capitan), IBM Power8 and Power9 CPUs, Cavium ThunderX and ThunderX2 ARM CPUs, Intel Xeon Phi Knights Landing and Knights Mill processors, Various Nvidia GPUs, and Fujitsu A64FX. As part of this work, we document the optimization techniques and performance portability impacts for future use in other LANL code bases.

2022 *SHF: Small: Methods, Workflows, and Data Commons for Reducing Training Costs in Neural Architecture Search on High-Performance Computing Platforms*
Source of Support: NSF #2223704
Total Amount: $623,999, Lead PI, with Silvina Caino-Lores (UTK) and Catherine Schuman (UTK).
Project Period: 10/01/2022 - 09/30/2025
Location of Project: University of Tennessee Knoxville

Description: This project addresses reducing high-performance computing resources to train neural networks while assuring explainable, reproducible, and nearly optimal neural networks. To this end, the researchers propose a flexible fitness-prediction method that uses parametric modeling to predict the future fitness of neural networks and allow for early termination of the training process.

2021 *OAC: Piloting the National Science Data Fabric: A Platform Agnostic Testbed for Democratizing Data Delivery*
Source of Support: NSF #2138811
Total Amount: $5,609,259 ($750,000 at University of Tennessee Knoxville), co-PI.
Project Period: 10/01/2021 - 09/30/2024
Location of Project: University of Tennessee Knoxville
Description: This project aims to build a National Science Data Fabric (NSDF), a testbed experimenting with critical technology needed to democratize data-driven sciences by constructing a CI platform designed for equitable access. NSDF connects storage, computing, and networking components with a software stack that empowers end-users with easy-to-use, integrated, and scaleable tools. Community-driven education and outreach will guarantee equitable access to all resources and engage an open network of universities, including minority-serving institutions, in a federated data fabric configurable for individual and shared scientific use.

2021 *Collaborative Research: Elements: SENSORY: Software Ecosystem for kNowledge diScOveRY - a data-driven framework for soil moisture applications*
Source of Support: National Science Foundation #2028923
Total Amount: $600,000 ($349,998 at University of Tennessee Knoxville), Lead PI. Collaborative research with Rodrigo Vargas, University of Delaware
Project Period: 06/01/2021 - 05/30/2024
Location of Project: University of Tennessee Knoxville
Description: This project connects multi-disciplinary advances across the scientific community (such as generating datasets at scale and supporting cloud-based cyberinfrastructures) to develop a data-driven software ecosystem for analyzing, visualizing, and extracting knowledge from the growing data collections (from fine-grained, in situ soil sensor information to coarse-grained, global satellite measurements) and releasing this knowledge to applications in environmental sciences.

2021 *Leveraging Kokkos Abstractions to Automate Checkpointing*
Source of Support: Argonne National Laboratory
Total Amount: $82,563, Single PI
Project Period: 05/01/2021 – 12/31/2023
Location of Project: University of Tennessee Knoxville
Description: This project focuses on what patterns the memory abstractions used by Kokkos form and how they can be efficiently captured and persistent with VELOC.

2020 *Leverage Containerized Environments for Reproducibility and Traceability of Scientific Workflows - the case study of Analytics for Neural Network Workflows*
Source of Support: Sandia National Laboratories
Total Amount: $400,000, Single PI
Project Period: 07/15/2020 – 07/14/2024

Location of Project: University of Tennessee Knoxville
Description: This project builds a prototype of a containerized environment that encapsulates each component of a scientific workflow (i.e., data and applications) in an individual container environment for transparent and automatic metadata collection and access, easy-to-read record trail, and tight connections between data and metadata.

2020 *Flux Scheduler Specializations: Improving Workflow Performance with Scheduler Structure and Policy Tuning*
Source of Support: Lawrence Livermore National Laboratory
Total Amount: $200,089, Single PI
Project Period: 04/01/2020 – 07/31/2023
Location of Project: University of Tennessee Knoxville
Description: This project studies how features of a user-level, highly configurable scheduler like Flux can best be leveraged to maximize workflow performance. The project aims to answer this question by developing a model that tunes scheduler settings to maximize workflow performance even under system stress conditions such as scheduler fragmentation and resource drains.

2020 *Augmenting Hatchet to support scalability and replicability solutions for HPC applications*
Source of Support: Lawrence Livermore National Laboratory
Total Amount: $216,000, Single PI
Project Period: 08/01/2020 - 07/31/2023
Location of Project: University of Tennessee Knoxville
Description: This project uses Hatchet and its features (e.g., query language) to study scalability and replicability problems in applications of interest to LLNL at a large scale and develops tooling to support the analysis and study of such problems to identify the source of the scalability and replicability problems.

2019 *SHF: Medium: Collaborative Research: ANACIN-X: Analysis and Modeling of Nondeterminism and Associated Costs in eXtreme Scale Applications*
Source of Support: National Science Foundation #1900888
Total Amount: $1,247,940 ($931,739,739 at University of Tennessee Knoxville), Lead PI. Collaborative research with Heike Jagode, University of Tennessee Knoxville, and Sanjukta Bhowmick, University of North Texas
Total Project Period: 06/01/2019 – 05/31/2024
Location of Project: University of Tennessee Knoxville
Description: This project advances the study of nondeterministic HPC applications by studying the recording costs of Record-and-replay (R&R) tools and defining strategy so that these tools can scale to the exascale domain.
Supplement: NSF Research Experiences for Undergraduates (REU), $16,000, Summer 2019 and Summer 2020

2018 *BIGDATA: IA: Collaborative Research: In Situ Data Analytics for Next Generation Molecular Dynamics Workflows*
Source of Support: National Science Foundation #1741057/#1841758
Total Amount: $2,293,043 ($1,079,986 at University of Tennessee Knoxville), Lead PI.

Collaborative research with Trilce Estrada, University of New Mexico; Ewa Deelman and Rafael Ferreira da Silva, University of Southern California; Michel Cuendet and Harel Weinstein, Weill Medical College of Cornell University
Project Period: 10/01/2017 – 09/30/2023
Location of Project: University of Tennessee Knoxville
Description: This interdisciplinary project tackles the data challenge of data analysis of molecular dynamics simulations on next-generation supercomputers. Specifically, this effort combines machine learning and data analytics approaches, workflow management methods, and high-performance computing techniques to analyze molecular dynamics data as it is generated.

Completed:

2020 *Flux Scheduler Specializations: Improving Workflow Performance with Scheduler Structure and Policy Tuning*
Source of Support: Lawrence Livermore National Laboratory
Total Amount: $200,089, Single PI
Project Period: 04/01/2020 – 07/31/2023
Location of Project: University of Tennessee Knoxville
Description: This project studies how features of a user-level, highly configurable scheduler like Flux can best be leveraged to maximize workflow performance. The project aims to answer this question by developing a model that tunes scheduler settings to maximize workflow performance even under system stress conditions such as scheduler fragmentation and resource drains.

2020 *Augmenting Hatchet to support scalability and replicability solutions for HPC applications*
Source of Support: Lawrence Livermore National Laboratory
Total Amount: $216,000, Single PI
Project Period: 08/01/2020 - 07/31/2023
Location of Project: University of Tennessee Knoxville
Description: This project uses Hatchet and its features (e.g., query language) to study scalability and replicability problems in applications of interest to LLNL at a large scale and develops tooling to support the analysis and study of such problems to identify the source of the scalability and replicability problems.

2020 *Collaborative Research: PPoSS: Planning: Performance Scalability, Trust, and Reproducibility: A Community Roadmap to Robust Science in High-throughput Applications*
Source of Support: National Science Foundation #2028923
Total Amount: $250,000 ($90,000 at University of Tennessee Knoxville), Lead PI. Collaborative research with Ewa Deelman, University of South California; Trilce Estrada, University of New Mexico; Mary Hall, University of Utah; and Victoria Stodden, University of Illinois at Urbana-Champaign
Project Period: 10/01/2020 - 09/30/2022
Location of Project: University of Tennessee Knoxville
Description: The project recruits a cross-disciplinary community working together in three virtual mini-workshops called virtual world cafes to define, design, implement, and use solutions for robust science.

2020 *Collaborative Research: EAGER: Advancing Reproducibility in Multi-Messenger Astrophysics*
Source of Support: National Science Foundation #2041977
Total Amount: $300,000 ($100,000 at University of Tennessee Knoxville), Lead PI. Collaborative research with Ewa Deelman, University of South California, and Duncan Brown, Syracuse University
Project Period: 08/01/2020-07/31/2022
Location of Project: University of Tennessee Knoxville
Description: The project provides the astrophysics community with a transformative building block to a roadmap for reproducible open science. Findings about the reproducibility process of the EHT and NICER results are captured and disseminated through documentation, data products, and methods used.

2020 *Study Performance Portability of the Vector Particle-In-Cell Project (VPIC) across architectures*
Source of Support: Los Alamos National Laboratory
Total Amount: $165,179, Single PI
Project Period: 05/18/2020 – 04/30/2023
Location of Project: University of Tennessee Knoxville
Description: This project studies aspects of performance portability associated with the Vector Particle-In-Cell Project or VPIC code across platforms by addressing questions such as "Is the execution of VPIC sensitive to new architectures on which it runs? How do we continue to extract as much performance as possible despite differences in hardware? What performance is lost when using a performance portability framework?"

2019 *JDRD: Empowering Training and Validation Stages in AI-Orchestrated Workflows*
Source of Support: Science Alliance, University of Tennessee Knoxville
Total Amount: $216,900, Single PI
Project Period: 10/01/2019 – 09/30/2021
Location of Project: University of Tennessee Knoxville
Description: This project studies AI-orchestrated workflows, including experimental, computational, and data manipulation steps in one or multiple domains, where an important component is one or more neural networks (NN) used for searching or decision-making. The project aims to transform the training process of NN in AI-orchestrated workflows from simulated data (clean, non-adversarial data) to deploying on real data (noisy, adversarial data) with the integration of mitigating methods.

2019 *EAGER: Reproducibility in Computational and Data-Enabled Science - Paradigms, Practices, and Infrastructure*
Source of Support: National Science Foundation #1941443
Total Amount: $300,000 ($149,997 at University of Tennessee Knoxville), PI. Collaborative research with Victoria Stodden, University of Illinois at Urbana-Champaign
Total Project Period: 08/16/2019 – 08/15/2022
Location of Project: University of Tennessee Knoxville

Description: This project seeks to improve understanding of how the scientific community can adapt to the increasing use of computing and large-scale data resources. One challenge is ensuring that computational results, such as those from simulations, are "reproducible" and the same results are obtained when one re-uses the same input data, methods, software, and analysis conditions. In 2019, the National Academies of Science, Engineering, and Medicine (NASEM) issued a report on "Reproducibility and Replication in Science" with recommendations. The project assesses the implications of these recommendations on the scientific discovery process for computationally- and data-enabled research.

2019 *Building a "Miniature" Version of the ORNL's Summit supercomputer for Computational Science Research at UTK*
Source of Support: 2019 IBM Global University Program Shared University Research Award
Total Amount: $472,536, Leading PI. Collaborative research with Jack Dongarra, Mark Dean, and Greg Peterson at the University of Tennessee Knoxville
Total Project Period: 06/21/2019 - 06/20/2024
Location of Project: University of Tennessee Knoxville
Description: The award enabled the purchase of a supercomputer for computational science applications at the University of Tennessee Knoxville.

2019 *Study of Data-intensive Workflows on Next-generation Systems with Emphasis on Memory Access*
Source of Support: Sandia National Laboratories
Total Amount: $99,999, Single PI
Project Period: 03/15/2019 – 07/31/2020
Location of Project: University of Tennessee Knoxville
Description: The project designs and implements a C++ suite of data-intensive mini-applications to study data management costs, emphasizing memory access times and use power consumption and replicability.

2019 *Moving Towards Self-Adjusting Scheduling Policies for High-Performance Workflows with Flux's Fully Hierarchical Scheduling*
Source of Support: Lawrence Livermore National Laboratory
Total Amount: $63,107, Single PI
Project Period: 02/08/2019 – 02/30/2020
Location of Project: University of Tennessee Knoxville
Description: The project tackles scheduler specializations by systematically studying fully hierarchical scheduling models with Flux and defining models supporting a given workflow to employ the best scheduler specialization strategy at runtime.

2018 *Driving Next-Generation Schedulers with Machine Learning-Based Application Patterns*
Source of Support: Lawrence Livermore National Laboratory
Total Amount: $225,663, Single PI
Project Period: 08/01/2018 – 07/31/2020
Location of Project: University of Tennessee Knoxville

Description: This project develops methods to identify and understand irregular HPC job patterns and integrates knowledge of these irregular HPC patterns into multi-objective schedulers. The work leverages the results of a previous award from Lawrence Livermore National Laboratory.

2018 *Collaborative: EAGER: Exploring and Advancing the State of the Art in Robust Science in Gravitational Wave Physics*
Source of Support: National Science Foundation #1823372
Total Amount: $300,000 ($75,000 at University of Tennessee Knoxville), PI. Collaborative research with Ewa Deelman, University of Southern California; Duncan Brown, Syracuse University; and Von Welch, Indiana University
Project Period: 05/31/2018 – 04/30/2020
Location of Project: University of Tennessee Knoxville
Description: The project develops and uses a survey to collect information about LIGO workflows composed of a series of experimental, computational, and data manipulation steps.

2017 *CIF21 DIBBs: PD: Cyberinfrastructure Tools for Precision Agriculture in the 21st Century*
Source of Support: National Science Foundation #1724843/#1854312
Total Amount: $574,999 ($339,497 at University of Tennessee Knoxville), Leading PI. Collaborative research with Rodrigo Vargas, University of Delaware
Project Period: 07/01/2017 – 06/30/2021
Location of Project: University of Tennessee Knoxville
Description: This interdisciplinary project applies computer science approaches and computational resources to large multidimensional environmental datasets and synthesizes this information into ecoinformatics, a branch of informatics that analyzes ecological and environmental science variables such as information on landscapes, soils, climate, organisms, and ecosystems.

2017 *Investigating Massively Scalable I/O-Aware Job Scheduling in Support of Flux (III)*
Source of Support: Lawrence Livermore National Laboratory
Total Amount: $112,014, Single PI
Project Period: 06/01/2017 – 05/31/2018
Location of Project: University of Delaware
Description: This project investigates distinct—yet complementary—techniques to overcome challenges that can preclude I/O-aware schemes from effectively scheduling massively large-scale systems. The work leverages the results of a previous award from Lawrence Livermore National Laboratory.

2016 *HAKER-HPC: HArnessing Knowledge for Environmental Research using High-Performance Computing (HPC) Solutions*
Source of Support: University of Delaware Research Foundation
Total Amount: $45,000, Senior Personnel Project Period: 12/01/2016 – 05/31/2018 Location of Project: University of Delaware
Description: This award aims to build preliminary results on developing HPC-based tools for analyzing moisture in soil data at a large scale.

2016 *Student Support: IEEE Cluster 2017 Conference*

Source of Support: National Science Foundation #1648617
Total Amount: $20,000, Single PI
Project Period: 08/01/2016 – 07/31/2018
Location of Project: University of Delaware
Description: This award supports 20 students from American institutions to attend the IEEE Cluster 2017 conference.

2016 *Performance Characterization and Optimization of the MapReduce-MPI Framework*
Source of Support: Argonne National Laboratory
Total Amount: $14,760, Single PI
Project Period 09/01/2016 – 09/30/2016
Location of Project: University of Delaware
Description: This project studies the performance of MapReduce-MPI on high-end clusters and identifies the performance bottlenecks for a selected number of popular benchmarks.

2016 *Development of a Scalable Method for Identifying Dietary Clusters in the National Health and Nutrition Examination Survey using MapReduce*
Source of Support: University of Delaware - UDRF
Total Amount: $38,500, PI (previous PI Mia Papas)
Project Period: 06/011/2016 – 05/31/2018
Location of Project: University of Delaware
Description: This award aims to build preliminary results on developing MapReduce-based tools for analyzing dietary data at a large scale.

2016 *Investigating Massively Scalable I/O-Aware Job Scheduling in Support of Flux (Part II)*
Source of Support: Lawrence Livermore National Laboratory
Total Amount: $164,271, single PI
Project Period: 03/31/2016 – 05/31/2017
Location of Project: University of Delaware
Description: This project investigates distinct—yet complementary—techniques to overcome challenges that can preclude I/O-aware schemes from effectively scheduling massively large-scale systems. The work leverages the results of a previous award from Lawrence Livermore National Laboratory.

2015 *BD Hubs: Collaborative Proposal: SOUTH: A Big Data Innovation Hub for the South Region*
Source of Support: National Science Foundation #1550305
Total Amount: $750,712 (Funds only to leading hub instructions: Georgia Tech and University of North Carolina, Chapel Hill), Senior Personnel
Project Period: 09/15/15 – 09/30/18
Location of Project: University of Delaware
Description: This award establishes the South Big Data Regional Innovation Hub (South BD Hub) with lead institutions, Georgia Institute of Technology (GT) and the University of North Carolina at Chapel Hill (UNC-CH). The South BD Hub serves as the primary vehicle for interdisciplinary, multi-stakeholder partnerships designed to pursue BD projects of interest to the South region by engaging academic institutions from 16 states in the South, including the University of Delaware

2015 *SHF: Medium: Collaborative Research: A Comprehensive Methodology to Pursue Reproducible Accuracy in Ensemble Scientific Simulations on Multi- and Many-core Platforms*
Source of Support: National Science Foundation #1513025/#1841552
Total Amount: $814,733 ($443,878 at University of Delaware and University of Tennessee Knoxville), Leading PI. Collaborative research with Michela Becchi, North Carolina State University
Total Project Period: 06/15/2015 – 05/31/2019
Location of Project: University of Delaware and University of Tennessee Knoxville
Description: This project tackles numerical errors due to limited arithmetic precision and non-determinism associated with multithreading; the goal is defining methodologies to enable reproducible accuracy of large ensemble simulations on exascale platforms.
Supplement: NSF REU, $16,000, Summer 2016

2015 *Student Support: IEEE Cluster 2015-2016 Conference*
Project Period: 08/01/2015 – 07/31/2018
Location of Project: University of Delaware
Description: This award supports 20 students from American institutions to attend the IEEE Cluster 2015 and Cluster 2016 conferences.

2015 *Comprehensive Study of I/O Performance at the Extreme Scale*
Source of Support: Army Research Office #W911NF-15-2-0033 Total Amount: $297,015, Single PI
Project Period: 06/01/2015 – 05/31/2018
Location of Project: University of Delaware
Description: This award involves studying aspects of I/O performance and in situ analysis for applications relevant to the Army.

2014 *Investigating Massively Scalable I/O-Aware Job Scheduling in Support of Flux*
Source of Support: Lawrence Livermore National Laboratory
Total Amount: $64,118, Single PI
Project Period: 11/05/2014 – 03/31/2016
Location of Project: University of Delaware
Description: This project investigates distinct—yet complementary—techniques to overcome challenges that can preclude I/O-aware schemes from effectively scheduling massively large-scale systems.

2014 *EAGER: Assessment of the Numerical Reproducibility in Large-Scale Scientific Simulations on Multicore Architectures*
Source of Support: National Science Foundation #1446794
Total Amount: $89,999, Single PI
Project Period: 06/15/2014 – 06/01/2016
Location of Project: University of Delaware
Description: This project studies the impact of rounding errors on result reproducibility when concurrent executions burst and workflow determinism vanishes in cutting-edge multicore architectures.

2014 *Evaluating, Analyzing, and Improving the Performance of Data-intensive Applications*
Source of Support: Argonne National Laboratory

Total Amount: $28,657, Single PI
Project Period: 06/09/2014 – 01/15/2015
Location of Project: University of Delaware
Description: This project studies various data-intensive computing frameworks, including MR-MPI and DataMPI, and investigates their performance characteristics, particularly in processing huge amounts of data.

2014 *Student Support: IEEE Cluster 2014 Conference*
Source of Support: National Science Foundation #1441397
Total Amount: $20,000, Single PI
Project Period: 07/01/2014 – 06/01/2016
Location of Project: University of Delaware
Description: This award supports 20 American university students attending the IEEE Cluster 2014 and Cluster 2015 Conferences.

2014 *Becoming the Online Resource Center for Ethics Education in Engineering and Science*
Source of Support: National Science Foundation #1355547
Total Amount: $1,199,918 ($209,239 at University of Delaware), Co-PI. Collaborative research with Thomas Powers, University of Delaware
Project Period: 02/01/2014 – 05/31/2018
Location of Project: University of Delaware
Description: This University of Delaware award subcontract supports the Online Ethics Center for Engineering and Science (OEC), which is an expansion of the existing National Academy of Engineering's (NAE) Online Ethics Center for Engineering and Research (OEC) to include international best practices in ethics for engineers and scientists at the global level.

2013 *SHF: Small: Collaborative Research: Modeling and Analyzing Big Data on Peta- and Exascale Distributed Systems supported by MapReduce Methodologies*
Source of Support: National Science Foundation #1318417
Total Amount: $528,038.00 ($459,000 at UD, PI, with Pietro Cicotti co-PI, SDSC)
Project Period: 09/01/2013 – 08/31/2017
Location of Project: University of Delaware
Description: This project develops transformative analysis methodology to model the properties of large scientific datasets distributed on petascale and exascale systems.
Supplements: NSF REU, $16,000, Summer 2014; NSF REU, $16,000, Summer 2015

2012 *CSR: Small: Collaborative: Pursuing High Performance on Clouds and Other Dynamically Heterogeneous Computing Platforms*
Source of Support: National Science Foundation #1217812
Total Amount: $500,000 ($192,487 at University of Delaware). Co-PI: Arnold L. Rosenberg, Northeastern University
Project Period: 10/01/2012 – 09/30/2015
Location of Project: University of Delaware
Description: This project develops a transformative computing paradigm that enables high-performance computing on computing clouds and many genres of computing grids.
Supplement: NSF REU, $8,000, Fall 2013

2012 *Scalable Aero-Load and Aero-Elasticity Solvers for Massively Parallel Heterogeneous Computing Architectures (Phase II)*
Source of Support: Air Force Office of Scientific Research Small Business Technology Transfer Program (AFOSR SBTT): Highly Scalable Computational-Based Engineering Algorithms for Emerging Parallel Machine Architectures (Topic BT13)
Total Amount: $700,000 ($162,000 at University of Delaware), PI. Collaborative research with Eric Kelmelis, EM Photonics
Project Period: 09/01/2012 – 08/31/2014
Location of Project: University of Delaware
Description: This project supports the development of innovative algorithms for scientific computing, modeling, and simulation in a multi-GPU environment with an emphasis on the parallelization of scientific applications across multiple GPUs.

2012 *Scalable Aero-Load and Aero-Elasticity Solvers for Massively Parallel Heterogeneous Computing Architectures (Phase I)*
Source of Support: Air Force Office of Scientific Research Small Business Technology Transfer Program (AFOSR SBTT) Program: Highly Scalable Computational-Based Engineering Algorithms for Emerging Parallel Machine Architectures (Topic BT13)
Total Amount: $200,000 ($59,997 at University of Delaware), PI. Collaborative research with Eric Kelmelis, EM Photonics
Project Period: 03/01/2012 – 03/01/2013
Location of Project: University of Delaware
Description: This award supports the development of innovative algorithms for scientific computing, modeling, and simulation in a multi-GPU environment, emphasizing parallelizing scientific applications across multiple GPUs.

2010 *Collaborative Research: CDI-Type II: From Data to Knowledge: The Quake-Catcher Network*
Source of Support: National Science Foundation #1027807
Total Amount: $1,841,104 ($101,513 at University of Delaware), PI. Collaborative research with Jesse F. Lawrence, Stanford University; Elizabeth S. Cochran, University of California, Riverside; Richard Allen, University of California, Berkeley; Jack Baker, Stanford University; Tomas Heaton, California Institute of Technology; Deborah Kilb, Scripps Institution of Oceanography
Project Period: 10/01/2010 – 09/30/2014
Location of Project: University of Delaware
Description: This award develops Volunteer Computing cyber-infrastructures to process and analyze large new seismic data sets in near-real time and foster collaboration between thousands of researchers and interested participants worldwide.
Supplement: NSF REU, $7,500, Fall 201; NSF REU, $6,000, Fall 2013

2010 *Collaborative Research: SoCS - ExSciTecH: An Interactive, Easy-to-Use Volunteer Computing System to Explore Science, Technology, and Health*
Source of Support: National Science Foundation #0968350
Total Amount: $683,199 ($308,719 at University of Delaware), Leading PI. Collaborative research with Garry M. Zoppetti, Millersville University, and Johan Cohoon, University of Virginia
Project Period: 09/01/2010 – 08/31/2014

Location of Project: University of Delaware
Description: This award supported the development of interactive methods for engaging new communities as volunteer citizen-scientists and building a mutually beneficial infrastructure for their interaction with professional scientists working on volunteer computing projects in biology and medicine.
Supplement: NSF REU, $7,500, Summer 2012; NSF REU, $7,500, Summer 2011

2010 *Collaborative Research: Accelerated Linear Algebra Solvers for Multi-Core GPU-Based Computing Architecture*
Source of Support: Air Force Office of Scientific Research Small Business Technology Transfer Program (AFOSR SBTT) program - Highly-Scalable Computational-Based Engineering Algorithms for Emerging Parallel Machine Architectures (Topic BT13)
Total Amount: $99,000 ($34,125 at University of Delaware), PI. Collaborative research with Eric Kelmelis, EM Photonics
Project Period: 06/08/2010 – 06/07.2011
Location of Project: University of Delaware
Description: This award supports the development of innovative algorithms for scientific computing, modeling, and simulation in a multi-GPU environment, emphasizing algorithms related to sparse and dense linear algebra problems.

2010 *ExSciTecH: An Interactive, Easy-to-Use Volunteer Computing System to Explore Science, Technology, and Health*
Source of Support: University of Delaware Research Foundation
Total Amount: $35,000, Single PI
Project Period: 06/01/2010 – 05/31/2012
Location of Project: University of Delaware
Description: This seed award aims to build an interactive, easy-to-use VC system to explore science, technology, and health that motivates and facilitates diverse volunteers to donate their resources to VC projects, aiding scientific discovery.
Supplement: University of Delaware Research Foundation REU, $3,500, single PI, Summer 2011

2009 *CDI-Type I: Bridging the Gap Between Next-Generation Hybrid High-Performance Computers and Physics-Based Computational Models for Quantitative Description of Molecular Recognition*
Source of Support: National Science Foundation #0941318
Total Amount: $538,740, co-PI. Collaborative research with Sandeep Patel, University of Delaware
Project Period: 10/01/2009 – 09/30/2013
Location of Project: University of Delaware
Description: This project designs and implements advanced algorithms and middleware packages for polarizable force fields on multicore and GPU systems supported by the MapReduce paradigm.

2009 *Acquisition of a Facility for Computational Approaches to Molecular-Scale Problems*
Source of Support: National Science Foundation #0922657
Total Amount: $451,051, Co-PIs: Douglas Doren, University of Delaware; Sandeep Patel, University of Delaware; and Dionisios Vlachos, University of Delaware

Project Period: 09/15/2009 – 09/14/2012
Location of Project: University of Delaware
Description: This award supports the acquisition of a hybrid-computing cluster with accelerated computing nodes for theoretical and experimental researchers at the University of Delaware to study several problems in chemical sciences.

2009 *Computer-Aided Design for Drugs on Emerging Hybrid High-Performance Computers*
Source of Support: Army Research Office #54723-CS
Total Amount: $306,750, single PI
Project Period: 06/01/2009 – 05/31/2013
Location of Project: University of Delaware
Description: This project developed accurate and efficient protein-ligand docking applications for hybrid computer systems in HPC, including multicore architectures and GPGPUs.
Supplements: ARO Undergraduate Research Program, $3,750, single PI, Summer 2012; ARO High School Apprenticeship Program, $3,000, single PI, Summer 2010

2009 *CRA Mentor, Distributed Mentor Project for Undergraduate Summer Research*
Source of Support: Computing Research Association (CRA)
Total Amount: $6,000, single PI
Total Project Period: 06/01/2009 – 08/31/2000
Location of Project: University of Texas El Paso
Description: This project supports two undergraduate students for ten weeks of research under my supervision. The research targets biological applications and their efficient migration to distributed systems.

2008 *Collaborative Research: Mathematical Models for RNA*
Source of Support: National Science Foundation #0800266
Total Amount: $621,193 ($205,561 at University of Delaware). Co-PIs: Ming-Ying Leung, University of Texas El Paso, and Kyle L. Johnson, University of Texas El Paso
Project Period: 06/01/2008 – 05/31/2012
Location of Project: University of Delaware
Description: This project develops probabilistic models to study the inversion distribution in RNA sequences and to combine the results with the general theory of excursions to maximize the prediction accuracy using an optimal RNA segment length. My group addresses the computational component with grid computing systems.

2008 *CRA Mentor, Distributed Mentor Project for Undergraduate Summer Research*
Source of Support: Computing Research Association (CRA)
Total Amount: $6,000, single PI
Total Project Period: 06/01/2008 – 08/31/2008
Location of Project: University of Texas El Paso
Description: This project supports two undergraduate students for ten weeks of research under my supervision. The research targets biological applications and their efficient migration to distributed systems.

2007 *Computational Prediction of RNA Viral Genome Structures*
Source of Support: National Institute of Health (NIH)

Total Amount: $581,329, co-PI. Collaborative research with Ming-Ying Leung, University of Texas El Paso
Project Period: 09/01/2007 – 08/31/2011
Location of Project: University of Texas El Paso
Description: This project designed and implemented mathematical methods and computation tools for RNA secondary structure prediction in viral genomics.

2006 *S-STEM - SHiPPER: Spreading High-Performance Computing Participation in Undergraduate Education and Research*
Source of Support: National Science Foundation #0631168
Total Amount: $275,856, Leading PI. Collaborative research with Patricia Teller, University of Texas El Paso
Project Period: 10/01/2006 – 01/31/2011
Location of Project: University of Texas El Paso
Description: The award creates and consolidates a community of undergraduate and graduate students pursuing advanced degrees in fields that combine expertise in high-performance computing and other scientific and engineering disciplines.

2006 *RNA Secondary Structure Prediction Using a Grid of Heterogeneous Computers*
Source of Support: Texas Higher Education Coordinating Board, Advanced Research Program #003661-0008-2006
Total Amount: $99,982. Co-PI: Min-Ying Leung, University of Texas El Paso
Total Project Period: 05/15/2006 – 05/14/2008
Location of Project: University of Texas El Paso
Description: This project builds an adaptive grid computing system that, at runtime, identifies and exploits computer resources across the University of Texas at El Paso campus to predict secondary structures of large numbers of RNA segments using various prediction programs.

2006 *Collaborative Research: DAPLDS: a Dynamically Adaptive Protein-Ligand Docking System based on Multiscale Modeling*
Source of Support: National Science Foundation #0506429/#0802650
Total Amount: $1,220,036 ($382,558 at University of Delaware and $273,068 at University of Texas El Paso). Co-PIs: Charles L. Brooks III, The Scripps Research Institute (TSRI) and D.P. Anderson, University of California Berkeley
Project Period: 09/01/2005 – 08/31/2009
Location of Project: University of Delaware and University of Texas El Paso
Description: This project explores the multiscale nature of algorithmic adaptations in protein-ligand docking and developing cyber-infrastructures based on computational methods and models that efficiently accommodate these adaptations.
Supplement: NSF REU, $6,250, Summer 2009

2005 *High-Performance Modular FEM/hp-FEM System (HERMES)*
Source of Support: University of Texas El Paso Seed Funds
Total Amount: $23,400, Co-PI. Collaborative research with Pavel Solin
Total Project Period: 06/01/2005 – 08/31/2005
Location of Project: University of Texas El Paso
Description: This project studies the feasibility of parallelization of FEM algorithms.

2005 *Performance via Autonomicity, Analysis, Virtualization, and Micro-partitioning, and Research in Life Sciences and Bioinformatics*
Source of Support: 2005 IBM Global University Program Shared University Research Award
Total Amount: $600,000, co-PI. Collaborative research with Patricia Teller
Total Project Period: 05/2005
Location of Project: University of Texas El Paso
Description: The award enables the purchase of a supercomputer for biology and bioinformatics applications at the University of Texas El Paso.

PUBLICATIONS

Mentored undergraduate co-authors are indicated with (#) in the author list; mentored graduate co-authors are indicated with (+) in the author list; mentored post-doctoral researchers are indicated with (*) in the author list. My policy is that my students or postdoctoral researchers' names appear first on publications. I use the same author order as in computational chemistry and put the leading author at the end of the authorship list.

BOOK CHAPTERS

[1] Pietro Cicotti, Sarp Oral, Gokcen Kestor, Roberto Gioiosa, Shawn Strande, **Michela Taufer**, James H. Rogers, Hasan Abbasi, Jason Hill, and Laura Carrington. Conquering Big Data with High Performance Computing. In Ritu Arora, editor, *Conquering Big Data with High Performance Computing*, pages 31–60. Springer, 2016.

[2] Trilce Estrada+ and **Michela Taufer**. Challenges in Designing Scheduling Policies in Volunteer Computing. In Christophe Cerin and Gilles Fedak, editors, *Desktop Grid Computing*.

[3] Roger S. Armen, Eric R. May, and **Michela Taufer**. Protein Docking. In David A. Padua, editor, *Encyclopedia of Parallel Computing*, pages 1638–1647. Springer, 2011.

JOURNAL ARTICLES

[1] Naweiluo Zhou*, Giorgio Scorzelli, Jakob Luettgau*, Rahul Reddy Kancharla, Joshua Kane, Robert Wheeler, Brendan Croom, Pania Newell, Valerio Pascucci, and **Michela Taufer**. Orchestration of Materials Science Workflows for Heterogeneous Resources at Large Scale. *International Journal of High-Performance Computing Applications (IJHPCA)*, 3-4(37):260–271, 2023. https://doi.org/10.1177/10943420231167800.

[2] Sanjukta Bhowmick, Parick Bell*, and **Michela Taufer**. A Survey of Graph Comparison Methods with Applications to Nondeterminism in High-Performance Computing. *International Journal of High-Performance Computing Applications (IJHPCA)*, 3-4(37):306–327, 2023. https://doi.org/10.1177/10943420231166610.

[3] Ria Patel#, Brandan Roachell#, Silvina Caíno-Lores*, Charles Ketron#, Jacob Leonard #, Nigel Tan+, Karan Vahi, Duncan Brown, Ewa Deelman, and **Michela Taufer**. Reproducibility of the First Image of a Black Hole in the Galaxy M87 from the Event Horizon Telescope (EHT) Collaboration. *IEEE Computing in Science and Engineering (CiSE)*, 5(24):42–52, 2022. https://doi.org/10.1109/MCSE.2023.3241105.

[4] Tu Mai Anh Do, Loïc Pottier, Rafael Ferreira da Silva, Silvina Caíno-Lores*, **Michela Taufer**, and Ewa Deelman. Performance Assessment of Ensembles of In Situ Workflows under Resource Constraints. *Journal of Concurrency and Computation: Practice and Experience (CCPE)*, 35(40):e7111, 2022. https://doi.org/10.1002/cpe.7111.

[5] Paula Olaya+, Dominic Kennedy#, Ricardo Llamas, Leobardo Valera*, Rodrigo Vargas, Jay Lofstead, and **Michela Taufer**. Building Trust in Earth Science Findings through Data Traceability and Results Explainability. *IEEE Trans. Parallel Distributed Syst. (TPDS)*, 34(2):704–717, 2023. 10.1109/TPDS.2022.3220539.

[6] Stephen Herbein, Tapasya Patki, Dong H. Ahn, Sebastian Mobo#, Clark Hathaway#, Silvina Caino-Lores*, James Corbett, David Domyancic, Thomas R. W. Scogland, Bronis R. de Supinski, and **Michela Taufer**. An Analytical Performance Model of Generalized Multi-Level Scheduling. *International Journal of High-Performance Computing Applications (IJHPCA)*, 36(3):289–306,

2022. 10.1177/10943420211051039.

[7] Michael R. Wyatt II+, Stephen Herbein, Todd Gamblin, and **Michela Taufer**. AI4IO: A Suite of AI-Based Tools for IO-Aware Scheduling. *International Journal of High Performance Computing Applications (IJHPCA)*, 36(3):370–387, 2022. 10.1177/10943420221079765.

[8] Nigel Tan+, Robert Bird, Guangye Chen, Scott V. Luedtke, Brian Albright, and **Michela Taufer**. Analysis of Vector Particle-In-Cell (VPIC) Memory Usage Optimizations on Cutting-Edge Computer Architectures. *Journal of Computational Science*, 60:101566, 2022. 10.1016/j.jocs.2022.101566.

[9] Ricardo M. Llamas, Leobardo Valera*, Paula Olaya+, **Michela Taufer**, and Rodrigo Vargas. Downscaling Satellite Soil Moisture Using a Modular Spatial Inference Framework. *Remote Sensing in Geology, Geomorphology and Hydrology*, 14(13):3137, 2022. 10.3390/rs14133137.

[10] Robert F. Bird, Nigel Tan, Scott V. Luedtke, Stephen Lien Harrell, Michela Taufer, and Brian J. Albright. VPIC 2.0: Next Generation Particle-in-Cell Simulations. *IEEE Trans. Parallel Distributed Syst.*, 33(4):952–963, 2022. 10.1109/TPDS.2021.3084795.

[11] Ariel Keller Rorabaugh*, Silvina Caıno-Lores*, Travis Johnston, and **Michela Taufer**. Building High-throughput Neural Architecture Search Workflows via a Decoupled Fitness Prediction Engine. *IEEE Trans. Parallel Distributed Syst. (TPDS)*, 33(11):2913–2926, 2022. 10.1109/TPDS.2022.3140681.

[12] Ariel Keller Rorabaugh*, Silvina Caıno-Lores*, Travis Johnston, and **Michela Taufer**. High Frequency Accuracy and Loss Data of Random Neural Networks Trained on Image Datasets. *Data in Brief, Elsevier*, 40:107780, 2022. https://doi.org/10.1016/j.dib.2021.107780.

[13] Nicholas C. Mucci, Katarina A. Jones, Mengyi Cao, Michael R. Wyatt, Shane Foye, Sarah J. Kauffman, Gregory R. Richards, **Michela Taufer**, Yoshito Chikaraishi, Shawn A. Steffan, Shawn R. Campagna, Heidi Goodrich-Blair, and Christopher R. Anderton. Apex Predator Nematodes and Meso-Predator Bacteria Consume Their Basal Insect Prey through Discrete Stages of Chemical Transformations. *mSystems*, 0(0):e00312–22, 2022. 10.1128/msystems.00312-22.

[14] Patrick Bell*, Kae Suarez+, Dylan Chapp+, Nigel Tan+, Sanjukta Bhowmick, and **Michela Taufer**. ANACIN-X: A Software Framework for Studying Non-Determinism in MPI Applications. *Software Impacts, Elsevier*, 10:100151, 2021.

[15] Duncan A. Brown, Karan Vahi, **Michela Taufer**, Von Welch, Ewa Deelman, Lorena A. Barba, and George K. Thiruvathukal. Reproducing GW150914: The First Observation of Gravitational Waves From a Binary Black Hole Merger. *IEEE Comput. Sci. Eng. (CiSE)*, 23(2):73–82, 2021. 10.1109/MCSE.2021.3059232.

[16] Tu Mai Anh Do, Loïc Pottier, Silvina Caíno-Lores*, Rafael Ferreira da Silva, Michel A. Cuendet, Harel Weinstein, Trilce Estrada, **Michela Taufer**, and Ewa Deelman. A Lightweight Method for Evaluating *in situ* Workflow Efficiency. *J. Comput. Sci. Elsevier*, 48:101259, 2021. 10.1016/j.jocs.2020.101259.

[17] Hector Carrillo-Cabada, Jeremy Benson, Asghar M. Razavi, Brianna Mulligan, Michel A. Cuendet, Harel Weinstein, **Michela Taufer**, and Trilce Estrada. A Graphic Encoding Method for Quantitative Classification of Protein Structure and Representation of Conformational Changes. *IEEE ACM Trans. Comput. Biol. Bioinform.*, 18(4):1336–1349, 2021. 10.1109/TCBB.2019.2945291.

[18] Dylan Chapp+, Nigel Tan+, Sanjukta Bhowmick, and **Michela Taufer**. Identifying Degree and Sources of Non-Determinism in MPI Applications Via Graph Kernels. *IEEE Trans. Parallel Distributed Syst. (TPDS)*, 32(12):2936–2952, 2021. 10.1109/TPDS.2021.3081530.

[19] Mario Guevara, **Michela Taufer**, and Rodrigo Vargas. Gap-free Global Annual Soil Moisture: 15 km grids for 1991–2018. *Earth System Science Data*, 13(4):1711–1735, 2021. 10.5194/essd-13-1711-2021.

[20] Dong H. Ahn, Ned Bass, Albert Chu, Jim Garlick, Mark Grondona, Stephen Herbein, Helgi I. Ingólfsson, Joseph Koning, Tapasya Patki, Thomas R. W. Scogland, Becky Springmeyer, and **Michela Taufer**. Flux: Overcoming Scheduling Challenges for Exascale Workflows. *Future Gener. Comput. Syst. Elsevier*, 110:202–213, 2020. 10.1016/j.future.2020.04.006.

[21] **Michela Taufer**, Trilce Estrada, and Travis Johnston. A Survey Of Algorithms For Transforming Molecular Dynamics Data Into Metadata For In Situ Analytics Based On Machine Learning Methods. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 378:20190063, 03 2020. 10.1098/rsta.2019.0063.

[22] Ricardo M. Llamas, Mario Guevara, Danny Rorabaugh*, **Michela Taufer**, and Rodrigo Vargas. Spatial Gap-Filling of ESA CCI Satellite-Derived Soil Moisture Based on Geostatistical Techniques and Multiple Regression. *Remote. Sens.*, 12(4):665, 2020. 10.3390/rs12040665.

[23] Dylan Chapp+, Victoria Stodden, and **Michela Taufer**. Building a Vision for Reproducibility in the Cyberinfrastructure Ecosystem: Leveraging Community Efforts. *Supercomput. Front. Innov.*, 7(1):112–129, 2020. 10.14529/jsfi200106.

[24] Tao Gao+, Yanfei Guo, Boyu Zhang, Pietro Cicotti, Yutong Lu, Pavan Balaji, and **Michela Taufer**. Memory-Efficient and Skew-Tolerant MapReduce Over MPI for Supercomputing Systems. *IEEE Trans. Parallel Distributed Syst. (TPDS)*, 31(12):2734–2748, 2020. 10.1109/TPDS.2019.2932066.

[25] Dylan Chapp+, Danny Rorabaugh*, Kento Sato, Dong H. Ahn, and **Michela Taufer**. A Three-Phase Workflow for General And Expressive Representations of Nondeterminism in HPC Applications. *Int. J. High Perform. Comput. Appl.*, 33(6), 2019. 10.1177/1094342019868826.

[26] Robert Searles, Stephen Herbein+, Travis Johnston, **Michela Taufer**, and Sunita Chandrasekaran. Creating a Portable, High-Level Graph Analytics Paradigm for Compute and Data-Intensive Applications. *Int. J. High Perform. Comput. Netw.*, 13(1):105–118, 2019. 10.1504/IJHPCN.2019.097054.

[27] Ewa Deelman, Tom Peterka, Ilkay Altintas, Christopher D. Carothers, Kerstin Kleese van Dam, Kenneth Moreland, Manish Parashar, Lavanya Ramakrishnan, **Michela Taufer**, and Jeffrey S. Vetter. The Future af Scientific Workflows. *Int. J. High Perform. Comput. Appl.*, 32(1):159–175, 2018. 10.1177/1094342017704893.

[28] Dylan Chapp+, Kento Sato, Dong H. Ahn, and **Michela Taufer**. Record-and-Replay Techniques for HPC Systems: A Survey. *Supercomput. Front. Innov.*, 5(1):11–30, 2018. 10.14529/jsfi180102.

[29] **Michela Taufer** and Arnold L. Rosenberg. Scheduling DAG-Based Workflows on Single Cloud Instances: High-Performance and Cost Effectiveness with a Static Scheduler. *Int. J. High Perform. Comput. Appl.*, 31(1):19–31, 2017. 10.1177/1094342015594518.

[30] Travis Johnston*, Boyu Zhang+, Adam Liwo, Silvia Crivelli, and **Michela Taufer**. *In situ* Data Analytics and Indexing of Protein Trajectories. *J. Comput. Chem.*, 38(16):1419–1430, 2017. 10.1002/jcc.24729.

[31] Boyu Zhang+, Trilce Estrada+, Pietro Cicotti, Pavan Balaji, and **Michela Taufer**. Enabling Scalable and Accurate Clustering of Distributed Ligand Geometries On Supercomputers. *Parallel Comput.*, 63:38–60, 2017. 10.1016/j.parco.2017.02.005.

[32] Victoria Stodden, Marcia McNutt, David H. Bailey, Ewa Deelman, Yolanda Gil, Brooks Hanson,

Michael A. Heroux, John P.A. Ioannidis, and **Michela Taufer**. Enhancing Reproducibility for Computational Methods. *Science*, 354(6317):1240–1241, 2016. 10.1126/science.aah6168.

[33] Stephen Herbein+, Sean McDaniel+, Norbert Podhorszki, Jeremy Logan, Scott Klasky, and **Michela Taufer**. Performance Characterization Of Irregular I/O At The Extreme Scale. *Parallel Comput.*, 51:17–36, 2016. 10.1016/j.parco.2015.10.009.

[34] Shuching Ou, Di Cui, Matthew Wezowicz#, **Michela Taufer**, and Sandeep Patel. Free Energetics of Carbon Nanotube Association in Aqueous Inorganic Nai Salt Solutions: Temperature Effects Using All-Atom Molecular Dynamics Simulations. *J. Comput. Chem.*, 36(16):1196–1212, 2015. 10.1002/jcc.23906.

[35] Samuel Schlachter+, Stephen Herbein+, Shuching Ou, Jeremy S. Logan, Sandeep Patel, and **Michela Taufer**. Pursuing Coordinated Trajectory Progression and Efficient Resource Utilization of GPU-Enabled Molecular Dynamics Simulations. *IEEE Des. Test*, 31(1):40–50, 2014. 10.1109/MDAT.2013.2284203.

[36] Jesse Lawrence, Elizabeth Cochran, Angela Chung, Anna Kaiser, Carl Christensen, Richard Allen, Jack Baker, Bill Fry, Thomas Heaton, Debi Kilb, and **Michela Taufer**. Rapid Earthquake Characterization Using MEMS Accelerometers and Volunteer Hosts Following the M 7.2 Darfield, New Zealand, Earthquake. *Bulletin of the Seismological Society of America*, 104, 02 2014. 10.1785/0120120196.

[37] **Michela Taufer**, Narayan Ganesan*, and Sandeep Patel. GPU-Enabled Macromolecular Simulation: Challenges and Opportunities. *Comput. Sci. Eng.*, 15(1):56–65, 2013. 10.1109/MCSE.2012.42.

[38] Kyle Benson#, Samuel Schlachter+, Trilce Estrada+, **Michela Taufer**, Jesse Lawrence, and Elizabeth Cochran. On the Powerful Use of Simulations in the Quake-Catcher Network to Efficiently Position Low-Cost Earthquake Sensors. *Future Gener. Comput. Syst. (FGCS)*, 29(8):2128–2142, 2013. 10.1016/j.future.2013.04.012.

[39] Boyu Zhang+, Daniel T Yehdego, Kyle L Johnson, Ming-Ying Leung, and **Michela Taufer**. Enhancement of Accuracy and Efficiency for RNA Secondary Structure Prediction by Sequence Segmentation and MapReduce. *BMC Structural Biology*, 13(1):1–24, 2013.

[40] Trilce Estrada+, Boyu Zhang+, Pietro Cicotti, Roger S. Armen, and **Michela Taufer**. A Scalable And Accurate Method For Classifying Protein-Ligand Binding Geometries Using A Mapreduce Approach. *Comput. Biol. Medicine*, 42(7):758–771, 2012. 10.1016/j.compbiomed.2012.05.001.

[41] Giorgos Arampatzis, Markos A. Katsoulakis, Petr Plechác, **Michela Taufer**, and Lifan Xu+. Hierarchical Fractional-Step Approximations and Parallel Kinetic Monte Carlo Algorithms. *J. Comput. Phys.*, 231(23):7795–7814, 2012. 10.1016/j.jcp.2012.07.017.

[42] Brad A. Bauer, Joseph E. Davis, **Michela Taufer**, and Sandeep Patel. Molecular Dynamics Simulations of Aqueous Ions at the Liquid-Vapor Interface Accelerated Using Graphics Processors. *J. Comput. Chem.*, 32(3):375–385, 2011. 10.1002/jcc.21578.

[43] Narayan Ganesan*, Brad A. Bauer, Timothy R. Lucas, Sandeep Patel, and **Michela Taufer**. Structural, Dynamic, And Electrostatic Properties of Fully Hydrated DMPC Bilayers from Molecular Dynamics Simulations Accelerated with Graphical Processing Units (GPUs). *J. Comput. Chem.*, 32(14):2958–2973, 2011. 10.1002/jcc.21871.

[44] Obaidur Rahaman, Trilce Estrada+, Douglas J. Doren, **Michela Taufer**, Charles L. Brooks III, and Roger S. Armen. Evaluation of Several Two-Step Scoring Functions Based on Linear Interaction Energy, Effective Ligand Size, and Empirical Pair Potentials for Prediction of

Protein-Ligand Binding Geometry and Free Energy. *J. Chem. Inf. Model.*, 51(9):2047–2065, 2011. 10.1021/ci1003009.

[45] John J Rosskopf, John H Upton III, Lizette Rodarte, Tammy A Romero, Ming-Ying Leung, **Michela Taufer**, and Kyle L Johnson. A 3 Terminal Stem–loop structure in Nodamura Virus RNA2 Forms an Essential cis-acting Signal for RNA Replication. *Virus research*, 150(1-2):12–21, 2010.

[46] Trilce Estrada+, **Michela Taufer**, and David P. Anderson. Performance Prediction and Analysis of BOINC Projects: An Empirical Study with EmBOINC. *J. Grid Comput.*, 7(4):537–554, 2009. 10.1007/s10723-009-9126-3.

[47] **Michela Taufer**, Abel Licon+, Roberto Araiza, David Mireles#, F. H. D. van Batenburg, Alexander P. Gultyaev, and Ming-Ying Leung. PseudoBase++: an Extension of PseudoBase for Easy Searching, Formatting and Visualization of Pseudoknots. *Nucleic Acids Res.*, 37(Database-Issue):127–135, 2009. 10.1093/nar/gkn806.

[48] **Michela Taufer**, Roger S. Armen, Jianhan Chen, Patricia J. Teller, and Charles L. Brooks. Computational Multiscale Modeling in Protein–ligand Docking. *IEEE Engineering in Medicine and Biology Magazine*, 28(2):58–69, 2009. 10.1109/MEMB.2009.931789.

[49] **Michela Taufer**, Ming-Ying Leung, Thamar Solorio, Abel Licon+, David Mireles#, Roberto Araiza, and Kyle L. Johnson. RNAVLab: A Virtual Laboratory for Studying RNA Secondary Structures based on Grid Computing Technology. *Parallel Comput.*, 34(11):661–680, 2008. 10.1016/j.parco.2008.08.002.

[50] Trilce Estrada+, Olac Fuentes, and **Michela Taufer**. A Distributed Evolutionary Method to Design Scheduling Policies for Volunteer Computing. *SIGMETRICS Perform. Evaluation Rev.*, 36(3):40–49, 2008. 10.1145/1481506.1481515.

[51] **Michela Taufer**, Chahm An, Andreas Kerstens, and Charles L. Brooks III. Predictor@Home: A "Protein Structure Prediction Supercomputer" Based on Global Computing. *IEEE Trans. Parallel Distributed Syst. (TPDS)*, 17(8):786–796, 2006. 10.1109/TPDS.2006.110.

[52] **Michela Taufer**, Michael Crowley, Daniel J. Price, Andrew A. Chien, and Charles L. Brooks III. Study of a Highly Accurate and Fast Protein-Ligand Docking Method based on Molecular Dynamics. *Concurr. Comput. Pract. Exp.*, 17(14):1627–1641, 2005. 10.1002/cpe.949.

[53] Kim K. Baldridge, Jerry P. Greenberg, Wibke Sudholt, Stephen Mock, Ilkay Altintas, Céline Amoreira, Yohann Potier, Adam Birnbaum, Karan Bhatia, and **Michela Taufer**. The Computational Chemistry Prototyping Environment. *Proc. IEEE*, 93(3):510–521, 2005. 10.1109/JPROC.2004.842747.

[54] Pietro Cicotti, **Michela Taufer**, and Andrew A. Chien. DGMonitor: A Performance Monitoring Tool for Sandbox-Based Desktop Grid Platforms. *J. Supercomput.*, 34(2):113–133, 2005. 10.1007/s11227-005-2336-y.

REFEREED CONFERENCES, SYMPOSIUMS, AND WORKSHOPS

[1] Harshita Sahni, Hector Carrillo-Cabada, Ekaterina Kots, Silvina Caino-Lores*, Jack Marquez*, Ewa Deelman, Michel Cuendet Harel Weinstein, **Michela Taufer**, and Trilce Estrada. Online Boosted Gaussian Learners for in-situ Detection and Characterization of Protein Folding States in Molecular Dynamics Simulations. In *Proceedings of the 19th IEEE International Conference on e-Science (eScience)*, pages 1–10, Limassol, Cyprus, October 2023. IEEE Computer Society. *(Acceptance Rate: 40/81, 50%).*

[2] Nigel Tan+, Bogdan Nicolae, Jakob Luettgau*, Jack Marquez*, Keita Teranishi, Nicolas Morales, Sanjukta Bhowmick, **Michela Taufer**, and Franck Cappello. Scalable Checkpointing of Applications with Sparsely Updated Data. In *Proceedings of the 52nd International Conference on Parallel Processing (ICPP)*, pages 1–10, Salt Lake City, UT, USA, August 2023. ACM. *(Acceptance Rate: 29%).*

[3] Georgia Channing*, Ria Patel#, Paula Olaya+, Ariel Rorabaugh*, Osamu Miyashita, Silvina Caino-Lores*, Catherine Schuman, Florence Tama, and **Michela Taufer**. Composable Workflow for Accelerating Neural Architecture Search Using In Situ Analytics for Protein Characterization. In *Proceedings of the 52nd International Conference on Parallel Processing (ICPP)*, pages 1–10, Salt Lake City, UT, USA, August 2023. ACM. *(Acceptance Rate: 29%).*

[4] Paula Olaya+, Jakob Luettgau*, Camila Roa+, Ricardo Llamas, Rodrigo Vargas, Sophia Wen, I-Hsin Chung, Seetharami Seelam, Yoonho Park, Jay Lofstead, and **Michela Taufer**. Enabling Scalability in the Cloud for Scientific Workflows: An Earth Science Use Case. In *Proceedings of IEEE CLOUD*, pages 1–10, Chicago, IL, USA, June 2023. IEEE Computer Society. *(Acceptance Rate: 21.6%).*

[5] Silvina Caino-Lores*, Michel A. Cuendet, Jack Marquez* Trilce Estrada, Ewa Deelman, Harel Weinstein, and **Michela Taufer**. Runtime Steering of Molecular Dynamics Simulations Through In Situ Analysis and Annotation of Collective Variables. In *Proceedings of Platform for Advanced Scientific Computing (PASC) Conference*, pages 1–10, Davos, Switzerland, June 2023. ACM.

[6] Camila Roa+, Paula Olaya+, Ricardo Llamas, Rodrigo Vargas, and **Michela Taufer**. GEOtiled: A Scalable Workflow for Generating Large Datasets of High-Resolution Terrain Parameters. In *Proceedings of the 32nd International ACM Symposium on High-Performance Parallel and Distributed Computing (HPDC)*, pages 1–2, Orlando, Florida, USA, June 2023. ACM. *(Short paper).*

[7] Jakob Luettgau*, Heberth Martinez*, Glenn Tarcea, Giorgio Scorzelli, Valerio Pascucci, and **Michela Taufer**. Studying Latency and Throughput Constraints for Geo-Distributed Data in the National Science Data Fabric. In *Proceedings of the 32nd International ACM Symposium on High-Performance Parallel and Distributed Computing (HPDC)*, pages 1–2, Orlando, Florida, USA, June 2023. ACM. *(Short paper).*

[8] Stephanie Brink, Michael McKinsey, David Boehme, W. Daryl Hawkins, Connor Scully-Allison, Ian Lumsden+, Treece Burgess+, Vanessa Lama+, Katherine E. Isaacs, Jakob Luettgau*, **Michela Taufer**, and Olga Pearce. Thicket: Seeing the Performance Experiment Forest for the Individual Run Trees. In *Proceedings of the 32nd International ACM Symposium on High-Performance Parallel and Distributed Computing (HPDC)*, pages 1–10, Orlando, Florida, USA, June 2023. ACM. *(Acceptance Rate: 21/101, 20.7%).*

[9] Jakob Luettgau*, Giorgio Scorzelli, Valerio Pascucci, and **Michela Taufer**. Development of Large-Scale Scientific Cyberinfrastructure and the Growing Opportunity to Democratize Access to Platforms and Data. In *Proceedings of the 25TH International Conference On Human-Computer Interaction (HCII)*, Copenhagen, Denmark, July 2023. Springer.

[10] Jakob Luettgau*, Giorgio Scorzelli, Glenn Tarcea, Christine R. Kirkpatrick, Valerio Pascucci, and **Michela Taufer**. NSDF-Catalog: Lightweight Indexing Service for Democratizing Data Delivery. In *Proceedings of the 15th IEEE/ACM International Conference on Utility and Cloud Computing (UCC)*, pages 1–10, Seattle, WA, USA, December 2022. IEEE Computer Society. *(Acceptance Rate: 13/49, 26.5%).*

[11] Tu Mai Anh Do, Loic Pottier, Rafael Ferreira da Silva, Frederic Suter, Silvina Caíno-Lores*, **Michela Taufer**, and Ewa Deelman. Co-Scheduling Ensembles of In Situ Workflows. In *Pro-*

ceedings of the 17th Workshop on Workflows in Support of Large-Scale Science (WORKS), pages 1–10, Dallas, TX, USA, November 2022. IEEE Computer Society.

[12] Ian Lumsden+, Jakob Luettgau*, Vanessa Lama, Connor Scully-Allison, Stephanie Brink, Katherine E. Isaacs, Olga Pearce, and **Michela Taufer**. Enabling Call Path Querying in Hatchet to Identify Performance Bottlenecks in Scientific Applications. In *Proceedings of the 18th IEEE International Conference on e-Science (eScience)*, pages 1–10, Salt Lake City, Utah, USA, October 2022. IEEE Computer Society. *(Acceptance Rate: 37/59, 62.7%).*

[13] Paula Olaya+, Silvina Caino-Lores*, Vanessa Lama+, Ria Patel#, Ariel Rorabaugh*, Osamu Miyashita, Florence Tama, and **Michela Taufer**. Identifying Structural Properties of Proteins from X-ray Free Electron Laser Diffraction Patterns. In *Proceedings of the 18th IEEE International Conference on e-Science (eScience)*, pages 1–10, Salt Lake City, Utah, USA, October 2022. IEEE Computer Society. *Best paper candidate (Acceptance Rate: 37/59, 62.7%).*

[14] Jakob Luettgau*, Silvina Caino-Lores*, Kae Suarez+, Dong H. Ahn, Stephen Herbein, and **Michela Taufer**. Reproducing and Extending Analytical Performance Models of Generalized Hierarchical Scheduling. In *Proceedings of the Second Workshop on Reproducible Workflows, Data, and Security (ReWorDS)*, pages 1–6, Salt Lake City, Utah, USA, October 2022. IEEE Computer Society.

[15] Dominic Kennedy#, Paula Olaya+, Jay Lofstead, Rodrigo Vargas, and **Michela Taufer**. Augmenting Singularity to Generate Fine-grained Workflows, Record Trails, and Data Provenance. In *Proceedings of the 18th IEEE International Conference on e-Science (eScience)*, pages 1–2, Salt Lake City, Utah, USA, October 2022. IEEE Computer Society. *(Short paper).*

[16] Yeom Jae-Seung, Dong H. Ahn, Ian Lumsden+, Jakob Luettgau*, Silvina Caino-Lores*, and **Michela Taufer**. Ubique: A New Model for Untangling Inter-task Data Dependence in Complex HPC Workflows. In *Proceedings of the 18th IEEE International Conference on e-Science (eScience)*, pages 1–2, Salt Lake City, Utah, USA, October 2022. IEEE Computer Society. *(Short paper).*

[17] Glenn Tarcea, Brian Puchala, Tracy Berman, Giorgio Scorzelli, Valerio Pascucci, **Michela Taufer**, and John Allison. The Materials Commons Data Repository. In *Proceedings of the 18th IEEE International Conference on e-Science (eScience)*, pages 1–2, Salt Lake City, Utah, USA, October 2022. IEEE Computer Society. *(Short paper).*

[18] Ria Patel, Ariel Rorabaugh Keller*, Paula Olaya+, Silvina Caino-Lores*, Georgia Channing*, Catherine Schuman, Osamu Miyashita, Florence Tama, and **Michela Taufer**. A Methodology to Generate Efficient Neural Networks for Classification of Scientific Datasets. In *Proceedings of the 18th IEEE International Conference on e-Science (eScience)*, pages 1–2, Salt Lake City, Utah, USA, October 2022. IEEE Computer Society. *(Short paper).*

[19] Paula Olaya+, Jakob Luettgau*, Naweiluo Zhou*, Giorgio Scorzelli, Jay Lofstead, Valerio Pascucci, and **Michela Taufer**. NSDF-FUSE: A Testbed for Studying Object Storage via FUSE File Systems. In *Proceedings of the 31st International ACM Symposium on High-Performance Parallel and Distributed Computing (HPDC)*, pages 1–2, Minneapolis, Minnesota, June 2022. ACM. *(Short paper).*

[20] Jakob Luettgau, Paula Olaya+, Naweiluo Zhou, Giorgio Scorzelli, Valerio Pascucci, and **Michela Taufer**. NSDF-Cloud: Enabling Ad-Hoc Compute Clusters Across Academic and Commercial Clouds. In *Proceedings of the 31st International ACM Symposium on High-Performance Parallel and Distributed Computing (HPDC)*, pages 1–2, Minneapolis, Minnesota, June 2022. ACM. *(Short paper).*

[21] **Michela Taufer**, Ewa Deelman, Rafael Ferreira da Silva, Trilce Estrada, Mary Hall, and Miron Livny. A Roadmap to Robust Science for High-throughput Applications: The Developers' Perspective. In *Proceedings of the IEEE Cluster Conference (CLUSTER)*, pages 1–2, Portland, Oregon, September 2021. ACM. *(Short paper).*

[22] Ross Ketron#, Jacob Leonard#, Brandan Roachell#, Ria Patel#, Rebecca White, Silvina Caino-Lores, Nigel Tan+, Patrick Miles, Karan Vahi, Ewa Deelam, Duncan A. Brown, and **Michela Taufer**. A Case Study in Scientific Reproducibility from the Event Horizon Telescope (EHT). In *Proceedings of the 20th IEEE International Conference on eScience*, pages 1–2, Innsbruck, Austria, September 2021. IEEE Computer Society. *(Short paper).*

[23] **Michela Taufer**, Ewa Deelman, Rafael Ferreira da Silva, Trilce Estrada, and Mary Hall. A Roadmap to Robust Science for High-throughput Applications: The Scientists' Perspective. In *Proceedings of the 20th IEEE International Conference on eScience*, pages 1–2, Innsbruck, Austria, September 2021. IEEE Computer Society. *(Short paper).*

[24] Tu Mai Anh Do, Loïc Pottier, Rafael Ferreira da Silva, Silvina Caino-Lores+, **Michela Taufer**, and Ewa Deelman. Assessing Resource Provisioning and Allocation of Ensembles of In Situ Workflows. In *Proceedings of the Fourteen International Workshop on Parallel Programming Models and Systems Software for High-End Computing (P2S2)*, Chicago, Illinois, USA, August 2021. IEEE Computer Society.

[25] Stephanie Brink, Ian Lumsden#, Connor Scully-Allison, Katy Williams, Olga Pearce, Todd Gamblin, **Michela Taufer**, Katherine E. Isaacs, and Abhinav Bhatele. Usability and Performance Improvements in Hatchet. In *Proceedings of the IEEE/ACM International Workshop on HPC User Support Tools and Workshop on Programming and Performance Visualization Tools (HUST/ProTools@SC)*, pages 49–58, Atlanta, GA, USA, November 2020. IEEE Computer Society.

[26] Nigel Tan+, Robert F. Bird, Guangye Chen, and **Michela Taufer**. Optimize Memory Usage in Vector Particle-In-Cell (VPIC) to Break the 10 Trillion Particle Barrier in Plasma Simulations. In *Proceedings of the 21st International Conference on Computational Science (ICCS)*, pages 452–465, Krakow, Poland, June 16-18 2021. Springer. *(Best Track Paper Award).*

[27] Tu Mai Anh Do, Loïc Pottier, Stephen Thomas, Rafael Ferreira da Silva, Michel A. Cuendet, Harel Weinstein, Trilce Estrada, **Michela Taufer**, and Ewa Deelman. A Novel Metric to Evaluate In Situ Workflows. In *Proceedings of the 20th International Conference on Computational Science (ICCS)*, pages 538–553, Amsterdam, The Netherlands, June 3-5 2020. Springer.

[28] Michael R. Wyatt II+, Stephen Herbein, Kathleen Shoga, Todd Gamblin, and **Michela Taufer**. CanarIO: Sounding the Alarm on IO-Related Performance Degradation. In *Proceedings of the 2020 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, pages 73–83, New Orleans, LA, USA, May 19-22 2020. IEEE Computer Society. *(Acceptance Rate: 45/110, 24.7%).*

[29] Danny Rorabaugh*, Mario Guevara, Ricardo M. Llamas, Joy Kitson#, Rodrigo Vargas, and **Michela Taufer**. SOMOSPIE: A Modular SOil MOisture SPatial Inference Engine Based on Data-Driven Decisions. In *Proceedings of the 15th International Conference on eScience*, pages 1–10, San Diego, CA, USA, September 24-27 2019. IEEE Computer Society.

[30] **Michela Taufer**, Ewa Deelman, Stephen Thomas*, Michael R. Wyatt II+, Tu Mai Anh Do, Loïc Pottier, Rafael Ferreira da Silva, Harel Weinstein, Michel A. Cuendet, and Trilce Estrada. Characterizing In Situ and In Transit Analytics of Molecular Dynamics Simulations for Next-Generation Supercomputers. In *Proceedings of the 15th International Conference on eScience*,

pages 188–198, San Diego, CA, USA, September 24 - 27 2019.

[31] Dylan Chapp+, Danny Rorabaugh*, Duncan A. Brown, Ewa Deelman, Karan Vahi, Von Welch, and **Michela Taufer**. Applicability Study of the PRIMAD Model to LIGO Gravitational Wave Search Workflows. In *Proceedings of the 2nd International Workshop on Practical Reproducible Evaluation of Computer Systems (P-RECS), co-located with the ACM HPDC Conference*, pages 1–6, Phoenix, AZ, USA, June 2019. ACM.

[32] Ewa Deelman, Victoria Stodden, **Michela Taufer**, and Von Welch. Initial Thoughts on Cybersecurity And Reproducibility. In *Proceedings of the 2nd International Workshop on Practical Reproducible Evaluation of Computer Systems (P-RECS), co-located with the ACM HPDC Conference*, pages 13–15, Phoenix, AZ, USA, June 2019. ACM.

[33] Joshua Hoke Davis#, Tao Gao+, Sunita Chandrasekaran, Heike Jagode, Anthony Danalis, Jack J. Dongarra, Pavan Balaji, and **Michela Taufer**. Characterization of Power Usage and Performance in Data-Intensive Applications Using MapReduce over MPI. In *Proceedings of the International Conference on Parallel Computing (PARCO)*, volume 36 of *Advances in Parallel Computing*, pages 287–298, Prague, Czech Republic, September 10-13 2019. IOS Press.

[34] Trilce Estrada, Jeremy Benson, Hector Carrillo-Cabada, Asghar M. Razavi, Michel A. Cuendet, Harel Weinstein, Ewa Deelman, and **Michela Taufer**. Graphic Encoding of Macromolecules for Efficient High-Throughput Analysis. In *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics (BCB)*, pages 315–324, Washington DC, USA, August 29 - September 1 2018. ACM. *(Acceptance Rate: 27%).*

[35] Sean McDaniel-Gray+, David L. Boothe, Alfred B. Yu, Dale R. Shires, and **Michela Taufer**. Leveraging In Situ Data Analysis to Enable Computational Steering of Brain's Neocortex Simulations with GENESIS. In *Proceedings of the 20th IEEE International Conference on High Performance Computing and Communications; 16th IEEE International Conference on Smart City; 4th IEEE International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, pages 873–880, Exeter, United Kingdom, June 28-30 2018. IEEE Computer Society.

[36] Tao Gao+, Yanfei Guo, Boyu Zhang+, Pietro Cicotti, Yutong Lu, Pavan Balaji, and **Michela Taufer**. On the Power of Combiner Optimizations in MapReduce Over MPI Workflows. In *Proceedings of the 24th IEEE International Conference on Parallel and Distributed Systems (ICPADS)*, pages 441–448, Singapore, December 11-13 2018. IEEE Computer Society. *(Acceptance Rate: 257/97, 37.7%).*

[37] Xinyu Chen, Jeremy Benson, Matt Peterson, **Michela Taufer**, and Trilce Estrada. KeyBin2: Distributed Clustering for Scalable and In-Situ Analysis. In *Proceedings of the 47th International Conference on Parallel Processing (ICPP)*, pages 34:1–34:10, Eugene, OR, USA, August 13-16 2018. ACM. *(Acceptance Rate: 305/99, 32.5%).*

[38] Michael R. Wyatt II+, Stephen Herbein+, Todd Gamblin, Adam Moody, Dong H. Ahn, and **Michela Taufer**. PRIONN: Predicting Runtime and IO using Neural Networks. In *Proceedings of the 47th International Conference on Parallel Processing, (ICPP)*, pages 46:1–46:12, Eugene, OR, USA, August 13-16 2018. ACM. *(Acceptance Rate: 305/99, 32.5%).*

[39] Don H. Ahn, Ned Bass, Albert Chu, Jim Garlick, Mark Grondona, Stephen Herbein, Joseph Koning, Topasya Patki, Thomas R. W. Scogland, Becky Springmeyer, and **Michela Taufer**. Flux: Overcoming Scheduling Challenges for Exascale Workflows. In *Proceedings of the 13th Workshop on Workflows in Support of Large-Scale Science (WORKS)*, pages 1–8, Dallas, TX, USA, November 2018. IEEE Computer Society. *(Acceptance Rate: 8/19, 42%).*

[40] Thomas Kitson#, Paula Olaya#, Elizabeth Racca#, Michael R. Wyatt II+, Mario Guevara,

Rodrigo Vargas, and **Michela Taufer**. Data Analytics for Modeling Soil Moisture Patterns across United States Ecoclimatic Domains. In *Proceedings of the 2017 IEEE International Conference on Big Data (BigData)*, pages 4768–4770, Boston, MA, USA, December 11-14 2017. IEEE Computer Society. *(Short paper).*

[41] Tao Gao+, Yanfei Guo, Yanjie Wei, Bingqiang Wang, Yutong Lu, Pietro Cicotti, Pavan Balaji, and **Michela Taufer**. Bloomfish: A Highly Scalable Distributed K-mer Counting Framework. In *Proceedings of the 23rd IEEE International Conference on Parallel and Distributed Systems (ICPADS)*, pages 170–179, Shenzhen, China, December 15-17 2017. IEEE Computer Society. *(Acceptance Rate: 271/89, 32.8%).*

[42] Tao Gao+, Yanfei Guo, Boyu Zhang+, Pietro Cicotti, Yutong Lu, Pavan Balaji, and **Michela Taufer**. Mimir: Memory-Efficient and Scalable MapReduce for Large Supercomputing Systems. In *Proceedings of the 2017 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, pages 1098–1108, Orlando, FL, USA, May 29 – Jun 2 2017. IEEE Computer Society. *(Acceptance Rate: 116/616, 18.8%).*

[43] Michael R. Wyatt II+, Travis Johnston*, Mia Papas, and **Michela Taufer**. Development of a Scalable Method for Creating Food Groups Using the NHANES Dataset and MapReduce. In *Proceedings of the 7th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics (BCB)*, pages 118–127, Seattle, WA, USA, October 2-4 2016. ACM. *(Acceptance Rate: 47/112, 42%).*

[44] Ryan McKenna#, Stephen Herbein+, Adam Moody, Todd Gamblin, and **Michela Taufer**. Machine Learning Predictions of Runtime and IO Traffic on High-End Clusters. In *Proceedings of the 2016 IEEE International Conference on Cluster Computing (CLUSTER)*, pages 255–258, Taipei, Taiwan, September 12 – 16 2016. IEEE Computer Society. *(Acceptance Rate: 35% for full and short papers).*

[45] Stephen Herbein+, Dong H. Ahn, Don Lipari, Thomas R. W. Scogland, Marc Stearman, Mark Grondona, Jim Garlick, Becky Springmeyer, and **Michela Taufer**. Scalable I/O-Aware Job Scheduling for Burst Buffer Enabled HPC Clusters. In *Proceedings of the 25th ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC)*, pages 69–80, Kyoto, Japan, May 30 – Jun 4 2016. ACM. *(Acceptance Rate: 20/129, 15.5%).*

[46] Sean McDaniel+, David L. Boothe, Joshua C. Crone, Song Jun Park, Dale R. Shires, Alfred B. Yu, and **Michela Taufer**. Study of Neocortex Simulations with GENESIS on High Performance Computing Resources. In *Proceedings of the 22nd IEEE International Conference on Parallel and Distributed Systems (ICPADS)*, pages 924–931, Wuhan, China, December 14–16 2016. IEEE Computer Society. *(Acceptance Rate: 82/230, 35.6%).*

[47] Travis Johnston*, Connor Zanin#, and **Michela Taufer**. HYPPO: A Hybrid, Piecewise Polynomial Modeling Technique for Non-Smooth Surfaces. In *Proceedings of the 28th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD)*, pages 26–33, Los Angeles, CA, USA, October 26–28 2016. IEEE Computer Society. *(Acceptance Rate: 27/77, 35% – One of four Best Paper Candidates).*

[48] Jeremy Benson, Trilce Estrada, Arnold L. Rosenberg, and **Michela Taufer**. Scheduling Matters: Area-Oriented Heuristic for Resource Management. In *Proceedings of the 28th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD)*, pages 214–221, Los Angeles, CA, USA, October 26–28 2016. IEEE Computer Society. *(Acceptance Rate: 27/77, 35%).*

[49] Stephen Herbein+, Ayush Dusia, Aaron Landwehr, Sean McDaniel+, José Monsalve Diaz, Yang

Yang, Seetharami R. Seelam, and **Michela Taufer**. Resource Management for Running HPC Applications in Container Clouds. In *Proceedings of the 31st High Performance Computing - International Supercomputing Conference (ISC)*, volume 9697 of *Lecture Notes in Computer Science*, pages 261–278, Frankfurt, Germany, June 19-23 2016. Springer. *(Acceptance Rate: 25/60, 40%).*

[50] Boyu Zhang+, Trilce Estrada+, Pietro Cicotti, Pavan Balaji, and **Michela Taufer**. Accurate Scoring of Drug Conformations at the Extreme Scale. In *Proceedings of the 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)*, pages 817–822, Shenzhen, China, May 4–7 2015. IEEE Computer Society. *(Award Winners. Acceptance Rate: 5/15, 33%).*

[51] Dylan Chapp+, Travis Johnston*, and **Michela Taufer**. On the Need for Reproducible Numerical Accuracy through Intelligent Runtime Selection of Reduction Algorithms at the Extreme Scale. In *Proceedings of the 2015 IEEE International Conference on Cluster Computing (CLUSTER)*, pages 166–175, Chicago, IL, USA, September 8 – 11 2015. IEEE Computer Society. *(Acceptance Rate: 38/157, 24%).*

[52] Sean McDaniel+, Stephen Herbein+, and **Michela Taufer**. A Two-Tiered Approach to I/O Quality of Service in Docker Containers. In *Proceedings of the 2015 IEEE International Conference on Cluster Computing (CLUSTER)*, pages 490–491, Chicago, IL, USA, 2015. IEEE Computer Society.

[53] Ayush Dusia, Yang Yang, and **Michela Taufer**. Network Quality of Service in Docker Containers. In *Proceedings of the 2015 IEEE International Conference on Cluster Computing (CLUSTER)*, pages 527–528, Chicago, IL, USA, 2015. IEEE Computer Society.

[54] José Monsalve Diaz, Aaron Landwehr, and **Michela Taufer**. Dynamic CPU Resource Allocation in Containerized Cloud Environments. In *Proceedings of the 2015 IEEE International Conference on Cluster Computing (CLUSTER)*, pages 535–536, Chicago, IL, USA, 2015. IEEE Computer Society.

[55] Ryan McKenna#, Vivek K. Pallipuram*, Rodrigo Vargas, and **Michela Taufer**. From HPC Performance to Climate Modeling: Transforming Methods for HPC Predictions into Models of Extreme Climate Conditions. In *Proceedings of the 11th IEEE International Conference on eScience (eScience)*, pages 108–117, Munich, Germany, August 31 – September 4 2015. IEEE Computer Society.

[56] Travis Johnston*, Mohammad Alsulmi+, Pietro Cicotti, and **Michela Taufer**. Performance Tuning of MapReduce Jobs Using Surrogate-based Modeling. In *Proceedings of the Proceedings of the International Conference on Computational Science (ICCS)*, volume 51 of *Procedia Computer Science*, pages 49–59, Reykjavík, Iceland, June 1-6 2015. Elsevier. *(Acceptance Rate: 26%).*

[57] Trilce Estrada, Michael R. Wyatt II, and **Michela Taufer**. A Genetic Programming Approach to Design Resource Allocation Policies for Heterogeneous Workflows in the Cloud. In *Proceedings of the 21st IEEE International Conference on Parallel and Distributed Systems (ICPADS)*, pages 372–379, Melbourne, Australia, December 14-16 2015. IEEE Computer Society. *(Acceptance Rate: 82/230, 35.6%).*

[58] Vivek K. Pallipuram#, Trilce Estrada, and **Michela Taufer**. A Testing Engine for High-Performance and Cost-Effective Workflow Execution in the Cloud. In *Proceedings of the 44th International Conference on Parallel Processing (ICPP)*, pages 849–858, Beijing, China, September 1-4 2015. IEEE Computer Society. *(Acceptance Rate: 99/305, 32.5%).*

[59] Vivek K. Pallipuram*, Jeffrey DiMarco+, and **Michela Taufer**. Applying frequency Analysis

Techniques to DAG-based Workflows to Benchmark and Predict Resource Behavior on Non-dedicated Clusters. In *Proceedings of the 2014 IEEE International Conference on Cluster Computing (CLUSTER)*, pages 29–37, Madrid, Spain, September 22-26 2014. IEEE Computer Society. *(Acceptance Rate: 29/122, 23.8% - One of four Best Paper Candidates).*

[60] Feng Raoking, Joanne McGrath Cohoon, Kathryn Cooke, **Michela Taufer**, and Trilce Estrada#. Gender and Volunteer Computing: A Survey Study. In *Proceedings of the IEEE Frontiers in Education Conference (FIE)*, pages 1–5, Madrid, Spain, October 22-25 2014. IEEE Computer Society.

[61] Marcos Portnoi+, Samuel Schlachter+, and **Michela Taufer**. Study of the Network Impact on Earthquake Early Warning in the Quake-catcher Network Project. In *Proceedings of the International Conference on Computational Science (ICCS)*, volume 29 of *Procedia Computer Science*, pages 453–464, Cairns, Queensland, Australia, June 10 – 12 2014. Elsevier. *(Acceptance Rate: 197/65, 33%).*

[62] Michael Matheny#, Stephen Herbein#, Norbert Podhorszki, Scott Klasky, and **Michela Taufer**. Using Surrogate-based Modeling to Predict Optimal I/O Parameters of Applications at the Extreme Scale. In *Proceedings of the 20th IEEE International Conference on Parallel and Distributed Systems, (ICPADS)*, pages 568–575. IEEE Computer Society, December 16-19 2014. *(Acceptance Rate: 96/322, 29.8%).*

[63] Stephen Herbein#, Scott Klasky, and **Michela Taufer**. Benchmarking the Performance of Scientific Applications with Irregular I/O at the Extreme Scale. In *Proceedings of the 43rd International Conference on Parallel Processing Workshops (ICPPW)*, pages 292–301, Minneapolis, MN, USA, September 9-12 2014. IEEE Computer Society.

[64] Boyu Zhang+, Trilce Estrada+, Pietro Cicotti, and **Michela Taufer**. Enabling In-Situ Data Analysis for Large Protein-Folding Trajectory Datasets. In *Proceedings of the 2014 IEEE 28th International Parallel and Distributed Processing Symposium (IPDPS)*, pages 221–230, Phoenix, AZ, USA, May 19-23 2014. IEEE Computer Society. *(Acceptance Rate: 114/541, 21%).*

[65] Bahman Javadi, Boyu Zhang+, and **Michela Taufer**. Bandwidth Modeling in Large Distributed Systems for Big Data Applications. In *Proceedings of the 15th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT)*, pages 21–27, Hong Kong, China, December 9-11 2014. IEEE Computer Society.

[66] Stephen Herbein#, Michael Matheny#, Matthew Wezowicz#, Jaron T. Krogel, Jeremy Logan, J. Kim, Scott Klasky, and **Michela Taufer**. Performance Impact of I/O on QMCPack Simulations at the Petascale and Beyond. In *Proceedings of the 16th IEEE International Conference on Computational Science and Engineering (CSE)*, pages 92–99, Sydney, Australia, December 2013. IEEE Computer Society.

[67] Boyu Zhang+, Trilce Estrada+, Pietro Cicotti, and **Michela Taufer**. On Efficiently Capturing Scientific Properties in Distributed Big Data without Moving the Data: A Case Study in Distributed Structural Biology Using MapReduce. In *Proceedings of the 16th IEEE International Conference on Computational Science and Engineering (CSE)*, pages 117–124, Sydney, Australia, December 2013. IEEE Computer Society.

[68] Samuel Schlachter+, Stephen Herbein#, **Michela Taufer**, Shuching Ou, Sandeep Patel, and Jeremy S. Logan. Efficient SDS Simulations on Multi-GPU Nodes of XSEDE High-End Clusters. In *Proceedings of the 9th IEEE International Conference on eScience*, pages 116–123, Beijing, China, October 22-25, 2013 2013. IEEE Computer Society.

[69] Trilce Estrada+, Kathleen L. Pusecker, Manuel R. Torres, Joanne McGrath Cohoon, and

Michela Taufer. Benchmarking Gender Differences in Volunteer Computing Projects. In *Proceedings of the 9th IEEE International Conference on eScience (eScience)*, pages 342–349, Beijing, China, October 22-25 2013. IEEE Computer Society.

[70] Daniel T. Yehdego+, Boyu Zhang+, Vikram Kumar Reddy Kodimala, Kyle L. Johnson, **Michela Taufer**, and Ming-Ying Leung. Secondary Structure Predictions for Long RNA Sequences Based on Inversion Excursions and MapReduce. In *Proceedings of the 2013 IEEE International Symposium on Parallel & Distributed Processing – Workshop Proceedings*, pages 520–529, Cambridge, MA, USA, May 20-24 2013. IEEE Computer Society.

[71] Matthew Wezowicz#, Trilce Estrada+, Sandeep Patel, and **Michela Taufer**. Performance Dissection of a Molecular Dynamics Code across CUDA and GPU Generations. In *Proceedings of the 2013 IEEE International Symposium on Parallel & Distributed Processing – Workshop Proceedings*, pages 1355–1364, Cambridge, MA, USA, May 20-24 2013. IEEE Computer Society.

[72] Jeffrey DiMarco+ and **Michela Taufer**. Performance Impact of Dynamic Parallelism on Different Clustering Algorithms. In *Modeling and Simulation for Defense Systems and Applications VIII*, volume 8752, pages 97 – 104, Baltimore, Maryland, USA, May 2013. International Society for Optics and Photonics, SPIE.

[73] Daniel T. Yehdego+, Boyu Zhang, **Taufer, Michela**, Vikram Kumar Reddy Kodimala, Rahulsimham Vegesna, Sameera Viswakula, Kyle L. Johnson, and Ming-Ying Leung. Secondary Structure Predictions for Long RNA Sequences Based on Inversion Excursions: Preliminary Results. In *Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine (BCB)*, page 545–547, Orlando, Florida, USA, 2012. ACM.

[74] Boyu Zhang+, Daniel T. Yehdego+, Kyle L. Johnson, Ming-Ying Leung, and **Michela Taufer**. A Modularized MapReduce Framework to Support RNA Secondary Structure Prediction and Analysis Workflows. In *Proceedings of the 2012 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*, pages 86–93, Philadelphia, USA, October 4-7 2012. IEEE Computer Society.

[75] Michael Matheny#, Samuel Schlachter+, L. M. Crouse, E. T. Kimmel#, Trilce Estrada, Marcel Schumann, Roger S. Armen, Gary M. Zoppetti, and **Michela Taufer**. ExSciTecH: Expanding volunteer computing to Explore Science, Technology, and Health. In *Proceedings of the 8th IEEE International Conference on eScience*, pages 1–8, Chicago, IL, USA, October 8-12 2012. IEEE Computer Society.

[76] Trilce Estrada, Boyu Zhang, **Michela Taufer**, Pietro Cicotti, and Roger S. Armen. Reengineering High-throughput Molecular Datasets for Scalable Clustering Using MapReduce. In *Proceedings of the 14th IEEE International Conference on High Performance Computing and Communication & 9th IEEE International Conference on Embedded Software and Systems (HPCC-ICESS)*, pages 351–359, Liverpool, United Kingdom, June 25-27 2012. IEEE Computer Society.

[77] Trilce Estrada and **Michela Taufer**. On The Effectiveness of Application-Aware Self-Management for Scientific Discovery in Volunteer Computing Systems. In *Proceedings of the SC Conference on High Performance Computing Networking, Storage and Analysis (SC)*, pages 80–, Salt Lake City, UT, USA, November 11 - 15 2012. IEEE Computer Society. *(Acceptance Rate: 100/472, 21%).*

[78] Matthew Wezowicz, B. David Saunder, and **Michela Taufer**. Dealing with Performance/Portability and Performance/Accuracy Trade-offs in Heterogeneous Computing Systems: a Case Study with Matrix Multiplication Modulo . In *Modeling and Simulation for Defense Systems and Applications VII*, volume 8403, pages 39 – 48, Baltimore, Maryland, USA, April 2012.

International Society for Optics and Photonics, SPIE.

[79] Kyle Benson, Trilce Estrada, **Michela Taufer**, Jesse Lawrence, and Elizabeth Cochran. On the Powerful Use of Simulations in the Quake-Catcher Network to Efficiently Position Low-cost Earthquake Sensors. In *Proceedings of the IEEE 7th International Conference on eScience (eScience)*, pages 77–84, Stockholm, Sweden, December 5-8 2011. IEEE Computer Society. *(Acceptance Rate: 54/110, 50%).*

[80] Rolf Riesen, Kurt B. Ferreira, Maria Ruiz Varela, **Michela Taufer**, and Arun Rodrigues. Simulating Application Resilience at Exascale. In *Proceedings of the Euro-Par 2011: Parallel Processing Workshops (CCPI, CGWS, HeteroPar, HiBB, HPCVirt, HPPC, HPSS, MDGS, ProPer, Resilience, UCHPC, VHPC)*, volume 7156 of *Lecture Notes in Computer Science*, pages 221–230, Bordeaux, France, 2011. Springer.

[81] Trilce Estrada and **Michela Taufer**. Providing Quality of Science in Volunteer Computing. In *Proceedings of the 13th IEEE International Conference on High Performance Computing & Communication (HPCC)*, pages 68–77, Banff, Alberta, Canada, September 2-4 2011. IEEE Computer Society. *(Acceptance Rate: 59/271, 21.7%).*

[82] Narayan Ganesan, **Michela Taufer**, Brad A. Bauer, and Sandeep Patel. FENZI: GPU-Enabled Molecular Dynamics Simulations of Large Membrane Regions Based on the CHARMM Force Field and PME. In *Proceedings of the 25th IEEE International Symposium on Parallel and Distributed Processing (IPDPS) - Workshop Proceedings*, pages 472–480, Anchorage, Alaska, USA, May 2011. IEEE Computer Society.

[83] Narayan Ganesan, Roger D. Chamberlin, Jeremy Buhler, and **Michela Taufer**. Rolling Partial Prefix-sums to Speedup Evaluation of Uniform and Affine Recurrence Equations. In *Modeling and Simulation for Defense Systems and Applications VI*, volume 8060, pages 18 – 25, Orlando, Florida, USA, April 2011. International Society for Optics and Photonics, SPIE.

[84] Trilce Estrada, Roger S. Armen, and **Michela Taufer**. Automatic Selection of Near-Native Protein-Ligand Conformations using A Hierarchical Clustering And Volunteer Computing. In *Proceedings of the Proceedings of the First ACM International Conference on Bioinformatics and Computational Biology (BCB)*, pages 204–213, Niagara Falls, NY, USA, August 2-4 2010. ACM. *(Acceptance Rate: 37/136, 28%).*

[85] Narayan Ganesan, Roger D. Chamberlain, Jeremy Buhler, and **Michela Taufer**. Accelerating HMMER on GPUs by Implementing Hybrid Data and Task Parallelism. In *Proceedings of the Proceedings of the First ACM International Conference on Bioinformatics and Computational Biology (BCB)*, pages 418–421, Niagara Falls, NY, USA, August 2-4 2010. ACM. *(Acceptance Rate: 30/99, 33% - Short paper).*

[86] Abel Licon, **Michela Taufer**, Ming-Ying Leung, and Kyle L. Johnson. A Dynamic Programming Algorithm for Finding the Optimal Segmentation of an RNA Sequence in Secondary Structure Predictions. In *Proceedings of the ISCA 2nd International Conference on Bioinformatics and Computational Biology (BICoB)*, pages 165–170, Honolulu, Hawaii, USA, March 24-26 2010. ISCA. *(Acceptance Rate: 45%).*

[87] **Michela Taufer**, Omar Padron, Philip Saponaro, and Sandeep Patel. Improving Numerical Reproducibility and Stability in Large-Scale Numerical Simulations on GPUs. In *Proceedings of the 24th IEEE International Symposium on Parallel and Distributed Processing (IPDPS)*, pages 1–9, Atlanta, Georgia, USA, 19-23 April 2010. IEEE Computer Society. *(Acceptance Rate: 127/527, 24%).*

[88] Lifan Xu, **Michela Taufer**, Stuart Collins, and Dionisios G. Vlachos. Parallelization of Tau-leap

Coarse-grained Monte Carlo Simulations on GPUs. In *Proceedings of the 24th IEEE International Symposium on Parallel and Distributed Processing (IPDPS)*, pages 1–9, Atlanta, Georgia, USA, 19-23 April 2010. IEEE Computer Society. *(Acceptance Rate: 127/528, 24%).*

[89] Joseph E. Davis, Adnan Ozsoy, Sandeep Patel, and **Michela Taufer**. Towards Large-Scale Molecular Dynamics Simulations on Graphics Processors. In *Proceedings of the Bioinformatics and Computational Biology, First International Conference (BICoB)*, volume 5462 of *Lecture Notes in Computer Science*, pages 176–186, New Orleans, LA, USA, April 2009. Springer. *(Acceptance Rate: 30/72, 41.6%).*

[90] Trilce Estrada, **Michela Taufer**, and Kevin Reed. Modeling Job Lifespan Delays in Volunteer Computing Projects. In *Proceedings of the 9th IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGrid)*, pages 331–338, Shanghai, China, May 2009. IEEE Computer Society. *(Acceptance Rate: 57/271, 21%).*

[91] James Atlas, Trilce Estrada, Keith Decker, and **Michela Taufer**. Balancing Scientist Needs and Volunteer Preferences in Volunteer Computing Using Constraint Optimization. In *Proceedings of the 9th International Conference on Computational Science (ICCS)*, volume 5544 of *Lecture Notes in Computer Science*, pages 143–152, Baton Rouge, LA, USA, May 2009. Springer. *(Acceptance Rate: 30%).*

[92] Patrick McClory, Ezra Kissel, D. Martin Swany, and **Michela Taufer**. MNEMONIC: A Network Environment for Automatic Optimization and Tuning of Data Movement over Advanced Networks. In *Proceedings of the 18th International Conference on Computer Communications and Networks (ICCCN)*, pages 1–7, San Francisco, California, USA, August 2009. IEEE Computer Society.

[93] Trilce Estrada, **Michela Taufer**, Kevin Reed, and David P. Anderson. EmBOINC: An Emulator for Performance Analysis of BOINC Projects. In *Proceedings of the 23rd IEEE International Symposium on Parallel and Distributed Processing (IPDPS) – Workshop Proceedings*, pages 1–8, Rome, Italy, May 2009. IEEE Computer Society.

[94] Trilce Estrada, Olac Fuentes, and **Michela Taufer**. A Distributed Evolutionary Method to Design Scheduling Policies for Volunteer Computing. In *Proceedings of the 5th Conference on Computing Frontiers*, pages 313–322, Ischia, Italy, May 2008. ACM. *(Acceptance Rate: 30/110, 27%).*

[95] **Michela Taufer**, Thamar Solorio, Abel Licon, David Mireles, and Ming-Ying Leung. On the Effectiveness of Rebuilding RNA Secondary Structures from Sequence Chunks. In *Proceedings of the 22nd IEEE International Symposium on Parallel and Distributed Processing (IPDPS)*, pages 1–8, Miami, Florida USA, April 2008. IEEE Computer Society. *(Acceptance Rate: 10/25, 40%).*

[96] **Michela Taufer**, Andre Kerstens, Trilce Estrada, David A. Flores, Richard Zamudio, Patricia J. Teller, Roger S. Armen, and Charles L. Brooks III. Moving Volunteer Computing towards Knowledge-Constructed, Dynamically-Adaptive Modeling and Scheduling. In *Proceedings of the 21th International Parallel and Distributed Processing Symposium (IPDPS) – Workshop Proceedings*, pages 1–8, Long Beach, California, USA, March 2007. IEEE Computer Society. *(Acceptance Rate: 37/24, 65%).*

[97] **Michela Taufer**, Ming-Ying Leung, Kyle L. Johnson, and Abel Licon. RNAVLab: A Unified Environment for Computational RNA Structure Analysis based on Grid Computing Technology. In *Proceedings of the 21th International Parallel and Distributed Processing Symposium (IPDPS)*, pages 1–8, Long Beach, California, USA, March 2007. IEEE Computer Society. *(Acceptance Rate: 43%).*

[98] Richard Zamudio, Daniel Catarino, **Michela Taufer**, Brent Stearn, and Karan Bhatia. Topaz: Extending Firefox to Accommodate the GridFTP Protocol. In *Proceedings of the 21th International Parallel and Distributed Processing Symposium (IPDPS) – Workshop Proceedings*, pages 1–8, Long Beach, California, USA, March 2007. IEEE Computer Society.

[99] **Michela Taufer**, Andre Kerstens, Trilce Estrada, David A. Flores, and Patricia J. Teller. SimBA: A Discrete Event Simulator for Performance Prediction of Volunteer Computing Projects. In *Proceedings of the 21st International Workshop on Principles of Advanced and Distributed Simulation (PADS)*, pages 189–197, San Diego, California, USA, June 2007. IEEE Computer Society.

[100] Guillermo A. Lopez, **Michela Taufer**, and Patricia J. Teller. Evaluation of IEEE 754 Floating-point Arithmetic Compliance Across a Wide Range of Heterogeneous Computers. In *Proceedings of the Richard Tapia Celebration of Diversity in Computing Conference*, pages 1–4, Orlando, Florida, USA, October 2007. ACM.

[101] Trilce Estrada, David A. Flores, **Michela Taufer**, Patricia J. Teller, Andre Kerstens, and David P. Anderson. The Effectiveness of Threshold-Based Scheduling Policies in BOINC Projects. In *Proceedings of the Second International Conference on e-Science and Grid Technologies (eScience)*, pages 88–, Amsterdam, The Netherlands, December 2006. IEEE Computer Society. *(Acceptance Rate: 60/160, 37.5%).*

[102] Maria Gabriela Aguilera, Patricia J. Teller, **Michela Taufer**, and Felix Wolf. A Systematic Multi-Step Methodology for Performance Analysis of Communication Traces of Distributed Applications based on Hierarchical Clustering. In *Proceedings of the 20th International Parallel and Distributed Processing Symposium (IPDPS) – Workshop Proceedings*, Rhodes Island, Greece, April 2006. IEEE Computer Society.

[103] Trilce Estrada, Abel Licon, and **Michela Taufer**. compPknots: A Framework for Parallel Prediction and Comparison of RNA Secondary Structures with Pseudoknots. In *Proceedings of the Frontiers of High Performance Computing and Networking (ISPA) – Workshop Proceedings (FHPCN, XHPC, S-GRACE,GridGIS, HPC-GTP, PDCE, ParDMCom, WOMP, ISDF, and UPWN)*, volume 4331 of *Lecture Notes in Computer Science*, pages 677–686, Sorrento, Italy, December 2006. Springer.

[104] **Michela Taufer**, Patricia J. Teller, David P. Anderson, and Charles L. Brooks III. Metrics for Effective Resource Management in Global Computing Environments. In *Proceedings of the First International Conference on e-Science and Grid Technologies (eScience)*, pages 204–211, Melbourne, Australia, December 2005. IEEE Computer Society. *(Acceptance Rate: 54/171, 31.6%).*

[105] **Michela Taufer**, David P. Anderson, Pietro Cicotti, and Charles L. Brooks III. Homogeneous Redundancy: a Technique to Ensure Integrity of Molecular Simulation Results Using Public Computing. In *Proceedings of the 19th International Parallel and Distributed Processing Symposium (IPDPS)*, pages 9 –, Denver, CO, USA, April 2005. IEEE Computer Society. *(Acceptance Rate: 14/29, 47%).*

[106] **Michela Taufer**, Chahm An, Andreas Kerstens, and Charles L. Brooks III. Predictor@Home: A "Protein Structure Prediction Supercomputer" Based on Public-Resource Computing. In *Proceedings of the 19th International Parallel and Distributed Processing Symposium (IPDPS) – Workshop Proceedings (HiCOMB)*, pages 8 – 16, Denver, CO, USA, April 2005. IEEE Computer Society. *(Acceptance Rate: 10/32, 31.5%).*

[107] Derrick Kondo, **Michela Taufer**, Charles L. Brooks III, Henri Casanova, and Andrew A. Chien. Characterizing and Evaluating Desktop Grids: An Empirical Study. In *Proceedings of the 18th International Parallel and Distributed Processing Symposium (IPDPS)*, pages 26 –, Santa Fe, New

Mexico, USA, April 2004. IEEE Computer Society. *(Acceptance Rate: 31.7%).*

[108] Pietro Cicotti, **Michela Taufer**, and Andrew A. Chien. DGMonitor: A Performance Monitoring Tool for Sandbox-Based Desktop Grid Platforms. In *Proceedings of the 18th International Parallel and Distributed Processing Symposium (IPDPS) Workshop Proceedings – Workshop Proceedings*, pages 246 –, Santa Fe, New Mexico, USA, April 2004. IEEE Computer Society.

[109] **Michela Taufer**, Michael Crowley, Daniel J. Price, Andrew A. Chien, and Charles L. Brooks III. Study of a Highly Accurate and Fast Protein-Ligand Docking Algorithm Based on Molecular Dynamics. In *Proceedings of the 18th International Parallel and Distributed Processing Symposium (IPDPS) – Workshop Proceedings (HiCOMB)*, pages 188 – 196, Santa Fe, New Mexico, USA, April 2004. IEEE Computer Society.

[110] Bennet Uk, **Michela Taufer**, Thomas Stricker, Giovanni Settanni, Andrea Cavalli, and Amedeo Caflisch. Combining Task- and Data Parallelism to Speed up Protein Folding on a Desktop Grid Platform. In *Proceedings of the 3rd IEEE International Symposium on Cluster Computing and the Grid (CCGrid)*, pages 240–247, Tokyo, Japan, May 2003. IEEE Computer Society. *(Acceptance Rate: 39/114, 34.2%).*

[111] **Michela Taufer** and Thomas Stricker. A Performance Monitor Based on Virtual Global Time for Clusters of PCs. In *Proceedings of the 2003 IEEE International Conference on Cluster Computing (CLUSTER)*, pages 64–72, Kowloon, Hong Kong, China, December 2003. IEEE Computer Society. *(Acceptance Rate: 48/164, 29.3%).*

[112] Bennet Uk, **Michela Taufer**, Thomas Stricker, Giovanni Settanni, and Andrea Cavalli. Implementation and Characterization of Protein Folding on a Desktop Computational Grid - Is CHARMM a Suitable Candidate for the United Devices MetaProcessor? In *Proceedings of the 17th International Parallel and Distributed Processing Symposium (IPDPS)*, pages 50–, April, Nice, France 2003. IEEE Computer Society. *(Acceptance Rate: 119/407, 29.2%).*

[113] **Michela Taufer**, Thomas Stricker, Gerard Roos, and Peter Güntert. On the Migration of the Scientific Code Dyana from SMPs to Clusters of PCs and on to the Grid. In *Proceedings of the 2nd IEEE International Symposium on Cluster Computing and the Grid (CCGrid)*, pages 93–101, Berlin, Germany, May 2002. IEEE Computer Society. *(Acceptance Rate: 25.0%).*

[114] **Michela Taufer**, Egon Perathoner, Andrea Cavalli, Amedeo Caflisch, and Thomas Stricker. Performance Characterization of a Molecular Dynamics Code on PC Clusters: Is There Any Easy Parallelism in CHARMM? In *Proceedings of the 16th International Parallel and Distributed Processing Symposium (IPDPS)*, Fort Lauderdale, FL, USA, April 2002. IEEE Computer Society. *(Acceptance Rate: 98/258, 38%).*

[115] **Michela Taufer**, Thomas Stricker, and Roger Weber. Scalability and Resource Usage of an OLAP Benchmark on Clusters of PCs. In *Proceedings of the Fourteenth Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA)*, pages 83–94, Winnipeg, Manitoba, Canada, August 2002. ACM.

[116] **Michela Taufer** and Thomas Stricker. Accurate Performance Evaluation, Modelling and Prediction of a Message Passing Simulation Code based on Middleware. In *Proceedings of the ACM/IEEE Conference on Supercomputing (SC)*, pages 52–, Orlando, FL, USA, November 1998. IEEE Computer Society.

[117] Peter Arbenz, Martin Billeter, Peter Güntert, Peter Luginbühl, **Michela Taufer**, and Urs von Matt. Molecular Dynamics Simulations on Cray Clusters using the SCIDDLE-PVM Environment. In *Proceedings of the Parallel Virtual Machine - EuroPVM'96, Third European PVM Conference*, volume 1156 of *Lecture Notes in Computer Science*, pages 142–149, München, Germany, October

1996. Springer.

REPORTS

[1] Connor Scully-Allison, Ian Lumsden, Katy Williams, Jesse Bartels, **Michela Taufer**, Stephanie Brink, Abhinav Bhatele, Olga Pearce, and Katherine E. Isaacs. Designing an Interactive, Notebook-Embedded, Tree Visualization to Support Exploratory Performance Analysis. *CoRR*, abs/2205.04557, 2022.

[2] Ria Patel, Brandan Roachell, Silvina Caíno-Lores, Ross Ketron, Jacob Leonard, Nigel Tan, Duncan A. Brown, Ewa Deelman, and **Michela Taufer**. Reproducibility of the First Image of a Black Hole in the Galaxy M87 from the Event Horizon Telescope (EHT) Collaboration. *CoRR*, abs/2205.10267, 2022.

[3] Ariel Keller Rorabaugh, Silvina Caíno-Lores, Michael R. Wyatt II, Travis Johnston, and **Michela Taufer**. Peng4nn: An accurate performance estimation engine for efficient automated neural network architecture search. *CoRR*, abs/2101.04185, 2021.

[4] Robert F. Bird, Nigel Tan, Scott V. Luedtke, Stephen Lien Harrell, **Michela Taufer**, and Brian J. Albright. VPIC 2.0: Next generation particle-in-cell simulations. *CoRR*, abs/2102.13133, 2021.

[5] Mark Asch, François Bodin, Micah Beck, Terry Moore, **Michela Taufer**, and Jean-Pierre Vilotte. Cybercosm: New foundations for a converged science data ecosystem. *CoRR*, abs/2105.10680, 2021.

[6] Paula Olaya, Jay F. Lofstead, and **Michela Taufer**. Building containerized environments for reproducibility and traceability of scientific workflows. *CoRR*, abs/2009.08495, 2020.

[7] Duncan A. Brown, Karan Vahi, **Michela Taufer**, Von Welch, and Ewa Deelman. Reproducing GW150914: the first observation of gravitational waves from a binary black hole merger. *CoRR*, abs/2010.07244, 2020.

[8] Joshua Hoke Davis, Tao Gao, Sunita Chandrasekaran, and **Michela Taufer**. Studying the impact of power capping on mapreduce-based, data-intensive mini-applications on intel KNL and KNM architectures. *CoRR*, abs/1903.11694, 2019.

[9] Dylan Chapp, Danny Rorabaugh, Duncan A. Brown, Ewa Deelman, Karan Vahi, Von Welch, and **Michela Taufer**. Applicability study of the PRIMAD model to LIGO gravitational wave search workflows. *CoRR*, abs/1904.05211, 2019.

[10] Danny Rorabaugh, Mario Guevara, Ricardo M. Llamas, Joy Kitson, Rodrigo Vargas, and **Michela Taufer**. SOMOSPIE: A modular soil moisture spatial inference engine based on data driven decision. *CoRR*, abs/1904.07754, 2019.

[11] Travis Johnston, Boyu Zhang, Adam Liwo, Silvia Crivelli, and **Michela Taufer**. It-situ data analysis of protein folding trajectories. *CoRR*, abs/1510.08789, 2015.

POSTERS AND ABSTRACTS *(not in proceedings)*

[1] Silvina Caino-Lores, Michel Cuendet, Trilce Estrada, Ewa Deelman, Harel Weinstein, and **Taufer, Michela**. High-Throughput In-Situ Workflows for Ensemble Molecular Dynamics. In *Proceedings of the 18th IEEE International Conference on e-Science (eScience)*, pages 1–1, Salt Lake City, Utah, USA, October 2022. IEEE Computer Society.

[2] Nauweiluo Zhou, Jakob Luettgau, Rahul Reddy Kancharla, Joshua Kane, Brendan Croom, Robert Wheeler, Pania Newell, Giorgio Scorzelli, Valerio Pascucci, and **Taufer, Michela**. A Software

Framework for Scientific Workflow Orchestration at Large Scale. In *Proceedings of the 18th IEEE International Conference on e-Science (eScience)*, pages 1–1, Salt Lake City, Utah, USA, October 2022. IEEE Computer Society.

[3] Jakob Luettgau, Giorgio Scorzelli, Nauweiluo Zhou, Glenn Tarcea, Jay Lofstead, Valerio Pascucci, and **Taufer, Michela**. Toward a Lightweight Indexing Service for the National Science Data Fabric. In *Proceedings of the 18th IEEE International Conference on e-Science (eScience)*, pages 1–1, Salt Lake City, Utah, USA, October 2022. IEEE Computer Society.

[4] Clark Hathaway, Sebastian Mobo, Silvina Caíno-Lores, Travis Johnston, and **Michela Taufer**. A Framework for Linking Urban Traffic and Vehicle Emissions in Smart Cities. In *Poster at the 33rd ACM/IEEE International Conference for High Performance Computing and Communications conference (SC)*. IEEE Computer Society, November 2020. *(Best UG Poster Candidate)*.

[5] Ian Lumsden, Stephanie Brink, Michael R. Wyatt II, Todd Gamblin, and **Michela Taufer**. Enabling Graph-Based Profiling Analysis using Hatchet. In *Poster at the 33rd ACM/IEEE International Conference for High Performance Computing and Communications conference (SC)*. IEEE Computer Society, November 2020. *(Best UG Poster)*.

[6] Nigel Tan, Robert F. Bird, and **Michela Taufer**. Optimizing Vector Particle-In-Cell (VPIC) for Memory Constrained Systems Using Half-Precision. In *Poster at the 33rd ACM/IEEE International Conference for High Performance Computing and Communications conference (SC)*. IEEE Computer Society, November 2020. *(Best Graduate Poster Candidate)*.

[7] Paula Olaya, Jay Lofstead, and **Michela Taufer**. Containerized Environment for Reproducibility and Traceability of Scientific Workflows. In *Poster at the 33rd ACM/IEEE International Conference for High Performance Computing and Communications conference (SC)*. IEEE Computer Society, November 2020.

[8] Paula Olaya, Michael Wyatt, Silvina Caino-Lores, Piotr Luszczek, Osamu Miyashita, Florence Tama, and **Michela Taufer**. XPSI: XFEL-based Protein Structure Identifier. In *Poster at the 33rd ACM/IEEE International Conference for High Performance Computing and Communications conference (SC)*. IEEE Computer Society, November 2020.

[9] Nigel Tan, **Michela Taufer**, Scott Luedtke, Robert Bird, and Brian Albright. Expanding VPIC Portability to Large Scale GPU Systems. In *Abstract in Proceedings of the 62nd Annual Meeting of the APS Division of Plasma Physics*. IEEE Computer Society, November 2020.

[10] Rodrigo M Llamas, Mario Guevara, Danny Rorabaugh, **Michela Taufer**, and Rodrigo Vargas. Large-Scale Soil Moisture Modeling Based on Linear Geostatistics and Remotely Sensed Data. In *Abstract in AGU 100 – Advanced Earth and Space Science – Fall Meeting*, Washington DC, USA, December 2018. IEEE Computer Society.

[11] Dylan Chapp, Danny Rorabaugh, and **Michela Taufer**. Modeling Record-and-Replay for Nondeterministic Applications on Exascale Systems. In *Poster at ModSim 2018: Workshop on Modeling simulation of Systems and Applications*, Seattle, WA, USA, August 2018.

[12] Robert Searles, Stephen Herbein, Travis Johnston, **Michela Taufer**, and Sunita Chandrasekaran. Creating a Portable, High- Level Graph Analytics Paradigm For Compute and Data-Intensive Applications. In *Poster in 2018 GPU Technology Conference (GTC)*, San Jose, CA, USA, March 2018.

[13] S. Herbein, T. Patki, D. H. Ahn, D. Lipari, T. Dahlgren, D. Domyancic, and **Michela Taufer**. Fully Hierarchical Scheduling: Paving the Way to Exascale Workloads. In *Poster at the 29th ACM/IEEE International Conference for High Performance Computing and Communications con-*

ference (SC), Denver, CO, USA, November 2017. IEEE Computer Society. *(Best Poster Candidate).*

[14] S. Herbein, M. Matheny, M. Wezowicz, J. Kroger, J.S. Logan, J. Kim, S. Klasky, and **Michela Taufer**. Predictions of Large-scale QMCPack I/Os on Titan using Skel. In *Poster at the 24th ACM/IEEE International Conference for High Performance Computing and Communications conference (SC)*, Denver, CO, USA, November 2013. IEEE Computer Society. *(Acceptance Rate: 40%).*

[15] M. Wezowicz and **Michela Taufer**. On the Cost of a General GPU Framework - The Strange Case of CUDA 4.0 vs. CUDA 5.0. In *Poster at the 23th ACM/IEEE International Conference for High Performance Computing and Communications conference (SC)*. IEEE Computer Society, November 2012. *(Acceptance Rate: 47%).*

[16] T. Estrada, K. Pusecker, M. Torres, J. Cohoon, and **Michela Taufer**. Benchmarking Gender Differences in Voluntary Computer Projects. In *Poster at 2012 Grace Hopper Celebration of Women in Computing (GHC12)*, Baltimore, Maryland, USA, October 2012.

[17] B. Zhang, P. Cicotti, and **Michela Taufer**. MapReduce Clustering on Large Datasets using SSDs and Virtual Shared Memory. In *Poster at Extreme Science and Engineering Discovery Environment (XSEDE)*, Chicago, Illinois, USA, July 2012.

[18] T. Estrada, B. Zhang, R.S. Armen, and **Michela Taufer**. Study of Protein-ligand Binding Geometries using a Scalable and Accurate Octree-based Algorithm in MapReduce. In *Poster at the 21th ACM/IEEE International Conference for High Performance Computing and Communications conference (SC)*, Seattle, Washington, USA, November 2011. IEEE Computer Society. *(Acceptance Rate: 40%).*

[19] O. Rahaman, R. Armen, T. Estrada, D. Doren, **Michela Taufer**, and C. L. Brooks III. Binding Free Energy Prediction by Molecular Dynamics Based Docking and Volunteer Computing. In *Poster at Division of Computers in Chemistry for the 238th ACS National Meeting*, Washington, DC, USA, August 2009. IEEE Computer Society.

[20] N. Ganesan, S. Patel, and **Michela Taufer**. Simulations of Large Membrane Regions using GPU-enabled Computations - Preliminary Results. In *Poster at the 2010 Symposium on Application Accelerators in High Performance Computing (SAAHPC)*, Knoxville, Tennessee, USA, July 2010.

[21] L. Xu, S. Collin, **Michela Taufer**, and D.G. Vlachos. Parallelization of Tau-Leaping Coarse-Grained Monte Carlo Method for Efficient and Accurate Simulations on GPUs. In *Poster at the 19th ACM/IEEE International Conference for High Performance Computing and Communications conference (SC)*, Portland, Washington, USA, November 2009. IEEE Computer Society.

[22] K.S. Hogle, J.H. Upton, A. Licon, M.-Y. Leung, **Michela Taufer**, and K.L. Johnson. Role of RNA secondary structure in replication of Nodamura virus RNA2. In *Poster at the 27th Annual Meeting of American Society for Virology*, Cornell University, Ithaca, NY, USA, July 2008. IEEE Computer Society.

[23] T. Estrada, **Michela Taufer**, and K. Reed. Performance analysis of Volunteer Computing Trace. In *Poster ar the ACM/IEEE International Conference for High Performance Computing, Network, Storage, and Analysis conference (SC)*, Reno, Nevada, USA, November 2007. IEEE Computer Society. *(Acceptance Rate: 24.5%).*

[24] David A. Flores, Trilce Estrada, **Michela Taufer**, Patricia J. Teller, and Andre Kerstens. SimBA: a discrete event simulator for performance prediction of volunteer computing projects. In *Poster at the ACM/IEEE International Conference for High Performance Computing and Communications*

conference (SC), Tampa, FL, USA, November 2006. IEEE Computer Society.

[25] C. An, **Michela Taufer**, and C.L. Brooks III. Predictorhome: A Multiscale, Distributed Approach for Protein Structure Prediction. In *Poster at 229th ACM National Meeting*, San Diego, California, USA, March 2005. IEEE Computer Society.

[26] C. An, **Michela Taufer**, and C.L. Brooks III. Predictorhome: A Multiscale, Distributed Approach for Protein Structure Prediction. In *Poster at 6th Community Wide Experiment on the Critical Assessment of Techniques for Protein Structure Prediction (CASP6)*, Gaeta, Italy, December 2004.

[27] Matthew Wezowicz and **Michela Taufer**. On the Cost of a General GPU Framework: The Strange Case of CUDA 4.0 vs. CUDA 5.0. In *Poster at the ACM/IEEE International Conference for High Performance Computing and Communications conference (SC)*, pages 1535–1536, Salt Lake City, UT, USA, November 10-16 2012. IEEE Computer Society. *(Best UG Poster Candidate)*.

[28] Trilce Estrada, Boyu Zhang, Pietro Cicotti, Roger S. Armen, and **Michela Taufer**. Study of Protein-ligand Binding Geometries using a Scalable and Accurate Octree-based Algorithm in MapReduce. In *Poster at the ACM/IEEE International Conference for High Performance Computing and Communications conference (SC)*, pages 39–40, Seattle, WA, USA, November 12-18 2011. ACM.

THESIS

[1] **Michela Taufer**. *Inverting Middleware: Performance Analysis of Layered Application Codes in High Performance Distributed Computing*. PhD thesis, ETH Zurich, 2002.

[2] **Michela Taufer**. *Development of the Parallelization of the Software Package OPAL for the Simulation of Dynamic Molecules on Supercomputers*. Master's thesis, University of Padova, 1996.

PATENTS

System and Methods for Graphic Encoding of Macromolecules for Efficient High-Throughput Analysis UNMRI Ref. No. 2018-036-03.

TALKS

KEYNOTES

Dec 2022	In Situ Data Analytics For Next Generation Molecular Dynamics Workflows. 15 IEEE/ACM International Conference on Utility and Cloud Computing (UCC), Portland, OR, USA.
Sep 2022	Studying Degree And Sources Of Non-Determinism In MPI Applications Via Graph Kernels. Latin America High Performance Computing Conference (CARLA), Porto Alegre, RS, Brazil.
Sep 2022	In Situ Data Analytics For Next Generation Molecular Dynamics Workflows. Parallel Processing and Applied Mathematics (PPAM), Gdańsk, Poland.
Dec 2021	AI4IO: A Suite of AI-based Tools For IO-aware HPC Resource Management. International Conference on High Performance Computing, Data and Analytics (HiPC) – Virtual conference.

Oct 2021 In Situ Data Analytics for Next Generation Molecular Dynamics Workflows. Seventh International Conference on Big Data and Information Analytics (IEEE BigDIA 2021). October 29-31, 2021. Chongqing, China – Virtual conference.

Oct 2021 In Situ Data Analytics for Next Generation Molecular Dynamics Workflows. International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD). Oct 26-29, 2021 – Virtual conference.

Feb 2021 AI4IO: A Suite of AI-based Tools For IO-aware HPC Resource Management. First International Symposium on Checkpointing for Supercomputing (SuperCheck) – Virtual conference.

Nov 2020 In Situ Data Analytics for Next Generation Molecular Dynamics Workflows. Fifteen Workshop on Workflows in Support of Large-Scale Science (WORKS) – Virtual conference.

May 2020 In Situ Data Analytics for Next Generation Molecular Dynamics Workflows. EuroGraphics Symposium on Parallel Graphics and Visualization – Virtual conference.

Dec 2019 Scientific Applications and Heterogeneous Architectures – Data Analytics and the Intersection of HPC and Edge Computing. Thirteenth CHPC National Conference, Johannesburg, South Africa.

Aug 2019 Scientific Applications and Heterogeneous Architectures – Data Analytics and the Intersection of HPC and Edge Computing. EuroPar Conference, Göttingen, Germany.

May 2018 Modeling the Next-Generation High Performance Schedulers. ACM Conference on Principles of Advanced Discrete Simulation (PADS), Rome, Italy.

Feb 2018 Building the Next Generation of MapReduce Programming Models over MPI to Fill the Gaps between Data Analytics and Supercomputers. Ninth International Workshop on Programming Models and Applications for Multicores and Many-cores (PMAM), Vosendorf, Austria.

Sep 2017 Building the Next Generation of MapReduce Programming Models over MPI to Fill the Gaps between Data Analytics and Supercomputers. Keynote at the EuroMPI/USA 2017 Conference, Chicago, IL, USA.

Sep 2017 Challenges in Big Data Computing on HPC Platforms. Parallel Processing and Applied Mathematics (PPAM), Lublin, Poland.

Sep 2016 Who is Afraid of I/O? - Exploring I/O Challenges and Opportunities at the Exascale. IEEE Cluster Conference, Taipei, Taiwan.

Jun 2016 Who is Afraid of I/O? - Exploring I/O Challenges and Opportunities at the Exascale. Keynote at the 7th Workshop on Scientific Cloud Computing (ScienceCloud), Kyoto, Japan.

May 2015 The Numerical Reproducibility Fair Trade: Facing the Concurrency Challenges at the Extreme Scale. Fifth International Workshop on Accelerators and Hybrid Exascale Systems (AsHES), Hyderabad, India.

INVITED TALKS

Jun 2023 Building Trust in Scientific Applications through Data Traceability and Results Explainability. High Performance Computing Conference 2023, Cetraro, Italy.

Sep 2022 The Curious Case of Reproducing Scientific Results about Black Holes. Workshop on Clusters, Clouds, and Data for Scientific Computing (CCDSC). Chemin de Chanzé, France.

Jul 2022 The Curious Case of Reproducing Scientific Results about Black Holes. High Performance Computing Conference 2022, Cetraro, Italy.

Jul 2022 A Containerized Environment for Reproducibility and Traceability of Scientific Workflows. 2022 IEEE World Congress on SERVICES - Hybrid conference.

Jun 2022 AI4IO: A Suite of AI-Based Tools for IO-Aware HPC Resource Management. Sparse Days 2022, Saint-Girons, Ariège, France.

Oct 2021 AI4IO: A Suite of AI-Based Tools for IO-Aware HPC Resource Management. Seminar at the Università della Svizzera Italiana – Virtual seminar.

Sep 2020 In Situ Data Analytics for Next Generation Molecular Dynamics Workflows. SIGHPC ASCAN Chapter Seminar – Virtual seminar.

Sep 2020 AI4IO: A Suite of AI-Based Tools for IO-Aware HPC Resource Management. Joint Laboratory for Extreme Scale Computing (JLESC) Workshop – Virtual workshop.

Jul 2020 In Situ Data Analytics for Next Generation Molecular Dynamics Workflows. IBM TJ Watson – Virtual seminar.

Feb 2020 Cyberinfrastructure Tools for Precision Agriculture in the 21st Century. NSF CSSI PI Meeting. Seattle, WA, USA.

Feb 2020 Transparency and Reproducibility: Case Studies, Formalisms, and Structured Guidance in Scientific Applications at Scale. SIAM Conference on Parallel Processing for Scientific Computing (PP20), February 12-15, 2020, Seattle, Washington. USA.

Nov 2019 Algorithms for In Situ Data Analytics of Next Generation Molecular Dynamics Workflows. 5th International Workshop on Data Analysis and Reduction for Big Scientific Data (DRBSD-5), Denver, CO, USA.

Oct 2019 PRIONN: Predicting Runtime and I/O using Neural Networks. Big Data and Extreme-scale Computing Workshop, San Diego, CA, USA.

Sep 2019 Scientific Applications and Heterogeneous Architectures – Data Analytics and the Intersection of HPC and Edge Computing. IEEE eScience Conference, San Diego, CA, USA.

Jul 2019 Scientific Applications and Heterogeneous Architectures – Data Analytics and the Intersection of HPC and Edge Computing. Argonne Training Program on Extreme-Scale Computing, Chicago, IL, USA.

Jul 2019 Convergence of Data Generation and Analytics in the Era of Heterogeneous Applications and Edge Computing. Sandia National Laboratories, Albuquerque, NM, USA.

Jun 2019 In Situ Data Analytics for Next Generation Molecular Dynamics Workflows. HPC ISC Conference, Frankfurt, Germany.

Apr 2019 Convergence of Data Generation and Analytics in the Era of Heterogeneous Applications and Edge Computing. DoE Salishan Conference on High Speed Computing, Salishan, Oregon.

Apr 2019 Characterization of Power and Performance in Data-Intensive Applications using MapReduce over MPI. Joint Laboratory for Extreme Scale Computing (JLESC) Workshop, Knoxville, USA.

Apr 2019 Algorithms for In Situ Data Analytics of Next Generation Molecular Dynamics Workflows. Numerical Algorithms for High-Performance Computational Science Workshop, The Royal Society, London, UK.

Mar 2019 Filling the Gaps between Data Analytics and High Performance Computing with (some help from) MPI. Inaugural MPI-Beyond Workshop, University of Tennessee Chattanooga, Chattanooga, Tennessee, USA.

Feb 2019 In Situ Data Analytics for Next Generation Molecular Dynamics Workflows. Mini-symposium on "Computational Tools and Precision Medicine," SIAM CSE, Spokane, Oregon, USA.

Feb 2019 In Situ Data Analytics for Next Generation Molecular Dynamics Workflows. Big Data and Extreme-scale Computing Workshop, Kobe, Japan.

Nov 2018 PRIONN: Predicting Runtime and IO using Neural Networks and GPUs. NVIDIA booth at SC18. Dallas, Texas, USA.

Oct 2018 Cyberinfrastructure Tools for Precision Agriculture in the 21st Century. Big Data and Extreme-scale Computing Workshop, Indiana University, Bloomington, Indiana, USA.

Sep 2018 Modeling Record-and-Replay Costs for Nondeterministic Applications on Exascale Systems. Clusters and Computational Data for Scientific Computing Workshop, Lyon, France.

Aug 2018 In Situ Data Analytics for Next Generation Molecular Dynamics Workflows. Los Alamos National Laboratory, USA.

Jul 2018 Challenges in Big Data Analytics on High Performance Computing Systems. 13th Scheduling for Large Scale Systems Workshop, Cetraro, Italy.

Jul 2018 Building the Next-Generation HPC Schedulers. 13th Scheduling for Large Scale Systems Workshop, Cetraro, Italy.

Jun 2018 Building the Next-Generation HPC Schedulers. HPC ISC Conference, Frankfurt, Germany.

Jun 2018 Modeling the Next-Generation HPC Batch-Job Schedulers. 13th Scheduling for Large Scale Systems Workshop. Lawrence Berkeley National Laboratory, Berkeley, CA, USA.

May 2018 Convergence Opportunities and Limits in Big Data and Simulations. Workshop on Converging Simulation and Data-Driven Science. National Academy of Sciences, Engineering, and Medicine. Washington, D.C., USA.

Mar 2018 Transitioning Data Analytics of MD Simulations Toward the Exascale Era. Session on Data Analytics in HPC: An Applications' Perspective. SIAM Conference on Parallel Processing (SIAM PP), Tokyo, Japan.

Dec 2017 Challenges in Big Data Analytics on High Performance Computing Systems. Workshop on Data Intensive Computing. Shenzhen Institutes of Advanced Technology. Shenzhen, China.

Oct 2017 Challenges in Big Data Computing on HPC Platforms. Department of Electrical Engineering and Computer Science, The University of Tennessee at Knoxville, Knoxville, TN, USA.

Aug 2017 Impacts of Non-determinism on Numerical Reproducibility and Debugging at the Exascale. Analysis and Synthesis of Floating-point Programs Seminar, Dagstuhl, Germany.

Aug 2017 Who is Afraid of I/O? - Exploring I/O Challenges and Opportunities at the Exascale. Information Sciences Institute, Marina del Rey, CA, USA.

Apr 2017 Leveraging MapReduce and Machine Learning Technologies in Support of Big Data Analytics to Examine Food Nutrient Content. Value Institute, Christina Care Health System. Wilmington, DE, USA.

Apr 2017 Cyberinfrastructures for Big Data Analytics: Trends and Opportunities. University of Alabama, Birmingham. Birmingham, AL, USA.

Mar 2017 The Three Rs of Work in Scientific Papers: Repeatability, Replicability, and Reproducibility. High Performance Distributed Computing Technical Program Committee Workshop, Tampa, FL, USA.

Feb 2017 The Numerical Reproducibility Fair Trade: Facing the Concurrency Challenges at the Extreme Scale. SIAM Conference on Computational Science and Engineering (CSE17), Atlanta, GA, USA.

Feb 2017 Cyberinfrastructures for Big Data Analytics: Trends and Opportunities. Institute for Financial Services Analytics, University of Delaware, Newark, DE, USA.

Nov 2016 The Three Rs of Work in Scientific Papers: Repeatability, Replicability, and Reproducibility. Numerical Reproducibility at Exascale Workshop (NRE2016). In cooperation with SC16, Salt Lake City, UT, USA.

Nov 2016 Who is Afraid of I/O? - Exploring I/O Challenges and Opportunities at the Exascale. ExaIO Workshop. In cooperation with SC16, Salt Lake City, UT, USA.

Oct 2016 The Numerical Reproducibility Fair Trade: Facing the Concurrency Challenges at the Extreme Scale. 2016 CCL Workshop on Scalable Computing, University of Notre Dame, IN, USA.

Oct 2016 In-Situ Data Analytics and Indexing of Protein Trajectories. National Supercomputing Center - Guangzhou, Guangzhou, China.

Oct 2016 In-Situ Data Analytics and Indexing of Protein Trajectories. International Workshop on HPC Architecture, Software, and Application at an Extreme Scale. National Supercomputing Center - Wuxi, Wuxi, China.

Oct 2016 In Situ Data Analysis of Protein Trajectories. Clusters, Clouds, and Data for Scientific Computing (CCDSC). Chemin de Chanzé, France.

Aug 2016 In-Situ Data Analytics and Indexing of Protein Trajectories. Lawrence Livermore National Laboratory, Livermore, CA, USA.

Aug 2016 Who is Afraid of I/O? - Exploring I/O Challenges and Opportunities at the Exascale. Research Computing Center Seminar, University of Queensland, Brisbane, Australia.

Mar 2016 In-Situ Data Analysis of Protein-folding Trajectories. 251st ACS National Meeting & Exposition - Division of Computers in Chemistry: From Dynamics to Function & Back Again: Adventures in Simulating Biomolecules, San Diego, CA, USA.

Mar 2016 Resource Management for Running HPC Applications in Container Clouds. Recent Advances in HPDC Research Workshop, Pittsburg, PA, USA.

Mar 2016 Who is Afraid of I/O? - Exploring I/O Challenges and Opportunities at the Exascale. Rensselaer Polytechnic Institute, Troy, NJ, USA.

Dec 2015 In-Situ Data Analysis of Protein-folding Trajectories. University of Queensland, Brisbane, Australia.

May 2015 Enabling In-situ Analysis of Ligand Geometries in Drug Design Simulations on Supercomputers. 14th Workshop on High Performance Computational Biology, Hyderabad, India.

Apr 2015 Enabling In-Situ and Scalable Data Analysis of Folding Trajectories on Distributed Memory Systems. Novel Tools in Computational Chemistry Coding (NTC3) Meeting, Rutgers University, Piscataway, NJ, USA.

Mar 2015 Numerical Reproducibility Challenges on Extreme Scale Multi-threading GPUs. NVIDIA GPU Technology Conference, San Jose, CA, USA.

Mar 2015 Enabling Scalable Data Analysis of Large Computational Structural Biology Datasets on Distributed Memory Systems. 2015 Hot Topics in High-Performance Distributed Computing Workshop, IBM Almadena, California, USA.

Feb 2015 Enabling Scalable Data Analysis of Large Computational Structural Biology Datasets on Distributed Memory Systems. Delaware Bioinformatics Institute, Newark, DE, USA.

Jan 2015 Enabling Scalable Data Analysis of Large Computational Structural Biology Datasets on Distributed Memory Systems. National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA.

Oct 2014 Enabling Scalable Data Analysis of Large Computational Structural Biology Datasets on Distributed Memory Systems. Virginia Tech, Blacksburg, VA, USA.

Oct 2014 Enabling Scalable Data Analysis of Large Computational Structural Biology Datasets on Distributed Memory Systems. Rensselaer Polytechnic Institute (RPI), Troy, NY, USA.

Sep 2014 Performance and Cost Effectiveness of DAG-based Workflow Executions on the Cloud. Clusters, Clouds, and Data for Scientific Computing (CCDSC). Chemin de Chanzé, France.

Jul 2014 The Numerical Reproducibility Fair Trade: Facing the Concurrency Challenges at the Extreme Scale. Challenges in 21st Century Experimental Mathematical Computation. Institute for Computational and Experimental Research in Mathematics (ICERM). Providence, RI, USA.

May 2014 Enabling Scalable Data Analysis of Large Computational Structural Biology Datasets on Distributed Memory Systems. Stony Brook University, Stony Brook, NY, USA.

Apr 2014 Enabling Scalable Data Analysis of Large Computational Structural Biology Datasets on Distributed Memory Systems. Rutgers University, Piscataway, NJ, USA.

Mar 2014 Performance Impact of Dynamic Parallelism on Clustering Algorithms on GPUs. NVIDIA GPU Technology Conference, San Jose, CA, USA.

Mar 2014 Enabling Scalable Data Analysis of Large Computational Structural Biology Datasets on Distributed Memory Systems. University of Chicago, IL, USA.

Oct 2013 Enabling Scalable I/O and In-situ Analysis in Scientific Simulations at the Petascale. U.S. Army Research Laboratory at the Aberdeen Proving Ground, Aberdeen, MD, USA.

Sep 2013 On the Effectiveness of Application-aware Self-management for Scientific Discovery in Distributed Systems. ScalPerf '13 - Scalable Approaches to High Performance and High Productivity, Bertinoro, Italy.

Aug 2013 On the Effectiveness of Application-aware Self-management for Scientific Discovery in Volunteer Computing Systems. The University of Tennessee at Knoxville, Knoxville, TN, USA.

Aug 2013 On the Effectiveness of Application-aware Self-management for Scientific Discovery in Volunteer Computing Systems. Oak Ridge National Laboratory (ORNL), Oak Ridge, TN, USA.

May 2013 A Scalable and Accurate Method for Classifying Protein–ligand Binding Geometries using a MapReduce Approach. Novartis, Boston, MA, USA.

Mar 2013 Transforming Computing Algorithms and Paradigms in HPC to Enable more Science from our Day-to-day Simulations, Florida State University, Tallahassee, FL, USA.

Mar 2013 Application-aware Resource Management in Volunteer Computing. Workshop on Trends in High-Performance Distributed Computing, Rutgers University, Piscataway, NJ, USA.

Mar 2013 GPU-enabled Studies of Molecular Systems on Keeneland at ORNL - On pursuing high resource utilization and coordinated simulations' progression. NVIDIA GPU Technology Conference, San Jose, CA, USA. (With Sandeep Patel)

Oct 2012 Transforming Computing Algorithms and Paradigms in HPC to Enable more Science from our Day-to-day Simulations, Oak Ridge National Laboratory, Oak Ridge, TN, USA.

Oct 2012 Transforming Computing Algorithms and Paradigms in HPC to Enable more Science from our Day-to-day Simulations, Argonne National Laboratory, Chicago, IL, USA.

Jul 2012 Volunteer Computing for Drug Design, UD K-12 Engineering, University of Delaware, Newark, DE, USA.

May 2012 GPU-enabled Macromolecular Simulation: Challenges and Opportunities. NVIDIA GPU Technology Conference, San Jose, CA, USA. (With Sandeep Patel)

Mar 2012 GPU-enabled Macromolecular Simulation: Challenges and Opportunities, 2012 HPC Symposium at Lehigh University, Bethlehem, PA, USA.

Mar 2012 Reengineering High-throughput Molecular Datasets for Scalable Clustering using MapReduce, Workshop on Trends in High-Performance Distributed Computing, Vrije Universiteit, Amsterdam, Netherlands.

Feb 2012 GPU-enabled Macromolecular Simulation: Challenges and Opportunities, NVIDIA Headquarter, San Jose, CA, USA.

Dec 2011 GPU-enabled Macromolecular Simulation: Challenges and Opportunities, NVIDIA webinar, San Jose, CA, USA.

Mar 2011 Enabling Faster Large-Scale Simulations with GPU Programming, Aberdeen Army Research Laboratory, Aberdeen, MD, USA.

Oct 2010 Enabling Faster Molecular Dynamics Simulations and Protein Motif-Finding with GPU Programming. Enabling Discovery with Dell HPC GPU Solutions, Harvard Medical School, MA, USA.

Sep 2010 MD Simulations of Large Membranes. NVIDIA GPU Technology Conference, San Jose, CA, USA. (With Sandeep Patel and Narayan Ganesan)

Jun 2009 Computational Multi-Scale Modeling in Protein-Ligand Docking. Colloquium at IBM T.J. Watson, York Town, NY, USA.

Jan 2008 Computational Multi-Scale Modeling in Protein-Ligand Docking. Invited speaker at the 20th Annual CSU Biotechnology Symposium Information, Special Session on Interface between Computer Science and Biotechnology, Oakland CA, USA.

Apr 2007 DAPLDS: a Dynamically Adaptive Protein-Ligand Docking System based on Multi-Scale Modeling. Invited speaker at the Multiscale Modeling (MSM) PI Consortium Meeting, NIH, Bethesda, MD, USA.

Mar 2007 Moving Volunteer Computing towards Knowledge-Constructed, Dynamically-Adaptive Modeling and Scheduling. Department Colloquium - Department of Computer Science, Mississippi State University, Starkville, MS, USA.

Mar 2007 Moving Volunteer Computing towards Knowledge-Constructed, Dynamically-Adaptive Modeling and Scheduling. Department Colloquium - Department of Computer and Information Sciences, University of Delaware, Newark, DE, USA.

Mar 2007 Moving Volunteer Computing towards Knowledge-Constructed, Dynamically-Adaptive Modeling and Scheduling. Department Colloquium – Department of Computer Science, University of Pittsburgh, Pittsburgh, PA, USA.

Mar 2007 Moving Volunteer Computing towards Knowledge-Constructed, Dynamically-Adaptive Modeling and Scheduling. Department Colloquium - Department of Computer Science, University of New Mexico, Albuquerque, NM, USA.

Feb 2007 Moving Volunteer Computing towards Knowledge-Constructed, Dynamically-Adaptive Modeling and Scheduling. Department Colloquium – Department of Computer Science and Engineering, University of South Florida, Tampa, FL, USA.

Dec 2006 High-Performance Computing: An Increasingly Powerful Tool for Biomedical Science - what can HPC do for Cancer Research? Colloquium - San Antonio Cancer Institute Seminar Series - via AccessGrid.

Oct 2006 Predictor@Home: A "Protein Structure Prediction Supercomputer Based on Volunteer Computing." Invited speaker at the 19th Rocky Mountain Regional Meeting of The American Chemical Society, Tucson, AZ, USA.

Sep 2006 Moving Volunteer Computing Towards Data-Driven, Knowledge-Constructed Capabilities. Department Colloquium – Department of Computer Science at the University of Houston, Houston, TX, USA.

Apr 2006 Predictor@Home: A "Protein Structure Prediction Supercomputer" Based on Global Computing. Bioinformatics Colloquium – Universality of Texas at El Paso, El Paso, TX, USA.

Feb 2006 Predictor@Home: A "Protein Structure Prediction Supercomputer" Based on Global Computing. Colloquium – High Performance Computing Center, Texas Tech University, Lubbock, TX, USA.

Dec 2005 Metrics for Effective Resource Management in Global Computing Environments. Colloquium at National ICT, Australia, Sydney, Australia.

CONFERENCE/WORKSHOPS TALKS

Jun 2023 Runtime Steering of Molecular Dynamics Simulations Through In Situ Analysis and Annotation of Collective Variables. Platform for Advanced Scientific Computing (PASC) Conference. Davos, Switzerland.

Sep 2019 Characterization of Power Usage and Performance in Data-Intensive Applications using MapReduce over MPI. International Conference on Parallel Computing (ParCo) 2019 Conference. Prague, Czech Republic.

Dec 2018 On the Power of Combiner Optimizations in MapReduce over MPI Workflows. In IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS), Singapore.

Oct 2016 HYPPO: A Hybrid, Piecewise Polynomial Modeling Technique for Non-Smooth Surfaces. 28th IEEEE International Symposium on Computer Architecture and High-Performance Computing (SBAC-PAD), Los Angeles, CA, USA.

Jun 2016 Resource Management for Running HPC Applications in Container Clouds. International Supercomputing Conference (ISC), Frankfurt, Germany.

Apr 2016 In-Situ Data Analysis of Protein-folding Trajectories. Short talk at the Salishan Conference on High Speed Computing, Gleneden Beach, OR, USA.

Dec 2015 A Genetic Programming Approach to Design Resource Allocation Policies for Heterogeneous Workflows in the Cloud. 21st IEEE International Conference on Parallel and Distributed Systems (ICPADS), Melbourne, Australia.

Sep 2015 A Resource-selection Heuristic for High-performance and Cost-effective Workflow Execution on the Cloud. International Conference on Parallel Processing (ICPP), Beijing, China.

Aug 2015 From HPC Performance to Weather Modeling: Transforming Methods for HPC Predictions Into Models of Extreme Climate Conditions. Tenth IEEE International Conference on e-Science and Grid Technologies (eScience), Munich, Germany.

May 2015 Accurate Scoring of Drug Conformations at the Extreme Scale. Eight IEEE International Scalable Computing Challenge - Co-located with IEEE/ACM CCGrid, Shenzhen, China.

Sep 2014 Applying Frequency Analysis Techniques to DAG-based Workflows to Benchmark and Predict Resource Behavior on Non-Dedicated Clusters. IEEE Cluster Conference. Madrid, Spain.

Jun 2014 Study the Network Impact on Earthquake Early Warning in the Quake-Catcher Network Project. International Conference on Computational Science (ICCS), Cairns, Australia.

Dec 2013 Performance Impact of I/O on QMCPack Simulations at the Petascale and Beyond. 16th IEEE International Conferences on Computational Science and Engineering (CSE), Sydney, Australia.

Dec 2013 On Efficiently Capturing Scientific Properties in Distributed Big Data without Moving the Data - A Case Study in Distributed Structural Biology using MapReduce. 16th IEEE International Conferences on Computational Science and Engineering (CSE), Sydney, Australia.

Oct 2013 Efficient Sodium dodecyl sulfate (SDS) Simulations on Multi-GPU Nodes of XSEDE High-end Clusters. Eighth IEEE International Conference on e-Science and Grid Technologies (eScience), Beijing, China.

Oct 2013 Benchmarking Gender Differences in Volunteer Computing Projects. Third Workshop on Analyzing and Improving Collaborative eScience with Social Networks (eSoN). Beijing, China.

Oct 2012 ExSciTecH: Expanding Volunteer Computing to Explore Science, Technology, and Health. Second workshop on Analyzing and Improving Collaborative eScience with Social Networks (eSoN), Chicago, IL, USA.

Oct 2012 A Modularized MapReduce Framework to Support RNA Secondary Structure Prediction and Analysis Workflows. 2012 Computational Structural Bioinformatics Workshop (CSBW), Philadelphia, PA, USA.

Sep 2011 Providing Application-Level Quality of Science in Volunteer Computing. 13th IEEE High Performance Computing and Communications (HPCC) Conference, Banff, Canada.

May 2011 FEN ZI: GPU Enabled Molecular Dynamics Simulation of Large Membrane Regions Based on CHARMM Force Field and PME. Tenth IEEE Workshop on Hi-Performance Computational Biology (HiCOMB), Anchorage, AK, USA.

May 2011 FEN ZI: GPU Enabled Molecular Dynamics Simulation of Large Membrane Regions Based on CHARMM Force Field and PME. Tenth IEEE Workshop on Hi-Performance Computational Biology (HiCOMB), Anchorage, AK, USA.

Apr 2010 Improving Numerical Reproducibility and Stability in Large-Scale Numerical Simulations on GPUs. IEEE/ACM International Parallel and Distributed Processing Symposium (IPDPS), Atlanta, GA, USA.

Mar 2007 Moving Volunteer Computing towards Knowledge-Constructed, Dynamically-Adaptive Modeling and Scheduling. First Workshop on Large-Scale, Volatile Desktop Grids (PCGrid), Long Beach, CA, USA.

Dec 2006 The Effectiveness of Threshold-based Scheduling Policies in BOINC Projects. Second IEEE International Conference on e-Science and Grid Technologies (eScience), Amsterdam, The Netherlands.

Mar 2006 Web-based Tools to Facilitate Collaboration. International SUN Conference on Teaching and Learning, El Paso, Texas, USA.

Dec 2005 Metrics for Effective Resource Management in Global Computing Environments. First IEEE International Conference on e-Science and Grid Technologies (e-Science). Melbourne, Australia.

Apr 2005 Homogeneous Redundancy: a Technique to Ensure Integrity of Molecular Simulation Results Using Public Computing. 14th Heterogeneous Computing Workshop (HCW), Denver, CO, USA.

Apr 2005 Predictor@Home: A "Protein Structure Prediction Supercomputer" Based on Public- Resource Computing. Fourth IEEE International Workshop on High Performance Computational Biology (HiCOMB), Denver, CO, USA.

EXPERT PANELS

Nov 2023 HPC and Cloud Converged Computing: Merging Infrastructures and Communities. International Conference for High Performance Computing, Networking, Storage, and Analysis (SC), Denver, CO, USA.

Nov 2023 Unleashing the Power within Data Democratization: Needs, Challenges, and Opportunities. International Conference for High Performance Computing, Networking, Storage, and Analysis (SC), Denver, CO, USA.

Oct 2022 Will commercial AI be sufficient for scientific AI problems? Oak Ridge National Laboratory (ORNL) Core Universities AI workshop, Arlington, VA, USA.

Nov 2021 Hierarchical Parallelism for Exascale Computing. Second Workshop on Hierarchical Parallelism for Exascale Computing. International Conference for High Performance Computing, Networking, Storage, and Analysis (SC21), St. Louis, MO, USA.

Aug 2020 The Future of HPC Systems in the Presence of AI. (Keynote Panel) Smoky Mountains Computational Science and Engineering Conference, Virtual Conference.

Nov 2019 The National Academies Report on Reproducibility and Replicability in Science: Inspirations for the SC Reproducibility Initiative. The International Conference for High Performance Computing, Networking, Storage, and Analysis (SC19), Denver, CO, USA.

Jun 2019 Exascale and Beyond: Challenges in Productive and Sustainable Software. SIGHPC Platform for Advanced Scientific Computing (PASC) Conference. Zurich, Switzerland.

Aug 2018 Skills and Competencies for Modeling and Simulations. ModSim 2018: Workshop on Modeling & Simulation of Systems and Applications. Seattle, WA, USA.

Nov 2017 Reproducibility and Uncertainty in High Performance Computing? The International Conference for High Performance Computing, Networking, Storage, and Analysis (SC17), Denver, CO, USA.

Nov 2017 Blurring the Lines: High-End Computing and Data Science. The International Conference for High Performance Computing, Networking, Storage, and Analysis (SC17), Denver, CO, USA.

Sep 2017 MPI on Post-Exascale Systems. EuroMPI/USA 2017 Conference, Chicago, IL, USA. Sep 2016 HPC vs. Big Data: Different Worlds or Common Ground? IEEE Cluster Conference 2016, Taipei, Taiwan.

Feb 2016 Integration with the Scholarly Record: Case Studies and Lessons Learned. Panel moderator in AAAS workshop on Software Reproducibility, Washington DC, USA.

Jun 2008 A Day in the Life of a Researcher in Graduate School, Academia, and Industry. Panelist at CRA-W/CDC Systems Research Mentoring Workshop, University of Delaware, Newark, Delaware, USA.

Nov 2003 What we DO need to make Desktop Grids a Success in Practice. Panelist in panel discussion: "The Great Academia/Industry Grid Debate," 4th International Workshop on Grid Computing (Grid 2003), Phoenix, Arizona.

MINISYMPOSIUM SPEAKER

Jun 2023 Scalable GPU-Accelerated Incremental Checkpointing of Sparsely Updated Data. Mini-symposium on "Performance in I/O and Fault Tolerance for Scientific Applications." Platform for Advanced Scientific Computing (PASC) Conference, June 26-28, 2023, Davos, Switzerland.

Feb 2023 Robust Science Roadmap: Challenges in Software Systems and High Throughput Computing Applications. Mini-symposium on "A Roadmap to Robust Science for High-throughput Applications: Use Cases and Lessons Learned." SIAM Conference on Computational Science and Engineering (CSE), February 27–March 3, 2023, Amsterdam, The Netherlands.

Feb 2022 Managing High-Throughput Application Workloads: Findings and Recommendations. Mini-symposium on "A Roadmap to Robust Science for High-throughput Applications: The Developers' Perspective." SIAM Conference on Parallel Processing for Scientific Computing (PP22), February 23-26, 2022, Seattle, WA, USA.

Jul 2021 A4MD: In Situ Data Analytics for Next Generation Molecular Dynamics Workflows. Mini-symposium on "Building a Community Roadmap to Robust Science based on Performance Scalability, Trust, and Reproducibility in High-throughput Applications." Platform for Advanced Scientific Computing (PASC) Conference, July 5-9, 2021. Virtual Event.

Mar 2021 Reproducibility vs. Scalability in Containerized Workflows – The SOMPOSPIE Use Case. Mini-symposium on "Building a Community Roadmap to Robust Science in High-Throughput Applications." SIAM Conference on Computational Science and Engineering (CSE21), March 1-5, 2021. Virtual Event.

Feb 2020 Transparency and Reproducibility: Case Studies, Formalisms, and Structured Guidance in Scientific Applications at Scale. Mini-symposium on "Transparency, Reproducibility, Sustainability, and Security: The Four Pillars of the Next Generation Scientific Software Stack." SIAM Conference on Parallel Processing for Scientific Computing (PP20), February 12-15, 2020, Seattle, Washington, USA.

Mar 2019 In Situ Data Analytics for Next Generation Molecular Dynamics Workflows. Minisymposium on "Computational Tools and Precision Medicine." SIAM Conference on Computational Science and Engineering (CSE18), February 25 – March 1, 2019. Spokane, Oregon, USA

MINISYMPOSIUM ORGANIZER

Feb 2023 A Roadmap to Robust Science for High-throughput Applications: Use Cases and Lessons Learned. SIAM Conference on Computational Science and Engineering (CSE23), February 27–March 3, 2023, Amsterdam, The Netherlands.

Feb 2022 A Roadmap to Robust Science for High-throughput Applications: The Developers' Perspective. SIAM Conference on Parallel Processing for Scientific Computing (PP22), February 23-26, 2022, Seattle, WA, USA.

Jul 2021 Building a Community Roadmap to Robust Science based on Performance Scalability, Trust, and Reproducibility in High-throughput Applications. Platform for Advanced Scientific Computing (PASC) Conference, July 5-9, 2021. Virtual Event.

Mar 2021 Building a Community Roadmap to Robust Science in High-Throughput Applications. SIAM Conference on Computational Science and Engineering (CSE21), March 1-5, 2021. Virtual Event.

Feb 2020 Transparency, Reproducibility, Sustainability, and Security: The Four Pillars of the Next Generation Scientific Software Stack. SIAM Conference on Parallel Processing for Scientific Computing (PP20), February 12-15, 2020, Seattle, WA, USA.

Mar 2018 Data Analytics in HPC: An Applications' Perspective. SIAM Conference on Parallel Processing for Scientific Computing (PP18), March 7-10, 2018, Waseda University in Tokyo, Japan.

TUTORIALS

Jul 2019 Introduction of Practical Approaches to Data Analytics for HPC with Spark. ACM Europe Summer School, Barcelona Supercomputer Center – HPC Architectures for AI and Dedicated Applications, Barcelona, Spain. (Half-day tutorial)

Nov 2018 Introduction of Practical Approaches to Data Analytics for HPC with Spark. International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC18). (Half-day tutorial)

BIRDS OF A FEATHERS (BOFS)

Nov 2023 Hatem Ltaief, **Michela Taufer**, Lois Curfman McInnes, Christine Harvey, and Michael Bader. Meeting HPC Community Needs: How SIGHPC, TCPP, and SIAG-SC join efforts to engage communities and deliver services. BoF Session at the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), Denver, CO, USA.

Nov 2023 Valerio Pascucci, **Michela Taufer**, Christine Kirkpatrick, and Jakob Luettgau. A National Science Data Fabric to Democratize Data Access and Reusability. BoF Session at the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), Denver, CO, USA.

May 2023 Hatem Ltaief, **Michela Taufer**, Lois Curfman McInnes, Christine Harvey, and Michael Bader. Supercomputing with the Societies. BoF Session at the HPC ISC Conference, Hamburg, Germany.

May 2023 **Michela Taufer**, Jay Lofstead, Christine Kirkpatrick, Jakob Luettgau, Valerio Pascucci, and **Michela Taufer**. A National Science Data Fabric to Democratize Data Access and Reusability. BoF Session at the HPC ISC Conference, Hamburg, Germany.

Nov 2022 Jesus Carretero, Martin Schulz, Estela Suarez, **Michela Taufer**, and Michèle Weiland. Enabling I/O and Computation Malleability in High-Performance Computing. BoF Session at the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), Dallas, TX, USA.

Nov 2022 Daniel Milroy, Marquita Ellis, Sameer Shende, **Michela Taufer**, Ward Harold, and Yan Fisher. Converged Computing: Bringing Together HPC and Cloud Communities. BoF Session at the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), Dallas, TX, USA.

Nov 2015 Miriam Leeser, Dong Ahn, and bf Michela Taufer. Reproducibility of High-Performance Codes and Simulations – Tools, Techniques, Debugging. BoF Session at the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2015. Austin, TX, USA.

GROUP MEMBERS

RESEARCH ASSISTANT PROFESSORS

- *Jack Marquez* (Aug 2023 – present)
- *Jakob Luettgau* (Jan 2022 – present)

- *Silvina Caíno-Lores* (Sep 2022 – Apr 2023)

POST-DOCS

- *Jack Marquez* (Jul 2022 – Jul 2023)
- *Naweiluo Zhou* (Dec 2021 – Nov 2022)
- *Jakob Luettgau* (Oct 2021 – Aug 2022)
- *Ariel Rorabaugh* (Oct 2019 – Oct 2022)
- *Silvina Caíno-Lores* (Feb 2020 – Dec 2021)
- *Leobardo Valera* (Jan 2020 – Dec 2021)
- *Danny Rorabaugh* (Jan 2018 – Feb 2020)

- *Stephen Thomas* (Aug 2018 – May 2019)
  First placement: Engineer at Celgene
- *Wei-Fan Chien* (Sep 2016 – Mar 2017)
- *Travis Johnston* (Sep 2014 – Jul 2016)
  First placement: Researcher at Oak Ridge Nat. Lab.
- *Vivek Pallipuram* (Jan 2014 – Jul 2015)
  First placement: Assistant Professor at the U. Pacific
- *Trilce Estrada* (Jul 2012 – Jul 2013)
  First placement: Assistant Professor at U. New Mexico
- *Narayna Ganesan* (Jan 2010 – Jul 2011)
  First placement: Assistant Prof. at Stevens I. of Tech.
- *Arun Rajendran* (Apr 2008 – Jul 2008)

PROGRAMMERS

- *Heberth Martinez* (Jan 2023 – present)
- *Georgia Channing* (Jul 2022 – present)

- *Ria Patel* (Feb 2023 – Jun 2023)
- *Nick Bell* (Sep 2020 – Jul 2022)
- *Samuel Schlachter* (Jun 13 – Jul 14)
  First placement: CTO at SNAPCARD, Inc

STAFF

- *Grace Wisser* (Sep 2020 – present) Event Coordinator and Journal Manager
- *Barbara Fossum* (Sep 2019 – present) Outreach Coordinator
- *Lauren Whitnah* (Jun 2019 – present) Project Manager and Research Coordinator

PH.D. STUDENTS

- *Ria Patel* (Jun 2023 - present) Preliminary Research Exam on May 2025
- *Ian Lumsden* (Aug 2020 – present) Preliminary Research Exam on May 2024
- *Paula Olaya* (Aug 2020 – present) Preliminary Research Exam on Dec 2023
- *Nigel Tan* (Aug 2019 – present) Preliminary Research Exam on Dec 2023

- *Dylan Chapp* (UDel graduated, June 2020)
  Thesis Title: Modeling Non-Determinism of Scientific Applications. First placement: Software Engineering at Glodon, USA
- *Michael Wyatt* (UDel graduated, June 2020)
  Thesis Title: AI4IO: A Suite of AI-Based Tools for IO-Aware HPC Resource Management. First placement: Scientist at LLNL
- *Stephen Herbein* (UDel graduated, August 2018)
  Thesis Title: Scalable I/O-Aware Job Scheduling for Burst Buffer Enabled HPC Clusters. First placement: Scientist at LLNL. Award: IEEE TCSC Outstanding PhD Dissertation Award, 2019
- *Sean McDaniel* (UDel graduated, August 2018)
  Thesis Title: Computational Steering for Spike-coupled Neuronal Network Simulations on High-performance Computing Resources. First placement: Post-doctoral researcher at John Hopkins University

- *Boyu Zhang* (UDel graduated, May 2015)
  Thesis Title: Enabling Scalable Data Analysis for Large Computational Structural Biology Datasets on Large Distributed Memory Systems supported by the MapReduce Paradigm. First placement: Data analyst at Purdue University
- *Trilce Estrada* (UDel graduated, May 2012)
  Thesis Title: On the Effectiveness of Application-aware Self-management for Scientific Discovery on Volunteer Computing System. First placement: Post-doctoral associate at U. Delaware

MS STUDENTS

- *Jay Ashworth* (Aug 2023 - present)
- *Gabriel Laboy* (Aug 2023 - present)
- *Andrew Mueller* (Aug 2023 - present)

- *Treece Burgess* (Aug 2021 – July 2023)
- *Vanessa Lama* (May 2021 – May 2023)
  Thesis Title: Interactive Data Analysis of Multi-Run Performance Data.
- *Kae Suarez* (UTK graduated, May 2022)
- *Paula Olaya* (UTK graduated, May 2020)
  Thesis Title: Building containerized environments for reproducibility and traceability of scientific workflows.
- *Rachel Kraft* (UDel graduated, Dec 2018)
- *Dylan Chapp* (UDel graduated, May 2017)
  Thesis Title: Study of the Impact of Non-determinism on Numerical Reproducibility and Debugging at the Exascale. First placement: PhD Student, U. Delaware
- *Jeffrey DiMarco* (UDel graduated, May 2017)
  First placement: Software developer at Fidessa
- *Abel Licon* (UDel graduated, May 2010)
  Thesis Title: RNAVLab 2.0: Combining Web Applications, Grid Computing, and Dynamic Programming to Overcome Resource Limitations in RNA Secondary Structure Analysis. First placement: Researcher at Thermo Fisher Scientific
- *Joseph Davis* (UDel graduated, May 2009)
  Co-advised with Sandeep Patel. First placement: Scientist Siemens Healthcare
- *Prayook Tungjatooronrusame* (UTEP graduated, Dec 2006)
  Co-advised with Ming-Ying Leung
- *David Flores* (UTEP graduated, May 2007)
  Thesis Title: SimBA: A Discrete-event Simulator for Performance Prediction of Volunteer Computing Projects. Co-advised with Patricia Teller. First placement: Software developer at Ximis
- *Richard Zamudio* (UTEP graduated, May 2007)
  Thesis Title: TOPAZ: A Firefox Protocol Extension for GridFTP. First placement: Software developer at Rockwell Collins. Outstanding Thesis in Computer Science 2006-2007 (UTEP)

RESEARCH GRADUATE STUDENTS

- *Julius Plehn* (Jan 2023 – Jul 2023)
- *Maria Camila Roa* (Aug 2022 – Jul 2023)
- *Neil Lindquist* (Aug 2019 – Feb 2020)
- *Joseph Teague* (Aug 2018 – Dec 2019)

- *Mohammad Alsulmi* (Jan 2014 – Dec 2014)
- *Taylor Baldwin* (Jan 2014 – Aug 2014)
- *Marcos Portnoi* (Jan 2013 – Dec 2013)
- *William Killian* (Jun 2011 – May 2012)
- *Omar Padron* (Jun 2011 – May 2012)
- *Maria Ruiz* (Jan 2010 – Dec 2011)
- *Lifan Xu* (Jan 2019 – Dec 2010)
- *Kevin Kreiser* (Jun 2008 – Aug 2009)
- *Obaidur Rahaman* (Jun 2008 – May 2010)
- *James Atlas* (Jun 2008 – Aug 2009)
- *Adnan Ozsoy* (Aug 2008 – Dec 2008)
- *Roberto Araiza* (Jan 2007 – Dec 2007)

UNDERGRADUATE RESEARCH ASSISTANTS

- *Blake Milstead* (Aug 2023 – present)
- *Befikir Bogale* (Nov 2022 – present)
- *Seoyoung (Amy) An* (Aug 2022 – present) Co-mentor with Katie Schuman

- *Lauren Proctor* (Jan 2022 – May 2023)
- *Noah Dahle* (Feb 2023 – May 2023)
- *Brandon Roachell* (Jan 2021 – Dec 2022)
- *Ria Patel* (Feb 2020 – Dec 2022)
- *Dominic Kennedy* (Aug 2021 – Nov 2022)
- *Georgia Channing* (Jan 2022 – Jun 2022)
- *Cole Johnston* (Jan 2022 – Jun 2022)
- *Juliet Bradford* (Jan 2022 – May 2022)
- *Jacob Leonard* (Jan 2021 – present)
- *Ross Ketron* (Jan 2021 – present)
- *Clark Hathaway* (June 2020 – August 2021)
- *Sebastian Mobo* (Jun 2020 – May 2021)
- *Ian Lumsden* (Jun 2019 – Jul 2020)
- *Devon (Kae) Suarez* (Jan 2019 – May 2020)
- *Antonio Vega* (Aug 2019 – Dec 2019)
- *Matthew Dixson* (Jan 2019 – Aug 2019)
- *Josh Davis* (Aug 2017 – May 2019)
- *Joy Kitson* (Jun 2017 – May 2018)
- *John Bounds* (Jun 2016 – May 2018)
- *Paula Olaya* (Jun 2017 – Aug 2017)
  First placement: PhD Student, U.Tennessee Knoxville
- *Liz Racca* (Jun 2017 – Aug 2017)
- *Rachel Kraft* (Jun 2017 – Aug 2017)
  First placement: MS student, U. Delaware in CS at UDel
- *Connor Zanin* (Jan 2015 – May 2016)
  Senior Thesis: Tuning MapReduce with Surrogate- Based Modeling. First placement: PhD Student, Northeastern University
- *Ryan McKenna* (Aug 2014 – May 2016)
  Senior Thesis: Predicting Performance Variability in Parallel File Systems. First placement: PhD

Student, U. Massachusetts Amherst
- *Sean McDaniel* (Jan 2014 – Dec 2014) First placement: PhD Student, U. Delaware
- *Stephen Herbein* (Aug 2014 – Aug 2014)
  Senior Thesis: Benchmarking and Auto-tuning I/O Intensive Applications at the Extreme Scale.
  First placement: PhD Student, U. Delaware
- *Michael Matheny* (Jan 2012-Aug 2014)
  First placement: PhD Student at U. Utah
- *Samuel Schlachter* (Jun 2011 – May 2013)
  First placement: Researcher at U. Delaware
- *Matthew Wezowicz* (Aug 2011-Aug 2013)
- *Casey Casalnuovo* (Aug 2013 – Dec 2013)
- *Haley Northrup* (Aug 2013 – Dec 2013)
- *Ryan Huttman* (Aug 2011 – May 2013)
- *Reza Hammond* (Aug 2010 – May 2011) First placement: PhD Student at U. Delaware
- *Jason Park* (Jun 2010 – May 2011)
- *Kyle Benson* (Jun 2010 – Aug 2011)
  First placement: PhD Student, U. California Irvine
- *Dirk Mezger* (Aug 2010 – Dec 2010)
- *Dominik Kimmel* (Aug 2010 – Dec 2010)
- *Omar Padron* (Jun 2009- Aug 2009)
  First placement: PhD Student, U. Delaware
- *Philip Saponaro* (Jan 2009 – May 2010)
  Senior Thesis: An Efficient Arbitrary Precision Mathematical Library for Accurate and Fast MD Simulations in Single Precision GPUs. First placement: PhD Student, U. Delaware
- *Patrick McClory* (Jan 2008 – May 2009)
  First placement: PhD Student, U. Pittsburgh
- *Reed Matz* (Jun 2008 – Dec 2008)
- *Jason Parrott* (Aug 2007 – May 2008)
  First placement: Factset Research Systems
- *Robert Keller* (Aug 2007 – May 2008)
  First placement: Vanguard
- *Brenda Medina* (Jun 2008 – Aug 2008)
- *David Mireles* (Aug 2006 – Aug 2007, Jun 2008-Aug 2008)
- *Vladimir Soto* (Jan 2007 – Aug 2007)
- *Princess Trillo* (Jan 2007 – Aug 2007)
- *David Gomez-Leon* (Aug 2006 – May 2007)
- *Karina Escapita* (Jan 2006 – Aug 2007)
- *Guillermo Lopez* (Aug 2006 – Aug 2007)
- *Abel Licon* (Aug 2006 – Aug 2007)
  First placement: MS Student, U. Delaware
- *Daniel Catarino* (Jan 2006 – Dec 2006)
  First placement: Exxon Mobil

HIGH SCHOOL STUDENT

- *Lou Fogel* (Jun 2010 – Aug 2010)
  First placement: BS, Worcester Polytechnic Institute

VISITING SCHOLARS

- *Ahmed Bin Zaman* (May 2019 – Jul 2019)
- *Tao Gao* (Jan 2016 – Dec 2017)
- *Julian A. Uran* (Jun 2014 – Nov 2014)
- *Cindy Solano* (Jun 2013 – Aug 2013)
- *Daniel T. Yehdego* (Jun 2013 – Aug 2013)

Ph.D. / MS COMMITTEE MEMBER

- *Tanner Hobson* (PhD Thesis supervisor: Jian Huang)
  Ph.D. in Computer Science at U. Tennessee Knoxville, 2023
- *Ricardo llamas* (PhD Thesis supervisor: Rodrigo Vargas)
  Ph.D. in Plant Soil Sciences at U. Delaware, 2023
- *Zhixiu Lu* (PhD Thesis supervisor: Scott Emrich)
  Ph.D. in Computer Science at U. Tennessee Knoxville, 2023
- *Tu Mai Anh Do* (PhD Thesis supervisor: Ewa Deelman)
  Ph.D. in Computer Science at USC Information Sciences Institute, 2022
- *Angelica Walker* (PhD Thesis supervisor: Daniel Jacobson)
  Ph.D. in Data Sciences at U. Tennessee Knoxville, 2022
- *Craig Willis* (PhD Thesis supervisor: Victoria Stodden)
  Ph.D. in Information Sciences at U. Illinois Urbana-Champaign, 2020
- *Xi Luo* (Ph.D. Thesis supervisor: Jack Dongarra and George Bosilca)
  Ph.D. in Computer Science at U. Tennessee Knoxville, 2020
- *Mario Antonio Guevara Santamaria* (PhD Thesis supervisor: Rodrigo Vargas)
  Ph.D. in Plant Soil Sciences at U. Delaware, 2020
- *Thananon (Arm) Patinyasakdikul* (PhD Thesis supervisor: Jack Dongarra and George Bosilca)
  Ph.D. in Computer Science at U. Tennessee Knoxville, 2019
- *Reazul Hoque* (PhD Thesis supervisor: Jack Dongarra and George Bosilca)
  Ph.D. in Computer Science at U. Tennessee Knoxville, 2019
- *Robert Searles* (PhD Thesis supervisor: Sunita Chandrasekaran)
  Ph.D. in Computer Science at U. Delaware, 2019
- *Valentin Reis* (PhD Thesis supervisor: Denis Trystram)
  Ph.D. in Computer Science at Institute for Research in Computer Science and Automation (INRIA) 2018
- *Arnov Sinha* (MS Thesis supervisor: Sunita Chandrasekaran)
  MS in Computer Science at U. Delaware, 2017
- *Wei-Fan Chiang* (PhD Thesis supervisor: Ganesh Gopalakrishnan)
  Ph.D. in Computer Science at U. Utah, 2016
- *Fan Yang* (PhD Thesis supervisor: Paul Amer)
  Ph.D. in Computer Science at U. Delaware, 2015
- *Bryan Youse* (PhD Thesis supervisor: B. David Saunders)
  Ph.D. in Computer Science at U. Delaware, 2015
- *Kevin McCormick* (PhD Thesis supervisor: Li Liao)
  Ph.D. in Computer Science at U. Delaware, 2013
- *Daniel Orozco* (PhD Thesis supervisor: Guang R. Gao)
  Ph.D. in Computer Engineering at U. Delaware, 2012
- *Liang Gu* (PhD Thesis supervisor: Xiaoming Li)
  Ph.D. in Computer Engineering at U. Delaware, 2011

- *Kurt Ferreira* (PhD Thesis supervisor: Patrick Bridges)
  Ph.D. in Computer Science at U. New Mexico, 2011
- *Jayaraman Suresh Babu* (MS Thesis supervisor: Patricia J. Teller)
  MS in Computer Science at U. Texas El Paso, 2006
- *Maria Gabriela Aguilera* (MS Thesis supervisor: Patricia J. Teller)
  MS in Computer Science at U. Texas El Paso, 2005
- *Yash Dayal* (MS Thesis supervisor: Gregory Lush)
  MS in Electrical and Computer Engineering at U. Texas El Paso, 2005
- *Javed Bilal Khan* (MS Thesis supervisor: John Chessa)
  MS in Mechanical Engineering at U. Texas El Paso, 2005

STUDENT AWARDS
INTERNATIONAL LEVEL

- *Ian Lumsden*: First Place at the ACM Student Poster Competition (UG) at SC20
- *Sebastian Mobo and Clark Hathaway*: Third Place at the ACM Student Poster Competition (UG) at SC20
- *Nigel Tan*: Third Place at the ACM Student Poster Competition (Graduate) at SC20
- *Stephen Herbein*: IEEE TCSC Outstanding PhD Dissertation Award, 2019
- *Dylan Chapp*: Best Student Poster at the ACM HPDC Conference, 2019
- *Josh Davis*: Second Place at the ACM Student Poster Competition (UG) at SC18
- *Sean McDaniel*: First Place at the ACM Student Poster Competition (UG) at SC14
- *Stephen Herbein*: Second Place at the ACM Student Poster Competition (UG) at SC13
- *Matthew Wezowicz*: Second Place at the ACM Student Poster Competition (UG) at SC12
- *Philip Saponaro and Omar Padron*: Dr. Robert M. Panoff Award (UG) at SC09
- *Abel Licon*: Google Hispanic Scholarship Fund Scholarship, 2008
- *David Mireles*: Google Hispanic Scholarship Fund Scholarship, 2007
- *Daniel Catarino*: Google Hispanic Scholarship Fund Scholarship, 2006

UNIVERSITY LEVEL

- *Ria Patel*: U. Tennessee Knoxville Graduate Student Fellowships, 2023 – present
- *Ian Lumsden*: U. Tennessee Knoxville Graduate Student Fellowships, 2020 – present
- *Devon (Kae) Suarez*: U. Tennessee Knoxville Graduate Student Fellowships, 2020 – 2022
- *Nigel Tan*: U. Tennessee Knoxville Graduate Student Fellowships, 2019 – 2023
- *Neil Lindquist*: U. Tennessee Knoxville Graduate Student Fellowships, 2019 -2020
- *Paula Olaya*: U. Tennessee Knoxville Access and Diversity Fellowship, 2019
- *Joe Teague*: U. Tennessee Knoxville Graduate student fellowships, 2018 – 2019
- *Rachel Kraft*: U. Delaware University Graduate Scholar Award, 2017 – 2018
- *Sean McDaniel*: U. Delaware University Graduate Scholar Award, 2015 – 2017
- *Taylor Baldwin*: U. Delaware University Graduate Scholar Award, 2014 – 2015
- *Omar Padron*: U. Delaware University Graduate Scholar Award, 2011 – 2012
- *Trilce Estrada*: U. Delaware Graduate Fellow Award, 2010-2011
- *Philip Saponaro*: U. Delaware University Graduate Scholar Award, 2010 – 2011
- *Abel Licon*: U. Delaware University Graduate Scholar Award, 2009 – 2011
- *Trilce Estrada*: U. Delaware Alumni Enrichment Award, 2008

TEACHING

COURSES AT UTK (2018 –):

Graduate Courses – Semester, Course Title, Enrolment (Credits):

Sp23 COSC526: Introduction to Data Mining 30 (3)

Sp22 COSC526: Introduction to Data Mining 32 (3)

Sp21 COSC526: Introduction to Data Mining 19 (3)

Sp20 COSC526: Introduction to Data Mining 18 (3)

Fa18 COSC690(001) / COSC594(007): Big Data Analytics 28 (3)

Undergraduate Courses – Semester, Course Title, Enrolment (Credits):

Sp23 COSC426: Introduction to Data Mining 29 (3)

Sp22 COSC426: Introduction to Data Mining 24 (3)

COURSES AT UD (2007 – 2018):

Graduate Courses – Semester, Course Title, Enrolment (Credits):

Sp18 CISC879: Adv. Topics in Arch. and Softw. Systems: Big Data Analytics 10 (3)

Fa17 CISC879: Adv. Topics in Arch. and Softw. Systems: Big Data Analytics 9 (3)

Fa16 CISC879: Adv. Topics in Arch. and Softw. Systems: Big Data Analytics 23 (3)

Fa15 CISC879: Adv. Topics in Arch. and Softw. Systems: Big Data Analytics 11 (3)

Sp15 CISC663: Operating Systems 8 (3)

Fa14 CISC879: Adv. Topics in Arch. and Softw. Systems 13 (3)

Fa13 CISC663: Operating Systems 10 (3)

Fa12 CISC662: Computer Architecture 19 (3)

Sp12 CISC663: Operating Systems 7 (3)

Fa11 CISC662: Computer Architecture 13 (3)

Sp11 CISC879: High Performance Parallel Algorithms for Computational Science 9 (3)

Fa10 CISC662: Computer Architecture 27 (3)

Fa09 CISC662: Computer Architecture 27 (3)

Sp09 CISC849: High Performance Parallel Algorithms for Computational Science 9 (3)

Fa08 CISC662: Computer Architecture 25 (3)

Sp07 CISC849: Analysis of Bio. Simulations 7 (3)

Fa07 CISC 662: Computer Architecture 19 (3)

Undergraduate Courses – Semester, Course Title, Enrolment (Credits):

Sp17 CISC361: Operating Systems 44 (3)

Fa15 CISC361: Operating Systems 28 (3)

Sp15 CISC361: Operating Systems 40 (3)

Sp10 CISC361: Operating Systems 40 (3)

Fa09 CISC360: Computer Architecture 39 (3)

Fa08 CISC360: Computer Architecture 16 (3)

Fa07 CISC360: Computer Architecture 19 (3)

COURSES AT UTEP (2005 – 2007):

Graduate courses – Semester, Course Title:

Sp05, Sp06 CS5334: Parallel and Concurrent Programming

Sp07 CS5341: Analysis and Modeling of Biological Structures

Undergraduate courses – Semester, Course Title:

Sp07, Fa06, Sp06, Fa05 CS3320: Computer Architecture II

Fa06 CS3335: Systems Programming

SERVICES

EDITORIAL AFFILIATIONS

2020 — Editor in Chief, Future Generation Computer Systems (FGCS), Elsevier

2015 — Associate Editor, Journal of Parallel Computing (ParCo), Elsevier

2015 — Subject Area Editor, International Journal of High-Performance Computing Applications (IJHPCA), Sage

2015 — Subject Area Editor, Supercomputing Frontiers and Innovations Journal (SuperFI), Sage

2011 - 2015 Subject Area Editor, Journal of Parallel and Distributed Computing, Elsevier

STEERING AND ADVISORY COMMITTEES

2023— Member, Computing Community Consortium (CCC), part of the Computing Research Association (CRA)

2023— Member, International Advisory Board of the 2023 IUPAP Conference on Computational Physics

2021 — Member, Steering Committee of the South African CHPC National Conference

2021 — Member, Scientific Advisory Committee of the Helmholtz Imagining Platform and the Helmholtz Association (18 German research centers employing more than 40,000 staff)

2021 — Member, Advisory Board of the EuroHPC project ADMIRE — Adaptive multi-tier, intelligent data manager for Exascale

2020 —	Chair, Steering Committee of the IEEE International Conference on Cluster Computing (IEEE Cluster)
2020 —	Member, Advisory Board of the Anvil Project, Purdue University
2020	Chair, Steering Committee of the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC)
2019 —	Chair, Steering Committee of the NSF-founded South Big Data Hub Coordination Council
2018 —	Member, Steering Committee of the ACM PASC Conference
2016 — 2018	Member, Steering Committee of the NSF-founded South Big Data Hub Coordination Council
2016 —	Member, Steering Committee of the ISC High-Performance Computing
2015 —	Member, Steering Committee of the ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC)
2016 – 2019	Member, NSF Advisory Committee for Cyberinfrastructure (ACCI)
2016 - 2018	Member, Steering Committee of the IEEE International Parallel and Distributed Processing Symposium (IPDPS)
2014 – 2020	Member, Steering Committee of the IEEE International Conference on Cluster Computing (Cluster)
2014 – 2021	Member, Steering Committee of the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC)

SERVICES TO TECHNICAL SOCIETIES (Selected)

2022 —	Treasurer, IEEE Computer Society (CS)
2022 —	Vice-Chair, ACM Special Interest Group on High-Performance Computing (SIGHPC) – member elected
2019 – 2021	Chair, IEEE-CS Technical Meeting Request Committee (TMRC) for the Technical & Conference Activities Board
2019	Chair, ACM Senior Member Committee
2016 – 2022	Member-at-Large, ACM Special Interest Group on High-Performance Computing (SIGHPC) – member elected two consecutive times
2016 – 2018	Member, ACM Senior Member Committee – Chair in 2018
2015 – 2018	Member, Advisory Group on Reproducibility - Advisory to the SC Conference, ACM, and IEEE

CHAIR AND CO-CHAIR

Conferences / Workshops / Symposiums / Scholarships

2024	Program Chair of the 39th ISC High Performance Conference, June 2024, Hamburg, Germany
2023	Program Deputy Chair of the 38th ISC High Performance Conference, June 2023, Hamburg, Germany

2022 Technical Program Co-Chair of the IEEE eScience Conference. October 2022, Salt Lake City, Utah, USA

2021 Workshop Chair of the IEEE eScience Conference. September 2021, Innsbruck, Austria

Technical Program Track Chair of the IEEE International Parallel and Distributed Processing Symposium (IPDPS), May 2021, Portland, OR, USA

2019 General Chair of the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2019, Denver, CO, USA

2017 Finance Chair of the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2017, Denver, CO, USA

Technical Paper Area Co-Chair for Applications and Algorithms of the IEEE Cluster Conference, September 2017, Honolulu, Hawaii, USA

Workshop Chair of the 32nd ISC High-Performance Conference, June 2017, Frankfurt, Germany

General Chair of the IEEE International Parallel and Distributed Processing Symposium (IPDPS), May 2017, Orlando, FL, USA

SCALE Challenge Co-Chair of the 17th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing, to be held in May 2017, Madrid, Spain

2016 Panel Chair of the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2016. Salt Lake City, UT, USA

Technical Paper Area Chair of the "Applications" Track - 28th International Symposium on Computer Architecture and High-Performance Computing (SBAC-PAD)", October 26-28, 2016, Los Angeles, CA, USA

Technical Paper Area Chair of the "Multicore and Many-core Parallelism" Track – EuroPar, August 2016, Grenoble, France

Workshop co-Chair of the 31st ISC High-Performance Conference, June 19 – June 23, 2016, Frankfurt, Germany

Technical Paper Area Chair of the "Performance" Track – International Conference on Parallel Processing (ICPP), August 2016, Philadelphia, PA, USA

2015 Workshop Chair of the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2015. Austin, TX, USA

General co-Chair of the IEEE International Conference on Cluster Computing 2015, September 2015, Chicago, IL, USA

Technical Program Co-Chair of the 24th ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC), June 15-19, 2015, Portland, OR, USA

2014 Technical Program Co-Chair of the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2014. New Orleans, LA, USA

Technical Paper Area Chair of the "Cluster Design, Configuration and Administration" Track - IEEE International Conference on Cluster Computing (Cluster), September 2014, Madrid, Spain

2013 Technical Poster Chair of the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2013. Denver, CO, USA

2012 Birds of a Feather (BoF) Chair of the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2012, Salt Lake City, UT, USA

2011 Deputy Birds of a Feather (BoF) Chair of the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2011, Seattle, WA, USA

2010 Technical Paper Area Co-Chair of the System Software technical track at the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC). November 2010, New Orleans, LA, USA

Technical Paper Area Chair of "Distributed Systems and Applications" of the 12th IEEE International Conference on High-Performance Computing and Communications (HPCC), September 2010, Melbourne, Australia

2009 Technical Program vice-Chair of the topic "Distributed Systems and Applications" Track - 11th IEEE International Conference on High-Performance Computing and Communications (HPCC), June 2009, Soul, South Korea

Technical Program Chair of the 8th IEEE International Workshop on High-Performance Computational Biology (HiCOMB), May 2009, Rome, Italy

2003 Workshop co-Chair of the First Advanced Topics Workshop on Desktop Grids: Critical Systems and Applications Research (DGRID). November 2003, Phoenix, AZ, USA

COMMITTEE MEMBER

Conferences / Workshops / Symposiums / Scholarships

1. Tutorial Committee Member of the 35th ISC High-Performance Conference. May 21-25, 2023, Hamburg, Germany.
2. Technical Program Committee Member of the IEEE Cluster Conference (Cluster), September 2022, Heidelberg, Germany.
3. Technical Program Committee Member of the 51st IEEE International Conference on Parallel Processing (ICPP), August 2022, Bordeaux, France.
4. Technical Program Committee Member of the 34th IEEE International Parallel and Distributed Processing Symposium (IPDPS), May 2022, Nice, France.
5. Technical Program Committee Member of the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2021, St Louis, MO, USA.

6. Technical Program Committee Member of the IEEE International Conference on Cluster Computing (Cluster), September 2021, Portland, OR, USA.
7. Technical Program Committee Member of the IEEE International Conference on Parallel Processing (ICPP), August 2021, Argonne National Laboratory, IL, USA.
8. Technical Program Committee Member of the 41st IEEE International Conference on Distributed Computing Systems (ICDCS 2021), July 2021, Washington DC, USA.
9. Technical Program Committee Member of the 28th International Symposium on High-Performance Parallel and Distributed Computing (HPDC), June 2021. Stockholm, Sweden.
10. Technical Program Committee Member of the 21st IEEE/ACM International Symposium on Cluster, Cloud, and Internet Computing (CCGrid), May 2021, Melbourne, Australia.
11. Technical Program Committee Member of the 2020 IEEE International Symposium on Workload Characterization (IISWC), October 2020, Virtual Conference.
12. Technical Program Committee Member of the 40th IEEE International Conference on Distributed Computing Systems (ICDCS), July 2020, Singapore.
13. Technical Program Committee Member of the 3rd International Workshop on Reproducible Evaluation of Computer Systems (P-RECS), June 2020. Stockholm, Sweden.
14. Technical Program Committee Member of the 28th International Symposium on High-Performance Parallel and Distributed Computing (HPDC), June 2020. Stockholm, Sweden.
15. Tutorial Committee Member of the 35th ISC High-Performance Conference. June 21-25, 2020, Frankfurt, Germany.
16. Technical Program Committee Member of the 20th International Workshop on High-Performance Computational Biology (HiCOMB), May 2020, New Orleans, LA, USA.
17. Technical Program Committee Member of the International SC Asia Conference, February 2020, Singapore.
18. Technical Program Committee Member of the 5th International Workshop on Container Technologies and Container Clouds (WoC), December 2019. Davis, CA, USA.
19. Technical Program Committee Member of the 2nd International Workshop on Reproducible Evaluation of Computer Systems (P-RECS), June 2019. Phoenix, Arizona, USA.
20. Technical Program Committee Member of the 27th International Symposium on High-Performance Parallel and Distributed Computing (HPDC), June 2019. Phoenix, AZ, USA.
21. Technical Program Committee Member of the 32nd IEEE International Parallel and Distributed Processing Symposium (IPDPS), May 2019. Rio de Janeiro, Brazil.
22. Technical Program Committee Member of the 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), December 2018, Madrid, Spain.
23. Technical Program Committee Member of the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2018, Dallas, TX, USA.
24. Technical Program Committee Member of the 1st International Workshop on Reproducible Evaluation of Computer Systems (RECS), June 2018. Washington DC, USA.
25. Technical Program Committee Member of the 27th International Symposium on High-Performance Parallel and Distributed Computing (HPDC), June 2018. Phoenix, AZ, USA.
26. Technical Program Committee Member of the 33rd ISC High-Performance Conference. June 18-22,

2017, Frankfurt, Germany.

27. Technical Program Committee Member of the 32nd IEEE International Parallel and Distributed Processing Symposium (IPDPS), May 2018. Vancouver, Canada.
28. Workshop Committee Member of the International Conference on Computational Science (ICCS). June 11-13, 2017, Wuxi, China.
29. Panel Committee Member of the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2017. Denver, CO, USA.
30. Technical Program Committee Member of the 26th International Symposium on High-Performance Parallel and Distributed Computing (HPDC), June 2017, Washington DC, USA.
31. Technical Program Committee Member of the 32nd ISC High-Performance Conference. June 18-22, 2017, Frankfurt, Germany.
32. Workshop Committee Member of the International Conference on Computational Science (ICCS). June 12-14, 2017, Zurich, Switzerland.
33. Technical Program Committee of the 2017 IEEE International Symposium on Performance Analysis of Systems and Software. April 23-25, 2017, San Francisco Bay Area, California, USA.
34. Technical Program Committee of the 23rd IEEE International Conference on High-Performance Computing, Data and Analytics (HiPC), December 19-22, 2016, Hyderabad, India.
35. Workshop Committee Member of the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2016. Salt Lake City, UT, USA.
36. Student Cluster Competition Reproducibility Committee Member of the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2016. Salt Lake City, UT, USA.
37. Technical Program Committee Member of the System Track - IEEE International Conference on Cluster Computing, September 2016, Taipei, Taiwan.
38. Technical Program Committee Member ISC High-Performance Conference. June 19-23, 2016, Frankfurt, Germany.
39. Technical Program Committee Member of the 25th International Symposium on High-Performance Parallel and Distributed Computing (HPDC), June 2016, Kyoto, Japan.
40. Technical Program Committee Member of the 16th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid), May 2016, Cartagena, Colombia.
41. Technical Program Committee Member of the ACM International Conference on Computing Frontiers (CF), May 2016, Como, Italy.
42. Technical Program Committee Member of the 6th International Workshop on Adaptive Self-tuning Computing Systems (ADAPT), January 2016, Prague, Czech Republic.
43. Technical Program Committee Member of the 21st IEEE International Conference on Parallel and Distributed Systems (ICPADS), December 2015, Melbourne, Australia.
44. Technical Program Committee Member (Data Analytics and Visualization Track) of the IEEE/ACM International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2015. Austin, TX, USA.
45. Technical Program Committee Member of the IA$\hat{3}$ 2015: 5th Workshop on Irregular Applications: Architectures and Algorithms. November 2015. Austin, TX, USA.
46. Technical Program Committee Member of the EduHPC-15: Workshop on Education for High-

Performance Computing. November 2015. Austin, TX, USA.

47. Technical Program Committee Member of 11th IEEE International Conference on e-Science and Grid Technologies (eScience), August 2015, Munich, Germany.
48. Technical Program Committee Member of the 2014 ACM International Conference on Supercomputing (ICS), June 2015, Long Beach, USA.
49. Technical Program Committee Member of the ACM Computing Frontiers (CF), May 2015, Ischia, Italy.
50. Technical Program Committee Member of 2015 ACM/IEEE CS George Michael HPC Fellowship.
51. Technical Program Committee Member of Workshop and Tutorials at the 2015 Richard Tapia Celebration of Diversity in Computing Conference, Boston, MA, USA.
52. Technical Program Committee Member of the 13th IEEE International Conference on Ubiquitous Computing and Communications (IUCC) December 2014, Chengdu, China.
53. Technical Program Committee Member of 10th IEEE International Conference on e-Science and Grid Technologies (eScience), October 2014, Guarujá, San Paulo, Brazil.
54. Technical Program Committee Member of the 2nd Workshop on Parallel and Distributed Agent-Based Simulations (PADABS), 25-29 August 2014, Porto, Portugal.
55. Technical Program Committee Member of the 2014 ACM International Conference on Supercomputing (ICS), June 2014, Munich, Germany.
56. Technical Program Committee Member of the 23rd International Symposium on High-Performance Parallel and Distributed Computing (HPDC), June 2014, Vancouver, Canada.
57. Technical Program Committee Member of the ACM International Conference on Computing Frontiers 2014 (CF), May 2014, Cagliari, Italy.
58. Technical Program Committee Member of the 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid), May 2014, Chicago, USA.
59. Technical Program Committee Member (System Software Track) of the International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2013, Denver, Colorado, USA.
60. Technical Program Committee Member of the 13th International Workshop on High-Performance Computational Biology (HiCOMB), May 2013, Phoenix, Arizona, USA.
61. Technical Program Committee Member of the 6th IEEE/ACM International Conference on Utility and Cloud Computing (UCC), December 2013, Dresden, Germany.
62. Technical Program Committee Member of the 2013 IEEE International Conference on Big Data (IEEE Big Data 2013), October 6-9, 2013, Silicon Valley, CA, USA.
63. Technical Program Committee Member of the Workshop on Parallel Computational Biology (PBC), held in conjunction with PPAM 2013, September 8-11, 2013, Warsaw, Poland.
64. Technical Program Committee Member of the 1st Workshop on Parallel and Distributed Agent-Based Simulations (PADABS), August 2013, Aachen, Germany.
65. Technical Program Committee Member of the 22nd International Symposium on High-Performance Parallel and Distributed Computing (HPDC), June 2013, New York, NY, USA.
66. Technical Program Committee Member of the 12th International Workshop on High-Performance Computational Biology (HiCOMB), May 2013, Boston, MA, USA.

67. Technical Program Committee Member of the 13th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing (CCGrid), May 2013, Delft, The Netherlands.
68. Technical Program Committee Member of the International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2012, Salt Lake City, UT, USA.
69. Technical Program Committee Member (Programming Systems Track) of the International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2012, Salt Lake City, UT, USA.
70. Technical Program Committee Member of the High-Performance Computing Conference (HiPC), December 2012, Pune, India.
71. Technical Program Committee Member of the 5th IEEE/ACM International Conference on Utility and Cloud Computing (UCC), November 2012, Chicago, IL, USA.
72. Technical Program Committee Member of the Grace Hopper Conference (GHC) Panels, Workshops, and Presentations (PWP) Committee, October 2012, Baltimore, MD, USA.
73. Technical Program Committee Member of 8th IEEE International Conference on e-Science and Grid Technologies (eScience), October 2012, Chicago, IL, USA.
74. Technical Program Committee Member of Symposium on Application Accelerators in High-Performance Computing (SAAHPC), July 2012, Argonne National Laboratory, IL, USA.
75. Technical Program Committee Member of the 21st ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC), June 2012, Delft, The Netherlands.
76. Technical Program Committee Member of the 12th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing (CCGrid), May 2012, Ottawa, Canada.
77. Technical Program Committee Member of the 2012 ACM International Conference on Computing Frontiers (CF), May 15-17, 2012, Cagliari, Italy.
78. TCPP Travel Award Committee of the 26th IEEE International Parallel and Distributed Processing Symposium (IPDPS), May 21 – 25, 2012, Shanghai, China.
79. Technical Program Committee Member of the Workshop Innovative Parallel Computing: Foundations and Applications of GPU, Many-core, and Heterogeneous Systems (InPar), May 2012, San Jose, CA, USA.
80. Technical Program Committee Member (System Software Track) of the International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC), November 2011, Seattle, WA, USA.
81. Technical Program Committee Member of the 13th IEEE International Conference on High-Performance Computing and Communications (HPCC) in Biological/Molecular Computing Track, September 2 – 4, 2011, Banff, Alberta, Canada.
82. Technical Program Committee Member of the IEEE Cluster 2011 Conference (Cluster), September 26 – 30, 2011, Austin, TX, USA.
83. Technical Program Committee Member of the Workshop on Parallel Computational Biology (PBC), held in conjunction with PPAM 2011, September 11-14, 2011, Torun, Poland.
84. Technical Program Committee Member of the 11th IEEE International Symposium on Cluster Computing and Grid (CCGrid), May 23 – 26, 2011, Los Angeles, CA, USA.
85. Technical Program Committee Member of the 2011 Symposium on Application Accelerators in High-Performance Computing (SAAHPC), July 19 – 20, 2011, University of Tennessee Conference

Center, TN, USA.

86. Technical Program Committee Member of the 7th International Workshop on High-Performance Computational Biology (HiCOMB), May 16, 2011, Anchorage, AK, USA.
87. Technical Program Committee Member of the 4th Annual Workshop for General-Purpose Computation on Graphics Processing Units (GPGPU), March 5, 2011, Newport Beach, California, USA.
88. Technical Program Committee Member of the 18th Euromicro Conference on Parallel, Distributed and Network-Based Processing (PDP), February 9-11, 2011, Ayia Napa, Cyprus.
89. Technical Program Committee Member of the 2010 IEEE 6th International Conference on e-Science (eScience), December 7 – 10, 2010, Brisbane, Australia.
90. Technical Program Committee Member of the International Conference of Computer Design (ICCD), October 3-6, 2010, Amsterdam, The Netherlands.
91. Technical Program Committee Member of the 22nd International Symposium on Computer Architecture and High-Performance Computing (SBAC-PAD), October 2010, Petropolis, Brazil.
92. Technical Program Committee Member of the Workshop on Parallel Programming and Applications on Accelerator Clusters (PPAAC), September 2010, Heraklion, Greece.
93. Technical Program Committee Member of the 2010 Symposium on Application Accelerators in High-Performance Computing (SAAHPC), July 13 – 15, 2010, University of Tennessee Conference Center, TN, USA.
94. Technical Program Committee Member of the 2010 ACM International Symposium on High-Performance Distributed Computing (HPDC), June 2010, Chicago, IL, USA.
95. Technical Program Committee Member of the ACM Computing Frontiers Conference (CF), May 2010, Bertinoro, Italy.
96. Technical Program Committee Member of the 10th IEEE International Symposium on Cluster Computing and Grid (CCGrid), May 2010, Melbourne, Australia.
97. Technical Program Committee Member of the Second Workshop on Large-Scale, Volatile Desktop Grids (PCGrid), May 2010, Melbourne, Australia.
98. Technical Program Committee Member of the 18th Euromicro Conference on Parallel, Distributed and Network-Based Processing (PDP), February 2010, Pisa, Italy.
99. Technical Program Committee Member (System Software Track) of the International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC). November 2009, Portland, OR, USA.
100. Technical Program Committee Member of the International Conference of Computer Design (ICCD), November 4 – 7, 2009, Lake Tahoe, CA, USA.
101. Scholarship Committee Member of the Grace Hopper Celebration of Women in Computing 2009, October 2009, Tucson, AZ, USA.
102. Technical Program Committee Member of the 2009 IEEE International Conference on Cluster Computing (Cluster), August 29 – September 4, 2009, New Orleans, LA, USA.
103. Technical Program Committee Member of the 12th IEEE International Conference on Computational Science and Engineering (CSE), August 29 – 31, 2009, Vancouver, Canada.
104. Technical Program Committee Member of 9th IEEE International Symposium on Cluster Computing and Grid (CCGrid), May 2009, Shanghai, China.

105. Technical Program Committee Member of the Workshop on Using Emerging Parallel Architectures for Computational Science, held in conjunction with the ICCS 2009, May 2009, Baton Rouge, LA, USA.
106. Technical Program Committee Member of the Second Workshop on Large-Scale, Volatile Desktop Grids (PCGrid), May 2009, Rome, Italy.
107. Technical Program Committee Member of the 2009 Richard Tapia Celebration of Diversity in Computing Conference, April 2009, Portland, OR, USA.
108. Technical Program Committee Member of the 17th Euromicro Conference on Parallel, Distributed and Network-Based Processing (PDP), February 2009, Bauhaus-University Weimar in Thuringia, Germany.
109. Technical Program Committee Member of the Intl. Conference on Advanced Computing and Communications, December 2008, Chennai, India.
110. Technical Program Committee Member of the Computational Structural Bioinformatics Workshop 2008 November 2008, Philadelphia, PA, USA.
111. Technical Program Committee Member of the International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC). November 2008, Austin, TX, USA.
112. Technical Program Committee Member of the Grace Hopper Celebration of Women in Computing 2008, October 2008, Denver, CO, USA.
113. Technical Program Committee Member of the IEEE Intl. Conference on Computer Design (ICCD), October 2008, Lake Tahoe, CA, USA.
114. Technical Program Committee Member of the 10th IEEE International Conference on High-Performance Computing and Communications (HPCC), September 2008, DaLian, China.
115. Technical Program Committee Member of the International Conference on Computational Science (ICCS), June 2008, Krakow, Poland.
116. Technical Program Committee Member of the ACM Computing Frontiers (CF), May 2008, Ischia, Italy.
117. Technical Program Committee Member of the Global and Peer-to-Peer Computing (GP2PC), May 2008, Lyon, France.
118. Technical Program Committee Member of the 7th International Workshop on High-Performance Computational Biology (HiCOMB), May 2008, Miami, FL, USA.
119. Technical Program Committee Member of the 9th IEEE International Workshop on Parallel and Distributed Scientific and Engineering Computing (PDSEC), May 2008, Miami, FL, USA.
120. Technical Program Committee Member of the 2nd Workshop on Large-Scale, Volatile Desktop Grids (PCGrid), May 2008, Miami, Florida, USA.
121. Technical Program Committee Member of the 22nd IEEE International Parallel and Distributed Processing Symposium (IPDPS), May 2008, Miami, FL, USA.
122. Technical Program Committee Member of the Euromicro Conference on Parallel, Distributed and Network-based Processing (PDP) February 2008, Toulouse, France.
123. Technical Poster Committee Member of the International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC). November 2007, Reno, NV, USA.
124. Technical Program Committee Member of the First Computational Structural Bioinformatics Workshop, November 2007, San Jose, CA, USA.

125. Technical Poster Committee Member of the 2007 Richard Tapia Celebration of Diversity in Computing Conference, October 2007, Orlando, FL, USA.
126. Scholarship Committee Member of the Grace Hopper Celebration of Women in Computing 2007, October 2007, Orlando, FL, USA.
127. Technical Program Committee Member of the 5th IEEE International Symposium on Parallel and Distributed Processing and Applications (ISPA), August-September, 2007, Niagara Falls, Ontario, Canada.
128. Technical Program Committee Member of the International Conference on Computational Science 2007 (ICCS), May 2007, Beijing, China.
129. Technical Program Committee Member of the 6th International Workshop on Global and Peer-to-Peer Computing (GP2P), May 2007, Rio de Janeiro, Brazil.
130. Technical Program Committee Member of 26th IEEE International Performance Computing and Communications Conference (IPCCC), April 2007 - New Orleans, LA, USA.
131. Technical Program Committee Member of the First Workshop on Large-Scale, Volatile Desktop Grids (PCGrid), March 2007, Long Beach, CA, USA.
132. Technical Program Committee Member of the 4th IEEE International Symposium on Parallel and Distributed Processing and Applications (ISPA), December 2006, Sorrento, Italy.
133. Technical Program Committee Member of the 2nd IEEE International Conference on e-Science and Grid Technologies (eScience), December 2006, Amsterdam, The Netherlands.
134. Technical Program Committee Member of the 5th International Workshop on Global and Peer-to-Peer Computing (GP2P), May 2006, Singapore.
135. Technical Program Committee Member of the 20th IEEE International Parallel and Distributed Processing Symposium (IPDPS), April 2006, Rhodes, Greece.
136. Technical Program Committee Member of the 5th IEEE International Workshop on High-Performance Computational Biology (HiCOMB), April 2006, Rhodes, Greece.
137. Technical Program Committee Member of the 1st IEEE International Conference on e-Science and Grid Technologies (eScience), December 2005, Melbourne, Australia.
138. Technical Program Committee Member of the 2005 IEEE International Conference on Cluster Computing (Cluster), September 2005, Boston, MA, USA.
139. Technical Program Committee Member of the 2005 International Conference on High-Performance Computing and Communications (HPCC), September 2005, Sorrento, Italy.
140. Technical Program Committee Member of the 5th International Workshop on Global and Peer-to-Peer Computing (GP2P), May 2005, Cardiff, UK.

JOURNAL AND BOOK REFEREE

Since 2005, I have reviewed several articles for journal and book editors, including IEEE Transactions on Parallel and Distributed Systems, Journal of Computational Chemistry, Parallel and Distributed Computing, and Journal of Bioinformatics.

SERVICES TO THE UNIVERSITY, COLLEGE, AND DEPARTMENT (SELECTED)

At UTK:

2020 Member, Committee to review StAR proposals. (University Committee)

2019 – present Member, Committee to review faculty for their post-tenure performance (College Committee)

2018 – present Member, Faculty Search (Department) Committee

2018 Member, Organization Committee of the NIMBioS Investigative Workshop Scientific Collaboration Enabled by High-Performance Computing (University Committee)

At UD:

2016 – 2018 Chair, Publicity/Awards/Development Activity Committee (CIS Department)

2016 – 2018 Faculty Secretary, College of Engineering (CoE)

2007 – 2018 Coordinator, CIS/ECE booth at the Supercomputing Conference (CIS Department)

2017 – 2018 Member, Faculty Recruitment Committee (CIS Department)

2016 – 2017 Member, Ad-hoc College of Engineering Committee on Server Room Relocation (CoE)

2015 – 2016 Acting Director, Center for Bioinformatics & Computational Biology (CBCB), Delaware Biotechnology Institute (DBI)

2015 – 2016 Member, Faculty Recruitment Committee (ECE Department)

2014 – 2016 Member, Bioinformatics Steering Committee (Bioinformatics Program)

2014 – 2016 Advisor, Bioinformatics Student Association (Bioinformatics Program)

2015 Member, Undergraduate Committee (CIS Department)

2014 – 2015 Chair, Distinguished Speaker Series Committee (CIS Department)

2014 Chair, Faculty Recruitment Committee (CIS Department)

2012 – 2013 Member, Biomedical Engineering Graduate Committee (BME Program)

2012 – 2013 Chair, Distinguished Speaker Series Committee (CIS Department)

2010 Organizer, CIS Research Day (CIS Department)

2009 – 2013 Member, Bioinformatics Program Committee (University Committee)

2009 – 2011 Member, Research Computing Task Force (University Committee)

2009 – 2010 Member, Graduate Recruiting Committee (CIS Department)

2008 – 2009 Member, Graduate Committee (CIS Department)

2007 – 2009 Advisor, Student ACM Chapter (CIS Department)

2007 – 2008 Member, Graduate Recruiting Committee (CIS Department)

At UTEP:

2006 – 2007 Member, Bioinformatics Research Committee and Bioinformatics Colloquium Committee (University Committee)

2005 – 2007 Member, High-End Computing Along the Rio Grande Consortium

2005 – 2007 Member, Computer System Curriculum Committee and the Facilities Committee

2006 Member, NSF-CSEMS Scholarship Committee at the University of Texas at El Paso, 2006

SERVICE TO FEDERAL AGENCIES

Panelist for National Science Foundation (NSF), National Institutes of Health (NIH), Department of Energy (DoE), Army Research Office (ARO)

Note: The single panels are not provided to ensure the confidentiality of the review process.

PROFESSIONAL AFFILIATIONS

- ACM, ACM SIGHPC
- IEEE, IEEE-CS
- SIAM, SIAM-SC
- AAAS