

# Stephen Herbein

Computer Scientist, Livermore Computing  
Lawrence Livermore National Laboratory

☎ 302.689.3894  
✉ [stephen@herbein.net](mailto:stephen@herbein.net)  
<https://herbein.net>

---

## Education

- 2014-2018 **Ph.D., Computer Science**, *University of Delaware*, GPA 4.0.  
Advisor: Dr. Michela Taufer
- 2016 **M.S., Computer Science**, *University of Delaware*, GPA 4.0.
- 2014 **Honors B.S. with Distinction, Computer Science**, *University of Delaware*.  
GPA 3.932; Magna Cum Laude

---

## Research Interests

- Next-generation resource management and job scheduling
- Exascale workflows and HPC+Cloud convergence
- I/O-aware scheduling

---

## Experience

- 2021-Present **NVIDIA**, *Senior Systems Software Engineer*.  
MagLev: Creating infrastructure for autonomous vehicle development
- 2018-2021 **Lawrence Livermore National Laboratory**, *Computer Scientist*.
  - Flux: developing features, supporting users, and readying code for future Exascale system - El Capitan
  - HPC+Cloud: PI of \$175K project to find gaps in LLNL's converged computing capabilities
  - Exaworks: creating a portable, performant, and hardened workflow SDK
  - ZFP: developed python bindings for this LLNL compression library
- 2014-2018 **University of Delaware**, *Graduate Research Assistant*.  
Created next-generation job schedulers for HPC clusters utilizing the Flux resource manager
- 2014-2017 **Lawrence Livermore National Laboratory**, *Research Intern*.  
(Summers) **2014**: Created a discrete event simulator for the Flux resource manager  
**2015**: Designed an automatic job aggregator and hierarchical scheduler, resulting in a 4x speed up over the existing scheduler  
**2016**: Added dynamicity to my hierarchical scheduler to eliminate resource fragmentation  
**2017**: Integrated my hierarchical scheduler with the Uncertainty Quantification Pipeline (UQP), resulting in a 37% improvement in workload runtime
- 2013-2014 **University of Delaware**, *Undergraduate Research Assistant*.  
Applied auto-tuning techniques to improve I/O performance of scientific applications on clusters with a Lustre filesystem
- Summer 2013 **Oak Ridge National Laboratory**, *Research Intern (DOE SULI)*.  
Integrated the ADIOS I/O library into the QMCPack scientific application to improve performance on ORNL's supercomputer, Titan
- 2012-2013 **University of Delaware**, *Undergraduate Research Assistant*.  
Improved the efficiency of GPU accelerated codes in scientific workflows on large-scale clusters
- Summer 2012 **University of Houston**, *Research Intern (NSF REU)*.  
Optimized VolpexMPI library for use on large-scale clusters

---

## Skills

Languages Python, C, C++, Rust, Lua, Java

Parallel MPI, Apache Spark, OpenMP, OpenCL, CUDA  
Cloud Docker, AWS, Kubernetes  
Python pandas, numpy, scikit-learn, matplotlib, sphinx, cffi, cython, pyspark  
Web Development PHP, MySQL, HTML, CSS, Javascript, ColdFusion  
Dev Tools git, autotools, CMake, (z)sh, svn, Latex

---

## Awards & Honors

Fall 2022 Best Paper Award - International Conference on e-Science  
Fall 2021 R&D 100 Award: Flux Next-Generation Workload Management Software Framework - R&D World  
Summer 2019 Spot Award - LLNL ATDM Next Generation Computing Enablement  
Summer 2019 Dissertation Award - IEEE Technical Committee on Scalable Computing  
Summer 2016 Best Poster - Annual LLNL Poster Symposium  
Spring 2016 Outstanding Graduate Student Award - CIS Department  
Spring 2016 HPDC Student Travel Award for \$1000 - HPDC  
Spring 2016 Invited to attend and present at Salishan Conference - LLNL  
Fall 2014 ICPADS Student Travel Award for \$1000 - ICPADS/TCPP  
Spring 2014 Outstanding Senior Student Award - CIS Department  
Fall 2013 Undergraduate Student Research Competition Award - ACM  
Spring 2013 Outstanding Junior Student Award - CIS Department  
Fall 2012 General Honors Award - University of Delaware  
2011 - 2014 Presidential Achievement Scholarship - University of Delaware

---

## Peer-Reviewed Publications

### Conference Papers

- Best Paper** [1] Dong Ahn, Xiaohua Zhang, Jeffrey Mast, **Stephen Herbein**, Francesco Di Natale, Dan Kirshner, Samuel Ade Jacobs, Ian Karlin, Dan Milroy, Bronis De Supinski, Brian Van Essen, Jonathan Allen, and Felice C. Lightstone. Scalable Composition and Analysis Techniques for Massive Scientific Workflows. In *Proceedings of the 18th IEEE International Conference on e-Science*, eScience, Oct 2022.
- [2] Claudia Misale, Daniel J. Milroy, Carlos Eduardo Arango Gutierrez, Maurizio Drocco, **Stephen Herbein**, Dong H. Ahn, Zvonko Kaiser, and Yoonho Park. Towards standard Kubernetes scheduling interfaces for converged computing. In *Smoky Mountains Computational Sciences and Engineering Conference*, SMC2021, 2021.
- [3] Harsh Bhatia, Francesco Di Natale, Joseph Y. Moon, Xiaohua Zhang, Joseph R. Chavez, Fikret Aydin, Chris Stanley, Tomas Opielstrup, Chris Neale, Sara Kokkila Schumacher, Dong H. Ahn, **Stephen Herbein**, Timothy S. Carpenter, Sandrasegaram Gnanakaran, Peer-Timo Bremer, James N. Glosli, Felice C. Lightstone, and Helgi I. Ingólfsson. Generalizable Coordination of Large Multiscale Ensembles: Challenges and Learnings at Scale. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis*, SC, St. Louis, MO, November 2021.
- [4] Michael Wyatt, **Stephen Herbein**, Kathleen Shoga, Todd Gamblin, and Michela Taufer. CanarIO: Sounding the Alarm on IO-Related Performance Degradation. In *Proceedings of the 34th IEEE International Parallel and Distributed Processing Symposium*, IPDPS, New Orleans, LA, May 2020.
- [5] Samuel D. Pollard, Nikhil Jain, **Stephen Herbein**, and Abhinav Bhatele. Evaluation of an Interference-free Node Allocation Policy on Fat-tree Clusters. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis*, SC, pages 333–345, Dallas, TX, November 2018.

- [6] Michael Wyatt, **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*, Eugene, OR, August 2018. (Acceptance Rate: 99/305, 32.5%).
- [7] 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 IEEE International Conference on Cluster Computing, CLUSTER*, pages 255–258, September 2016. (Acceptance Rate: 57/162, 35%).
- [8] **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 International Symposium on High-Performance Parallel and Distributed Computing, HPDC*, pages 69–80, Kyoto, Japan, June 2016. (Acceptance Rate: 20/129, 15.5%).
- [9] **Stephen Herbein**, Ayush Dusia, Aaron Landwehr, Sean McDaniel, Jose Monsalve, Yang Yang, Seetharami R. Seelam, and Michela Taufer. Resource Management for Running HPC Applications in Container Clouds. In *Proceedings of 31st International Supercomputing Conference, ISC*, pages 261–278, Leipzig, Germany, June 2016. (Acceptance Rate: 25/60, 40%).
- [10] 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, Dec 2014. (Acceptance Rate: 96/322, 29.8%).
- [11] **Stephen Herbein**, Michael Matheny, Matthew Wezowicz, Jaron Krogel, Jeremy Logan, Jeongnim H. 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, Dec 2013.
- [12] 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 e-Science, eScience*, pages 116–123, Oct 2013. (Acceptance Rate: 41/98, 41.8%).

### Journal Papers

- [13] **Stephen Herbein**, Tapasya Patki, Dong H Ahn, Sebastian Mobo, Clark Hathaway, Silvina Caíno-Lores, James Corbett, David Domyancic, Thomas RW Scogland, Bronis R de Supinski, and Michela Taufer. An analytical performance model of generalized hierarchical scheduling. *The International Journal of High Performance Computing Applications*, 36(3):289–306, 2022.
- [14] II Michael R Wyatt, **Stephen Herbein**, Todd Gamblin, and Michela Taufer. AI4IO: A suite of AI-based tools for IO-aware scheduling. *The International Journal of High Performance Computing Applications*, 36(3):370–387, 2022.
- [15] 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 Generation Computer Systems*, 110:202 – 213, 2020.
- [16] 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. *International Journal of High Performance Computing and Networking*, 13:105–118, Jan 2019.
- [17] **Stephen Herbein**, Sean McDaniel, Norbert Podhorszki, Jeremy Logan, Scott Klasky, and Michela Taufer. Performance characterization of irregular I/O at the extreme scale. *Parallel Computing*,

51:17 – 36, Jan 2016. Special Issue on Parallel Programming Models and Systems Software for High-End Computing.

- [18] Sam 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 Design Test*, 31(1):40–50, Feb 2014.

### Workshop Papers

- [19] Aymen Al-Saadi, Dong H. Ahn, Yadu Babuji, Kyle Chard, James Corbett, Mihael Hategan, **Stephen Herbein**, Shantenu Jha, Daniel Laney, Andre Merzky, Todd Munson, Michael Salim, Mikhail Titov, Matteo Turilli, Thomas D. Uram, and Justin M. Wozniak. ExaWorks: Workflows for Exascale. In *Proceedings of the 18th IEEE Workshop on Workflows in Support of Large-Scale Science*, (WORKS), pages 50–57, 2021.
- [20] Claudia Misale, Maurizio Drocco, Daniel J. Milroy, Carlos Eduardo Arango Gutierrez, **Stephen Herbein**, Dong H. Ahn, and Yoonho Park. It's a Scheduling Affair: GROMACS in the Cloud with the KubeFlux Scheduler. In *Proceedings of the 3rd International Workshop on Containers and New Orchestration Paradigms for Isolated Environments in HPC*, (CANOPIE-HPC), pages 10–16, 2021.
- [21] **Stephen Herbein**, David Domyancic, Paul Minner, Ignacio Laguna, Rafael Ferreira da Silva, and Dong H. Ahn. MCEM: Multi-Level Cooperative Exception Model for HPC Workflows. In *Proceedings of the 9th International Workshop on Runtime and Operating Systems for Supercomputers*, ROSS, Phoenix, AZ, June 2019.
- [22] Dong H. Ahn, Ned Bass, Albert Chu, Jim Garlick, Mark Grondona, **Stephen Herbein**, Joseph Koning, Tapasya 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 10–19, Dallas, TX, November 2018.
- [23] Robert Searles\*, **Stephen Herbein\***, and Sunita Chandrasekaran. A Portable, High-Level Graph Analytics Framework Targeting Distributed, Heterogeneous Systems. In *Proceedings of the 3rd Workshop on Accelerator Programming Using Directives*, WACCPD, pages 79–88, Salt Lake City, UT, November 2016. \*Equal contribution.
- [24] **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 7th International Workshop on Parallel Programming Models and Systems Software for High-End Computing*, P2S2, pages 292–301, Sept 2014.

### Posters & Presentations

- [25] **Stephen Herbein**. Flux: Solving Exascale Workflow and Resource Challenges; Plus - How Open-Source Drives Our Project Design. In *Free and Open source Software Developers' European Meeting*, FOSDEM'21, Online, February 7 2021.
- Best Poster Candidate** [26] **Stephen Herbein**, Tapasya Patki, Dong H. Ahn, Don Lipari, Tamara Dahlgren, David Domyancic, and Michela Taufer. Fully Hierarchical Scheduling: Paving the Way to Exascale Workloads. In *Proceedings of the 29th ACM/IEEE International Conference for High Performance Computing and Communications*, SC, Denver, CO, November 2017. (Acceptance Rate: 98/169 58.0%).
- [27] **Stephen Herbein**. Exploring the Trade-off Space of Hierarchical Scheduling for Very Large HPC Centers. In *Proceedings of the 27th ACM/IEEE International Conference for High Performance Computing and Communications*, SC, Austin, TX, November 2015.
- [28] Sean McDaniel, **Stephen Herbein**, and Michela Taufer. A Two-Tiered Approach to I/O Quality of Service in Docker Containers. In *Proceedings of the IEEE International Conference on Cluster Computing*, CLUSTER, pages 490–491, September 2015.

- [29] **Stephen Herbein**. Enabling Fine-grained Gathering of Scientific Data in QMCPack Simulations on Titan. In *Proceedings of the 25th ACM/IEEE International Conference for High Performance Computing and Communications*, SC, Denver, CO, November 2013.
- [30] Matthew Wezowicz, Michael Matheny, **Stephen Herbein**, Jeremy Logan, Jeognim Kim, Jaron Krogel, Scott Klasky, and Michela Taufer. Predictions of Large-scale QMCPack I/Os on Titan using Skel. In *Proceedings of the 25th ACM/IEEE International Conference for High Performance Computing and Communications*, SC, Denver, CO, November 2013.

### Theses

- [31] **Stephen Herbein**. *Advanced Schedulers For Next-Generation HPC Systems*. Thesis, University of Delaware, Newark, DE, 2018.
- [32] **Stephen Herbein**. *Benchmarking and Auto-tuning I/O Intensive Applications at the Extreme Scale*. Thesis, University of Delaware, Newark, DE, May 2014.

### Tutorials

- February, 2021 **Using Flux to Overcome Scheduling challenges of Exascale Workflows**,  
*Exascale Computing Project Annual Meeting*.
- February, 2020 **Flux: Using Next-Generation Resource Management and Scheduling Infrastructure for Exascale Workflows**,  
*Exascale Computing Project Annual Meeting*.
- February, 2019 **Flux: Using Next-Generation Resource Management and Scheduling Infrastructure for Exascale Workflows**,  
*Exascale Computing Project Annual Meeting*.
- November, 2018 **Introduction of Practical Approaches to Data Analytics for HPC with Spark**,  
*International Conference for High Performance Computing, Networking, Storage, and Analysis*.

### Invited Talks

- March, 2019 **Flux: Using Next-Generation Resource Management and Scheduling Infrastructure for Exascale Workflows**,  
*University of Tennessee, Knoxville*.
- December, 2015 **Towards Resource QoS In Container Clouds**,  
*IBM Thomas J. Watson Research Center*,  
6th Student Workshop on Cloud and Data Services

### Professional Services & Activities

#### Committee Member

- 2019-2021 Planning Committee: Mobile App - SuperComputing (SC)  
2017 Lead Student Volunteer - Information Booth at SuperComputing (SC)

#### Conference (Sub-)Reviewer

- 2021 SC, IJHPCA  
2020 PASC, HPC Asia, IJHPCA, JPDC  
2018 SC, HPDC, PARCO  
2017 IPDPS  
2016 PARCO, Cluster, COM-HPC, SBAC-PAD  
2015 WoC

#### Other

- 2012-2016 Student Volunteer at SuperComputing (SC)

---

## Memberships

- Association for Computing Machinery (ACM)
- Institute of Electrical and Electronics Engineers (IEEE)

---

## Open-Source Software Projects

- flux-core Next-generation resource manager
- flux-sched Next-generation batch job scheduler
- zfp Compressed floating-point and integer arrays
- BREWCOP BREWCOP is a Raspberry pi that Electronically Weighs COffee Pots