

Robert Underwood

✉ rr.underwood94@gmail.com • 🌐 robertu94.github.io
📄 github.com/robertu94

Education

Clemson University

PhD in Computer Science, GPA 3.92/4.0

Dissertation: Approachable Error Bounded Lossy Compression

Passed Qualifying Exam: May 2018

Co-Advisers: Dr. Amy Apon, Dr. Jon Calhoun, and Dr. Franck Cappello

Clemson, SC

December 2021

Clemson University

Master of Science in Computer Science, GPA 4.0/4.0

Concentration: Systems and Implementation

Clemson, SC

August 2018

Clemson University, Calhoun Honors College

Bachelor of Science, Summa Cum Laude in Computer Science, GPA 4.0/4.0

Honors Thesis: Automation in the Classroom, Adviser: Dr. Jacob Sorber

Clemson, SC

December 2016

Publications

- [1] Zizhe Jian, Sheng Di, Jinyang Liu, Kai Zhao, Xin Liang, Haiying Xu, Robert Underwood, Shixun Wu, and Zizong Chen. “CliZ: Optimizing Lossy Compression for Climate Datasets with Adaptive Fine-tuned Data Prediction”. In: 38th IEEE International Parallel & Distributed Processing Symposium. **co-Author, To Appear**. IEEE. May 2024. DOI: TBD.
- [2] Avinash Maurya, Robert Underwood, M Mustafa Rafique, Franck Cappello, and Bogdan Nicolae. “DataStates-LLM: Lazy Asynchronous Checkpointing for Large Language Models”. In: International Symposium on High-Performance Parallel and Distributed Computing. **co-Author, To Appear**. ACM. May 2024. DOI: TBD.
- [3] Robert Underwood, Meghana Madhyastha, Randal Burns, and Bogdan Nicolae. “Evostore: Towards Scalable Storage of Evolving Learning Models”. In: International Symposium on High-Performance Parallel and Distributed Computing. **First Author, To Appear**. ACM. May 2024. DOI: TBD.
- [4] Grant Wilkins, Sheng Di, Jon Calhoun, Kibaek Kim, Robert Underwood, Richard Mortier, and Franck Cappello. “FedSZ: Leveraging Floating-Point Lossy Compression for Federated Learning Communications”. In: 44th IEEE International Conference on Distributed Computing Systems. **co-Author, To Appear**. IEEE. May 2024. DOI: TBD.
- [5] Robert Underwood, Meghana Madhyastha, Randal Burns, and Bogdan Nicolae. “Understanding Patterns of Deep Learning Model Evolution in Network Architecture Search”. In: 30th IEEE International Conference on High Performance Computing, Data, and Analytics. **First Author**. IEEE. Dec. 2023. DOI: 10.1109/HiPC58850.2023.00025.

- [6] Arkapraba Ganguli, Robert Underwood, Julie Bessac, David Krasowska, Jon C. Calhoun, Sheng Di, and Franck Cappello. "A Lightweight, Effective Compressibility Estimation Method for Error-bounded Lossy Compression". In: *2023 IEEE International Conference on Cluster Computing*. **First two authors contributed equally**. ACM. Oct. 2023. DOI: 10.1109/CLUSTER52292.2023.00028.
- [7] Xin Liang, Kai Zhao, Sheng Di, Sihuan Li, Robert Underwood, Ali M. Gok, Jiannan Tian, Junjing Deng, Jon C. Calhoun, Dingwen Tao, Zizhong Chen, and Franck Cappello. "SZ3: A Modular Framework for Composing Prediction-Based Error-Bounded Lossy Compressors". In: *IEEE Transactions on Big Data* 9.2 (Aug. 2023). **Co-author**, pp. 485–498. DOI: 10.1109/TBDATA.2022.3201176.
- [8] Robert Underwood, Chunhong Yoon, Ali Gok, Sheng Di, and Franck Cappello. "ROIBIN-SZ: Fast and Science-Preserving Compression for Serial Crystallography". In: *Syncrotron Radiation News* (July 2023). **Invited, Not Peer Reviewed**. DOI: 10.1080/08940886.2023.2245722.
- [9] Meghana Madhyastha, Robert Underwood, Randal Burns, and Bogdan Nicolae. "DStore: A Lightweight Scalable Learning Model Repository with Fine-Grain Tensor-Level Access". In: *Proceedings of the 37th International Conference on Supercomputing*. **Co-Author**. June 2023, pp. 133–143. DOI: 10.1145/3577193.3593730.
- [10] Robert Underwood, Julie Bessac, David Krasowska, Jon C Calhoun, Sheng Di, and Franck Cappello. "Black-box statistical prediction of lossy compression ratios for scientific data". In: *The International Journal of High Performance Computing Applications* (June 2023). **Published, First two authors contributed equally**, p. 10943420231179417. DOI: 10.1177/10943420231179417.
- [11] Robert Underwood and Bogdan Nicolae. "MPIGDB: A Flexible Debugging Infrastructure for MPI Programs". In: *13th Workshop on AI and Scientific Computing at Scale using Flexible Computing Infrastructures*. **Presented at HPDC/FCRC-23**. ACM. June 2023. DOI: 10.1145/3589013.3596675.
- [12] Julie Bessac, Robert Underwood, and Sheng Di. "Discussion on "Saving Storage in Climate Ensembles: A Model-Based Stochastic Approach"". In: *Journal of Agricultural, Biological and Environmental Statistics* 28.2 (May 2023). **Published**, pp. 358–364. DOI: 10.1007/s13253-023-00540-7.
- [13] Robert Underwood, Sheng Di, Jon C. Calhoun, Amy Apon, and Franck Cappello. *OptZConfig: Using Optimization Configure Error Bounded Lossy Compressors*. **First Author**. IEEE, Dec. 2022. DOI: 10.1109/TPDS.2022.3154096.
- [14] Robert Underwood, Julie Bessac, Sheng Di, and Franck Cappello. "Understanding the Effects of Modern Compressors on the Community Earth Science Model". In: *2022 IEEE/ACM 8th International Workshop on Data Analysis and Reduction for Big Scientific Data (DRBSD)*. **First Author, Best Paper**. Nov. 2022, pp. 1–10. DOI: 10.1109/DRBSD56682.2022.00006.

- [15] Julie Bessac, David Krasowska, Jon C. Calhoun, Robert Underwood, Sheng Di, and Franck Cappello. "Exploring Lossy Compressibility through Statistical Correlations of Scientific Datasets". In: *The 7th International Workshop on Data Analysis and Reduction for Big Scientific Data*. Co-Author. IEEE. Nov. 2021. doi: 10.1109/DRBSD754563.2021.00011.
- [16] Robert Underwood, Victorina Malvoso, Sheng Di, Jon C. Calhoun, Amy Apon, and Franck Cappello. "Productive and Performant Generic Lossy Data Compression with LibPressio". In: *The 7th International Workshop on Data Analysis and Reduction for Big Scientific Data*. Presented at SC21. IEEE. Nov. 2021. doi: 10.1109/DRBSD754563.2021.00005.
- [17] Dakota Fulp, Alexandra Poulos, Robert Underwood, and Jon C. Calhoun. "ARC: An Automated Approach to Resiliency for Lossy Compressed Data via Error Correcting Codes". In: *Proceedings of 30th International ACM Symposium on High-Performance Parallel and Distributed Computing*. Co-Author. ACM. June 2021. doi: 10.1145/3431379.3460638.
- [18] Jiannan Tian, Sheng Di, Kai Zhao, Cody Rivera, Megan Hickman, Robert Underwood, Sian Jin, Xin Liang, Jon Calhoun, Dingwen Tao, and Franck Cappello. "cuSZ: An Efficient GPU Based Error-Bounded Lossy Compression Framework for Scientific Data". In: *Proceedings of 29th International Conference on Parallel Architectures and Compilation Techniques*. Co-Author. ACM. Atlanta, Georgia (virtual), Oct. 2020. doi: 10.1145/3410463.3414624.
- [19] Robert Underwood, Sheng Di, Jon C. Calhoun, and Franck Cappello. "FRaZ: A Generic High-Fidelity Fixed-Ratio Lossy Compression Framework for Scientific Floating-point Data". In: *proceedings of the 9th international conference on performance engineering*. Presented virtually at IPDPS 2020. IEEE. New Orleans, Louisiana (virtual), May 2020, pp. 1–11. doi: 10.1109/IPDPS47924.2020.00065.
- [20] Robert Underwood, Jason Anderson, and Amy Apon. "Measuring Network Latency Variation Impacts to High Performance Computing Application Performance". In: *Proceedings of the 9th International Conference on Performance Engineering*. Presented at ICPE 2018. ACM/SPEC. Berlin, Germany, Apr. 2018, pp. 1–12. doi: 10.1145/3184407.3184427.

Peer-Reviewed Academic Poster Presentations

Statistical Prediction of Lossy Compression Ratios for 3D Scientific Data Dallas, Texas
Supercomputing 2022, Best Poster ACM Student Research Competition November 2022

David Krasowska, **Robert Underwood**, Julie Bessac, Sheng Di, Frank Cappello

LibPressio: A Unifying Data Compression Interface for Users and Developers Virtual
United States Department of Energy Data Days May 2022

Robert Underwood, Sheng Di, and Franck Cappello

Approachable Error Bounded Lossy Compression Virtual
Supercomputing 2021 November 2021

Robert Underwood

Research Experience

Argonne National Laboratory

Lemont, IL

Post Doctoral Appointee

2022-

- Researched applications and approaches of lossy compression to ensure data integrity.
- Researched data movement and checkpointing systems

Clemson University

Clemson, SC

Clemson Data Intensive Computing Environments

2016-2021

- Researched and modeled reliability and performance of applications using lossy compression
- Developed techniques to understand the impacts of lossy compression on AI applications
- Designed experiments to analyze performance of high performance computing systems
- Designed models to understand and improve the reliability of computer infrastructure
- Researched trade-offs of compressor agnostic tooling for scientific applications

Argonne National Laboratory

Lemont, IL

Under Dr. Franck Cappello

Summer-Fall 2019

- Researched the design of optimization based techniques for enforcing user-level error bounds
- Designed and implemented LibPressio – a generic abstraction between compression libraries
- Contributed to the design and implementation of SZ – a lossy compression framework – for CPUs and GPUs

Clemson University

Clemson, SC

Clemson PERSIST Lab

2015-2016

- Designed and developed an automated grading framework using Python, C, Raspberry Pi, and Docker
- System used modular design, supports process isolation, and multiple test formats

Significant Software

LibPressio

<https://github.com/robertu94/libpressio>

2019-present

- High-performance generic abstraction for compression of dense tensors
- Supports 45+ of plugins for compressors and analysis in collaboration with 17 institutions world-wide
- ~200 unique monthly downloads
- Significant plugins include: LibPressio-Opt (automatic configuration of compression), a parallel compression runtime, and the external metrics and compressors framework
- Significant integrations include: Python bindings, HDF5-filters, R bindings, ADIOS2, Spack, Z-checker

SZ

<https://szcompressor.org/>

2019-present

- One of the leading open and transparent Lossy Compression Frameworks for scientific data
- the SZ framework is a R&D 500 Award Winner for 2021
- Contributed an early design of SZ for GPUs and the modular SZ-3 and SZ-4
- Implemented the python bindings for SZ

mpigdb

<https://github.com/robertu94/mpigdb>

2023-present

- Open Source Parallel Debugger for MPI based codes
- Exposes capabilities of underlying GDB with extensions for MPI

Teaching and Mentoring Experience

Argonne National Laboratory

Mentoring

2022-present

- (2023) Co-Mentored Isita Talukdar (Undergraduate Electrical Engineering and Computer Science at University of California Berkley) with Amrajit Singh and Kento Sato of Riken Center for Computational Science on aspects of software development, debugging, and research communication in the context of integrating LibPressio with TeZIP (a complex, heterogenous, AI-based, compressor implemented in Python)
- (2023) Mentored Alexandra Poulos (PhD Candidate at Clemson University) on aspects of software development, debugging, and research communication in the context of performing sensitivity analysis on compression estimation techniques.
- (2022) Co-Mentored David Krasowska (Undergraduate Computer Engineering at Clemson University) with Julie Bessac on aspects of software development, research communication, and how to quickly understand large code bases in the context of extending a compressor agnostic predictor compress-ability to 3d data sets. David took 1st in the ACM student poster competition at SC22 on a project I mentored him on and supervised the writing of a workshop paper and journal article. He is now a graduate student at Northwestern University under Peter Dinda.
- (2022) Co-Mentored Arkaprabha Ganguli (PhD Candidate in Statistics at Michigan State University) with Julie Bessac on how to design and implementation of statistical measures to estimate lossy compress-ability without compressors. He is now a Postdoc at Argonne National Laboratory.
- (2022-2023) Co-Mentored Meghana Madhyastha (PhD Candidate in Computer Science at Johns Hopkins University) with Bogdan Nicolae on various aspects of the Datastates project from designing lock-free concurrent data structures to accessing data pointers on GPUs, to navigating subtleties using HPC machines.

Clemson University

Mentoring

Clemson, SC

Summer 2021

- Mentored one female and one male, undergraduate student on projects that led to two ACM student research poster submissions and later journal submission.
- Provided training on git, python, C++, lossy compression, and scientific experiment design

Clemson University

CPSC/ECE 3220: Operating Systems

Clemson, SC

Fall 2018

- Graduate Teacher of Record, produced all lectures and most materials
- Junior/Senior level course - 50 Students enrolled, Completed (78%), Course GPA (2.42)
- Course materials <https://robertu94.github.io/cpsc3220-f18/>
- Anonymous Student Assessment Responses:
 - Response Rate (92.3%), Would Recommend (72.2%)
 - Median Results: Effective Instructor (4/5), Helpful Feedback (4/5), Relative Difficulty (5/5)
 - Selected Student Comments:
 - “Definitely. One of the best professors I’ve had at Clemson.”
 - “Yes. He is very knowledgeable [sic] and works very hard to impart that knowledge to others.”
 - “Yes, it is obvious that Mr. Underwood is passionate about operating systems and is extensively knowledgeable about computer science in general. This course felt overwhelming at times, but I definitely learned a lot through it.”

Relevant Coursework

Clemson University **Clemson, SC**
EES 883: Resilient Infrastructure Systems *Spring 2018*

- Prepared a NSF grant proposal submitted by my adviser to NSF and funded by NSF
- Constructed and quantified uncertainty in a queuing theory and population based model of Infrastructure systems
- Designed experiments for statistical model validation

Clemson University **Clemson, SC**
CPSC 827: Language Translation *Fall 2016*

- Implemented a subset of Python from a yacc-able version of the full Python 2.7 grammar in C++, flex, and bison
- Included: a.s.t. generator; type system; function, global, nested, and returning scope; and primitive exceptions
- Designed and implemented using Object Oriented principals with 55 classes, over 3600 SLOC, in less than 2 months

Clemson University **Clemson, SC**
CPSC 820: Parallel Architecture *Fall 2016*

- Researched and presented on the design and implementation of Linux Bridge, OpenVSwitch, DPDK, SRIOV, and MACVLAN
- Designed and conducted experiment to quantify latency variation in RDMA using InfiniBand layers 1, 2, and 4

Clemson University **Clemson, SC**
CPSC 822: Case Study in Operating Systems: Linux *Spring 2016*

- Designed and developed:
 - Graphics driver for an AMD Radeon-like device with frame buffer, fifo, and dma interfaces
 - System call to unconditionally kill a process
 - Disk scheduler for a SCSI disk controller
- Analyzed and debugged performance issues in the Linux kernel
- Worked with a complex system with limited documentation

Clemson University **Clemson, SC**
CPSC 840: Design and Analysis of Algorithms *Spring 2016*

- Analyzed and designed amortized, randomized, and approximation algorithms to solve problems.
- Designed time and space efficient data structures

Work Experience

The Boeing Company

Information Technology Intern

Charleston, SC

Summer 2016, 2017

- Developed improvements for a web based portal system in HTML, Python, and JavaScript
- Developed the user interface for a materials database using HTML and JavaScript
- Designed, developed, and led development on a resource management tool using C#, HTML, and JavaScript.
- Worked on the Network Automation, Tooling, and Standards Integration Team

Unitrends, Inc

Software Development Intern

Columbia, SC

2014-2016

- Developed GPU offloading for AES encryption using Nvidia CUDA.
- Designed and developed automated configuration scripts for testing environments using Ansible.
- Designed and developed new cloud infrastructure using LVM, Linux, and Docker
- Designed and developed a Dynamic Alert System in Python
- Worked on the Alerts System in PHP, BASH, C, PERL, SQL
- Worked on the internal Customer Incident Analysis web portal using Django, Postgresql, HTML, CSS, and JavaScript

Computer Skills

Advanced: Bash, Bourne Shell, C, C++, Docker, Linux Kernel and Userspace, Python, Vim

Intermediate: Ansible, Cuda, Git, Hadoop, JAVA, JavaScript, Keras, HDF5, Julia, L^AT_EX, Linux Profiling (perf), LLVM-libtooling, MPI, OpenCL, OpenMP, SQL, SaltStack, SciKit Learn, Spack, Systemd, Tensorflow 2.x, SWIG

Basic: Apache Spark, ARM assembly, CSS3 C#, Haskell, HTML5, Perl, Puppet, RCpp/RInside, Rust, SNMP, SVN, FreeBSD, PHP

Professional Affiliations

Association for Computing Machinery: Professional Member 2022-Present

Joint Laboratory for Extreme Scale Computing: 2018-Present

Association for Computing Machinery: Student Member 2014-2021

Professional Service

JLESC Early Career Ambassador for Argonne National Laboratory: 2023

Session Chair: JLESC 2022, JLESC 2023

Reviewer: ICPE 2017, ICCCN 2017, PABS 2017, SC17, IEEE CLOUD 2018, IEEE TSE 2018, IPDPS 2018, IPDPS 2019, IEEE CLUSTER 2020, Sustainable Computing 2021, IEEE Big Data 2021, PacificVis 2022, HPDC 2023, IEEE TC 2023

Volunteer: SC2018, SC21

Honors

- Clemson Outstanding Ph.D. in Computer Science Award, 2021
- Graduate Student Research Lighting Talk Competition Faculty Award, 2020
- Department of Energy Office of Science Graduate Student Research Award, 2019
- Fellowship, National Research Traineeship: Resilient Infrastructure Systems 2017-2020
- National Science Foundation Graduate Research Fellowship Honorable Mention 2017
- Faculty Scholarship Award, Clemson University 2016
- Benefitfocus Scholarship 2015-2016
- McAlister Scholarship 2015-2016
- Palmetto Fellows Recipient 2013-2016
- President's List at Clemson University 2013-2016
- Outstanding Sophomore in Computer Science at Clemson University 2015
- Order of the Arrow, Vigil Honor 2013
- Eagle Scout 2010