Lipeng Wan

Computer Science and Mathematics Division
Oak Ridge National Laboratory
Oak Ridge, TN, 37830, USA

Phone: (865) 919-0473
Email: wanl@ornl.gov or lipengwan86@gmail.com
Linkedin: https://www.linkedin.com/in/lipengwan/

Education

University of Tennessee, Knoxville, TN, USA — Aug. 2011 - May. 2016
Ph.D., Computer Science

Southeast University, Nanjing, Jiangsu, China — Sep. 2008 - Mar. 2011
Master of Engineering, Information and Communication Engineering

Nanjing University of Science and Technology, Nanjing, Jiangsu, China — Sep. 2004 - Jun. 2008
Bachelor of Engineering, Communication Engineering

Research Interests

Big Data Management and Analytics, High-Performance Computing, Scientific Machine Learning

Work Experiences

Computer Scientist, Oak Ridge National Laboratory — Nov. 2018 - present
Postdoctoral Research Associate, Oak Ridge National Laboratory — July. 2016 - Oct. 2018
Graduate Research Assistant, University of Tennessee, Knoxville — Aug. 2011 - May. 2016
Student Intern, Oak Ridge National Laboratory — May. 2015 - Aug. 2015
Student Intern, Oak Ridge National Laboratory — May. 2014 - Aug. 2014

Teaching Experiences

Teaching Assistant, CS 560 (Big Data Software Systems), University of Tennessee, Knoxville

Teaching Assistant, CS 361 (Operating Systems), University of Tennessee, Knoxville

Teaching Assistant, ECE 451 (Computer Systems Architecture), University of Tennessee, Knoxville

Research Grants

Current

- "Enabling Scalable Analytics Using Multiprecision Refactoring"
  - Sponsor: Oak Ridge National Laboratory
  - Amount: $ 640,000
  - Award Period: Oct. 2020 - Sep. 2022
  - Role: Lead PI
- "SIRIUS 2: Science-driven Data Management for Multi-tiered Storage 2.0"
  - Sponsor: U.S. Department of Energy

  - Amount: $ 2,750,000
  - Award Period: Oct. 2020 - Sep. 2025
  - Lead PI: Scott Klasky (Oak Ridge National Laboratory)
  - Role: Senior Personnel
- "CODAR: Co-Design for Online Data Analysis and Reduction at the Exascale"
  - Sponsor: U.S. Department of Energy
  - Amount: $ 4,000,000
  - Award Period: Oct. 2016 - Sep. 2023
  - Lead PI: Ian Foster (Argonne National Laboratory)
  - Role: Personnel
- "Storage System and I/O risk mitigation for ADIOS on Exascale Machine"
  - Sponsor: U.S. Department of Energy
  - Amount: $ 250,000
  - Award Period: Oct. 2020 - Sep. 2021
  - Lead PI: Scott Klasky (Oak Ridge National Laboratory)
  - Role: Personnel

Publications

Journal Papers

1. **Lipeng Wan**, Axel Huebl, Junmin Gu, Franz Poeschel, Ana Gainaru, Ruonan Wang, Jieyang Chen, Xin Liang, Dmitry Ganyushin, Todd Munson, Ian Foster, Jean-Luc Vay, Norbert Podhorszki, Kesheng Wu, and Scott Klasky. Improving I/O Performance for Exascale Applications through Online Data Layout Reorganization. *IEEE Transactions on Parallel and Distributed Systems (***TPDS***)*, 2021.
2. Xin Liang, Ben Whitney, Jieyang Chen, **Lipeng Wan**, Qing Liu, Dingwen Tao, James Kress, Dave Pugmire, Matthew Wolf, Norbert Podhorszki and Scott Klasky. MGARD+: Optimizing Multilevel Methods for Error-Bounded Scientific Data Reduction. *IEEE Transactions on Computers (***TC***)*, 2021.
3. Eric Suchyta, Scott Klasky, Norbert Podhorszki, Matthew Wolf, Abolaji Adesoji, CS Chang, Jong Youl Choi, Philip E. Davis, Julien Dominski, Stéphane Ethier, Ian Foster, Kai Germaschewski, Berk Geveci, Chris Harris, Kevin A Huck, Qing Liu, Jeremy Logan, Kshitij Mehta, Gabriele Merlo, Shirley V Moore, Todd Munson, Manish Parashar, David Pugmire, Mark S Shephard, Cameron W Smith, Pradeep Subedi, **Lipeng Wan**, Ruonan Wang and Shuangxi Zhang. The Exascale Framework for High Fidelity Coupled Simulations (EFFIS): Enabling Whole Device Modeling in Fusion Science. *The International Journal of High Performance Computing Applications*, 2021.
4. Ian Foster, Mark Ainsworth, Julie Bessac, Franck Cappello, Jong Youl Choi, Sheng Di, Zichao Di, Ali M Gok, Hanqi Guo, Kevin A Huck, Christopher Kelly, Scott Klasky, Kerstin Kleese van Dam, Xin Liang, Kshitij Mehta, Manish Parashar, Tom Peterka, Line Pouchard, Tong Shu, Ozan Tugluk, Hubertus van Dam, **Lipeng Wan**, Matthew Wolf, Justin M Wozniak, Wei Xu, Igor Yakushin, Shinjae Yoo and Todd Munson. Online Data Analysis and Reduction: An Important Co-Design Motif for Extreme-Scale Computers. *The International Journal of High Performance Computing Applications*, 2020.
5. William F.Godoy, Norbert Podhorszki, Ruonan Wang, Chuck Atkins, Greg Eisenhauer, Junmin Gu, Philip Davis, Jong Youl Choi, Kai Germaschewski, Kevin Huck, Axel Huebl, Mark Kim, James Kress, Tahsin, Kurc, Qing Liu, Jeremy Logan, Kshitij Mehta, George Ostrouchov, Manish Parashar, Franz Poeschel, David Pugmire, Eric Suchyta, Keichi Takahashi, Nick Thompson, Seiji Tsutsumi, **Lipeng Wan**, Matthew Wolf, Kesheng Wu and Scott Klasky. ADIOS 2: The Adaptable Input Output System. A Framework for High-Performance Data Management. *SoftwareX*, 2020.

6. Jeremy Logan, Mark Ainsworth ,Chuck Atkins, Jieyang Chen, Jong Youl Choi, Junmin Gu, James Kress, Greg Eisenhauer, Berk Geveci, William Godoy, Mark Kim, Tahsin Kurc, Qing Liu, Kshitij Mehta, George Ostrouchov, Norbert Podhorzski, David Pugmire, Eric Suchyta, Nicolas Thompson, Ozan Tugluk, **Lipeng Wan**, Ruonan Wang, Ben Whitney, Matthew Wolf, Kesheng Wu and Scott Klasky. Extending the Publish/Subscribe Abstraction for High-Performance I/O and Data Management at Extreme Scale. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*, 2020.
7. **Lipeng Wan**, Zhibo Wang, Zheng Lu, Yunhe Feng, Hairong Qi, Wenjun Zhou and Qing Cao. Approximate and Sublinear Spatial Queries for Large-Scale Vehicle Networks. *IEEE Transactions on Vehicular Technology*, 2018.
8. **Lipeng Wan**, Qing Cao, Feiyi Wang and Sarp Oral. Optimizing Checkpoint Data Placement with Guaranteed Burst Buffer Endurance in Large-Scale Hierarchical Storage Systems. *Journal of Parallel and Distributed Computing*, 2017.
9. **Lipeng Wan**, Qing Cao and Wenjun Zhou. Optimizing the Performance of Sensor Network Programs through Estimation-Based Code Profiling. *Pervasive and Mobile Computing*, 2017.
10. Qing Cao, Yunhe Feng, Zheng Lu, Hairong Qi, Leon M Tolbert, **Lipeng Wan**, Zhibo Wang and Wenjun Zhou. Approximate Cardinality Estimation (ACE) in Large-Scale Internet of Things deployments. *Ad Hoc Networks*, 2017.
11. Jilong Liao, Zhibo Wang, **Lipeng Wan**, Qing Cao and Hairong Qi. Smart Diary: A Smartphone-Based Framework for Sensing, Inferring and Logging Users' Daily Life. *IEEE Sensors Journal*, 2014.

Conference Papers

1. Xin Liang, Qian Gong, Jieyang Chen, Ben Whitney, **Lipeng Wan**, Qing Liu, David Pugmire, Rick Archibald, Norbert Podhorszki and Scott Klasky. Error-Controlled, Progressive, and Adaptable Retrieval of Scientific Data with Multilevel Decomposition. *International Conference for High Performance Computing, Networking, Storage and Analysis (***SC***)*, 2021.
2. Xinying Wang, **Lipeng Wan**, Jieyang Chen, Qian Gong, Ben Whitney, Jinzhen Wang, Ana Gainaru, Qing Liu, Norbert Podhorszki, Dongfang Zhao, Feng Yan, Scott Klasky. Unbalanced Parallel I/O: An Often-Neglected Side Effect of Lossy Scientific Data Compression. *IEEE/ACM The 7th International Workshop on Data Analysis and Reduction for Big Scientific Data (***DRBSD***)*, 2021.
3. Qian Gong, Xin Liang, Ben Whitney, Jong Youl Choi, Jieyang Chen, **Lipeng Wan**, Stephane Ethier, Seung-Hoe Ku, Michael Churchill, Choong-Seock Chang, Mark Ainsworth, Ozan Tugluk, Todd Munson, David Pugmire, Richard Archibald and Scott Klasky. Maintaining Trust in Reduction: Preserving the Accuracy of Quantities of Interest for Lossy Compression. *Smoky Mountains Computational Sciences and Engineering Conference (***SMC***)*, 2021.
4. Franz Poeschel, William F Godoy, Norbert Podhorszki, Scott Klasky, Greg Eisenhauer, Philip E Davis, **Lipeng Wan**, Ana Gainaru, Junmin Gu, Fabian Koller, Rene Widera, Michael Bussmann and Axel Huebl. Transitioning from File-Based HPC Workflows to Streaming Data Pipelines with openPMD and ADIOS2. *Smoky Mountains Computational Sciences and Engineering Conference (***SMC***)*, 2021.
5. David Pugmire, James Kress, Jieyang Chen, Hank Childs, Jong Youl Choi, Dmitry Ganyushin, Berk Geveci, Mark Kim, Scott Klasky, Xin Liang, Jeremy Logan, Nicole Marsaglia, Kshitij Mehta, Norbert Podhorszki, Caitlin Ross, Eric Suchyta, Nick Thompson, Steven Walton, **Lipeng Wan** and Matthew Wolf. Visualization as a Service for Scientific Data. *Smoky Mountains Computational Sciences and Engineering Conference (***SMC***)*, 2021.
6. Jieyang Chen, **Lipeng Wan**, Xin Liang, Ben Whitney, Qing Liu, David Pugmire, Nicholas Thompson, Jong Youl Choi, Matthew Wolf, Todd Munson, Ian Foster and Scott Klasky. Accelerating Multigrid-based Hierarchical ScientificData Refactoring on GPUs. *35th IEEE International Parallel & Distributed Processing Symposium (***IPDPS***)*, 2021.
7. **Lipeng Wan**, Matthew Wolf, Feiyi Wang, Jong Youl Choi, George Ostrouchov, Jieyang Chen, Norbert Podhorszki, Jeremy Logan, Kshitij Mehta, Scott Klasky and Dave Pugmire. I/O Performance Characterization and Prediction through Machine Learning on HPC Systems. *Cray User Group (***CUG***)*, 2020.
8. Bingbing Li, Santosh Pandey, Haowen Fang, Yanjun Lyv, Ji Li, Jieyang Chen, Mimi Xie, **Lipeng Wan**, Hang Liu and Caiwen Ding. FTRANS: Energy-Efficient Acceleration of Transformers using FPGA. *ACM/IEEE International Symposium on Low Power Electronics and Design (***ISLPED***)*, 2020.

9. Subhendu Behera, **Lipeng Wan**, Frank Mueller, Matthew Wolf and Scott Klasky. Orchestrating Fault Prediction with Live Migration and Checkpointing. *Proceedings of the 29th International Symposium on High-Performance Parallel and Distributed Computing (***HPDC***)*, 2020.
10. **Lipeng Wan**, Kshitij V. Mehta, Scott A. Klasky, Matthew Wolf, H. Y. Wang, W. H. Wang, J. C. Li and Zhihong Lin. Data Management Challenges of Exascale Scientific Simulations: A Case Study with the Gyrokinetic Toroidal Code and ADIOS. *The 10th International Conference on Computational Methods (***ICCM***)*, 2019.
11. Jieyang Chen, David Pugmire, Matthew Wolf, Nicholas Thompson, Jeremy Logan, Kshitij Mehta, **Lipeng Wan**, Jong Youl Choi, Ben Whitney and Scott Klasky. Understanding Performance-Quality Trade-offs in Scientific Visualization Workflows with Lossy Compression. *IEEE/ACM The 5th International Workshop on Data Analysis and Reduction for Big Scientific Data (***DRBSD***)*, 2019.
12. Jong Youl Choi, Jeremy Logan, Kshitij Mehta, Eric Suchyta, William Godoy, Nicholas Thompson, **Lipeng Wan**, Jieyang Chen, Norbert Podhorszki, Matthew Wolf, Scott Klasky, Julien Dominski and Choong-Seock Chang. A Co-Design Study of Fusion Whole Device Modeling Using Code Coupling. *IEEE/ACM The 5th International Workshop on Data Analysis and Reduction for Big Scientific Data (***DRBSD***)*, 2019.
13. Scott Klasky, Matthew Wolf, Mark Ainsworth, Chuck Atkins, Jong Choi, Greg Eisenhauer, Berk Geveci, William Godoy, Mark Kim, James Kress, Tahsin Kurc, Qing Liu, Jeremy Logan, Arthur B Maccabe, Kshitij Mehta, George Ostrouchov, Manish Parashar, Norbert Podhorszki, David Pugmire, Eric Suchyta, **Lipeng Wan** and Ruonan Wang. A View from ORNL: Scientific Data Research Opportunities in the Big Data Age. *IEEE 38th International Conference on Distributed Computing Systems (***ICDCS***)*, 2018.
14. **Lipeng Wan**, Matthew Wolf, Feiyi Wang, Jong Youl Choi, George Ostrouchov and Scott Klasky. Analysis and Modeling of the End-to-End I/O Performance on OLCF's Titan Supercomputer. *IEEE 19th International Conference on High Performance Computing and Communications (***HPCC***)*, 2017 (**Best Paper Finalist Award**).
15. **Lipeng Wan**, Matthew Wolf, Feiyi Wang, Jong Youl Choi, George Ostrouchov and Scott Klasky. Comprehensive Measurement and Analysis of the User-Perceived I/O Performance in a Production Leadership-Class Storage System. *IEEE 37th International Conference on Distributed Computing Systems (***ICDCS***)*, 2017.
16. Scott Klasky, Eric Suchyta, Mark Ainsworth, Qing Liu, Ben Whitney, Matthew Wolf, Jong Choi, Ian Foster, Mark Kim, Jeremy Logan, Kshitij Mehta, Todd Munson, George Ostrouchov, Manish Parashar, Norbert Podhorszki, David Pugmire and **Lipeng Wan**. Exacution: Enhancing Scientific Data Management for Exascale. *IEEE 37th International Conference on Distributed Computing Systems (***ICDCS***)*, 2017.
17. Jeremy Logan, Jong Youl Choi, Matthew Wolf, George Ostrouchov, **Lipeng Wan**, Norbert Podhorszki, William Godoy, Scott Klasky, Erich Lohrmann, Greg Eisenhauer, Chad Wood and Kevin Huck. Extending Skel to Support the Development and Optimization of Next Generation I/O Systems. *IEEE 37th International Conference on Cluster Computing (***CLUSTER***)*, 2017.
18. **Lipeng Wan**, Feiyi Wang, Sarp Oral, Devesh Tiwari, Sudharshan S. Vazhkudai and Qing Cao. A Practical Approach to Reconciling Availability, Performance, and Capacity in Provisioning Extreme-scale Storage Systems. *International Conference for High Performance Computing, Networking, Storage and Analysis (***SC***)*, 2015.
19. **Lipeng Wan**, Zheng Lu, Qing Cao, Feiyi Wang, Sarp Oral and Bradley Settlemyer. SSD-Optimized Workload Placement with Adaptive Learning and Classification in HPC Environments. *30st International Conference on Massive Storage Systems and Technology (***MSST***)*, 2014.
20. **Lipeng Wan**, and Qing Cao. Towards Instruction Level Record and Replay of Sensor Network Applications. *IEEE 21st International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (***MASCOTS***)*, 2013.
21. Yanjun Yao, **Lipeng Wan**, Qing Cao, and Rukun Mao. PhoneCon: Voice-driven SmartPhone Controllable Wireless Sensor Networks. *IEEE 31st International Performance Computing and Communications Conference (***IPCCC***)*, 2012.

Book Chapters

1. Yanjun Yao, **Lipeng Wan** and Qing Cao. System Architecture and Operating Systems for Wireless Sensor Networks. *The Art of Senosr Networks, Springer*, 2012.

Posters

1. Subhendu Behera, **Lipeng Wan**, Frank Mueller, Matthew Wolf and Scott Klasky. An Adaptive Checkpoint Model for Large-Scale HPC Systems. *International Conference for High Performance Computing, Networking, Storage and Analysis (***SC***)*, 2019.
2. **Lipeng Wan**, Qing Cao, and Wenjun Zhou. Estimation-based Profiling for Code Placement Optimization in Sensor Network Programs. *IEEE International Symposium on Performance Analysis of Systems and Software (***ISPASS***)*, 2015.
3. **Lipeng Wan**, Zhibo Wang, Zheng Lu, Hairong Qi, Wenjun Zhou, and Qing Cao. Towards Approximate Spatial Queries for Large-scale Vehicle Networks. *ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (***SIGSPATIAL GIS***)*, 2014.

Professional Service

Journal Reviewer

IEEE Transactions on Parallel and Distributed Systems

ACM Transactions on Storage

Journal of Parallel and Distributed Computing

IEEE Access

Parallel Computing

Concurrency and Computation: Practice & Experience

IEEE Wireless Communications Magazine

Peer-to-Peer Networking and Applications

Pervasive and Mobile Computing

Expert Systems with Applications

Statistical Analysis and Data Mining

KSII Transactions on Internet and Information Systems

International Journal of Distributed Sensor Networks

Conference Technical Program Committee

Workshop on Fault Tolerance for HPC at eXtreme Scale, 2021

IEEE International Conference on High Performance Computing and Communications, 2020, 2021

IEEE Global Communications Conference: Communication & Information Systems Security, 2019, 2020

Awards and Honors

Exceptional Contribution Award, Oak Ridge National Laboratory, 2021

Outstanding Contribution Award, Oak Ridge National Laboratory, 2019

Best Paper Finalist Award, IEEE 19th International Conference on High Performance Computing and Communications (**HPCC**), 2017

Outstanding Post-Doc Award, Oak Ridge National Laboratory, 2017