Jens **Domke**

Supercomputing Performance Research Team · Team Principal

jens.domke@riken.jp | https://domke.gitlab.io/ | jdomke | domke | jens-domke-7a346a253 | @SPRT_RCCS | Jens Domke

"Docker containers are kind of neat. They are also kind of a craven surrender to the rotting mess of excessive software complexity." —John Carmack

Research and Work Experiences

Team Principal

Kobe, Japan

Supercomputing Performance Research Team, RIKEN Center for Computational Science (R-CCS) — *July 2022 – present*

- Note: Team Principal position is equivalent to a full Professor in Japan
- Team budget planing and decisions (labor, equipment, and travel) and employment of multiple researchers, postdoctoral researchers, trainees, and students
- Leading subgroups and assisting in the co-design of future large-scale HPC/AI deployments at RIKEN
- Conceptualizing and developing a prototype for RIKEN-wide, on-prem AI model serving infrastructure for internal generative AI services
- Model training and improving AI frameworks: porting PyTorch to aarch64 (e.g. designed CUDA emulator), replacing quantizations with floating-point compressors in llama.cpp
- Conducting and supervising performance modelling, performance studies (HPC, AI, and & cloud/DC systems), numerical and floating-point precision analyses, and tool developments
- Software tuning and deployments: application benchmarking (e.g. benchpark), compiler improvements (e.g. MLIR), HPC system software (e.g. MPI collectives), and DL/AI/ML frameworks research and development
- Organizing conferences, workshops, and HPC/AI community gatherings; and hosting of international delegations

Postdoctoral Researcher (until Mar.'21) and Research Scientist (from Apr.'21)

Kobe, Japan

High Performance Big Data Research Team, RIKEN Center for Computational Science (R-CCS) — *Apr. 2019 – June 2022*

- Researching alternative CPU, memory, and network architectures to steer co-design based on architecture simulators (e.g. gem5) and machine code analyzers, and analyze applications for their applicability to TensorCores-like architectures
- Prototype and advise the portability of AI frameworks (i.e. Tensorflow and PyTorch) to aarch64/CPU-based architectures
- Benchmarking processor- and network-related aspects of current HPC installation and testbeds at R-CCS
- Supervise internship projects on failover approaches for HPC interconnects, compiler-based tuning of GPU kernels, Neural Network-based compression, and Machine Learning for Checkpoint/Restart

Science Tokyo Research Fellow

Tokyo, Japan

Institute of Science Tokyo (formerly Tokyo Institute of Technology) — *May 2019 – present*

- Performing interconnect simulations and benchmarking newest HPC hardware

RIKEN R-CCS Visiting Researcher

Kobe, Japan

RIKEN Center for Computational Science (R-CCS) — *Feb. 2019 – Mar. 2019*

- Assisting and consulting in the early stages of the HPC system procurement
- Benchmarking network-related aspects of current HPC installation at R-CCS

Research Staff

Tokyo, Japan

Matsuoka & Endo Laboratory, Global Scientific Information and Computing Center, Tokyo Institute of Technology — *Apr. 2017 – Mar. 2019*

- Rewiring an existing 672-node, multi-petaflop/s supercomputer at TokyoTech to evaluate the novel HyperX topology
- Research in the area of HPC interconnects, topology design, and simulation frameworks
- Planning and consulting activities for future AI/HPC supercomputers, such as TSUBAME4 and AIST's ABCI system
- Supervising and mentoring undergraduate and graduate students of the Matsuoka Laboratory

Summer Student Intern

Livermore, CA, USA

Center for Applied Scientific Computing, Lawrence Livermore National Laboratory — *July 2016 – Sept. 2016*

- Applying modern network routing techniques to large-scale production HPC systems and measuring the impact
- Analyzing the disadvantages of existing routing algorithms and develop solutions to allow a more efficient and workload-specific traffic routing
- Combining two existing network simulation frameworks to improve simulation accuracy and scalability of statically routed large-scale HPC topologies

Tokyo Tech Research Fellow

Tokyo, Japan

Tokyo Institute of Technology — *Feb. 2015 – Mar. 2017*

- Performing network simulations and benchmarking newest InfiniBand hardware

Research Associate

Dresden, Germany

Faculty of Computer Science, Technische Universität Dresden

Oct. 2014 – Mar. 2017

- Research in the area of high performance interconnection networks at the Institute of Computer Engineering
- Management of the InfiniBand fabric of the 2065-node Taurus supercomputer
- Implementation of software tools for monitoring and performance analysis
- Integration into the teaching activities of the institute and student supervision
- Public relations work at local and international conferences and exhibitions, such as SC/ISC

Research Associate

Tokyo, Japan

Matsuoka Laboratory, Global Scientific Information and Computing Center, Tokyo Institute of Technology

Sept. 2013 – Sept. 2014

- Development of new deadlock-free oblivious routing algorithms

Visiting Researcher

Tokyo, Japan

Matsuoka Laboratory, Global Scientific Information and Computing Center, Tokyo Institute of Technology

Oct. 2012 – Sept. 2013

- Enhancing the deadlock-free oblivious routing algorithm and focusing on research related to routing algorithms, network topologies and network faults
- Providing knowledge and support in the field of performance analysis tools, like Vampir and Scalasca, for an application scalability study

Research Associate

Knoxville, TN, USA

Joint Institute for Computational Sciences (JICS) –University of Tennessee and Oak Ridge National Laboratory (ORNL)

Aug. 2011 – Aug. 2012

- Enhancing the functionality of the Vampir framework, from the Technische Universität Dresden, to support running in a mixed environment of X86 processors and General-Purpose computation on Graphics Processing Units processors
- Supporting the analysis process of large scale applications on Jaguar (Jaguar was the third fastest HPC system in the Top500 list, Nov. 2011)

Research Associate

Dresden, Germany

Center for Information Services and High Performance Computing (ZIH), Technische Universität Dresden

Feb. 2011 – Aug. 2011

- Extending the Vampir/VampirTrace tool set within an European–Russian project, called HOPSA

Student Research Assistant

Dresden, Germany

Center for Information Services and High Performance Computing (ZIH), Technische Universität Dresden

Nov. 2007 – Dec. 2010

- Implementation/parallelization of different sparse linear algebra algorithms for Cell BE
- BenchIT Project: Designing and testing of measurement kernels
- Performed measurements on a variety of HPC installations: SGI Altix 4700, Linux Networx Evolocity II PC-Farm, IBM BladeCenter QS21, NEC SX6 (ZIH, Dresden), NEC SX8 (HLRS, Stuttgart)
- Winner (Team) of the Cluster Challenge at Supercomputing 08 (Austin, Texas)

Vocational Internship

Dresden, Germany

Center for Information Services and High Performance Computing at Technische Universität Dresden

Aug. 2007 – Sept. 2007

- Build and tuned a cluster of Playstation 3 consoles and implemented different sparse linear algebra algorithms

Officer Candidate

Bayreuth / Fürstenfeldbruck, Germany

German Armed Forces

Oct. 2003 – June 2004

- In early 2004, resigned as officer candidate and prospective air force pilot, and moved to basic military service

Education

Doctor Candidate in Computer Science

Dresden, Germany

Faculty of Computer Science, Technische Universität Dresden

Oct. 2014 – Mar. 2017

- Topic: Routing on the Channel Dependency Graph: A New Approach to Deadlock-Free, Destination-Based, High-Performance Routing for Lossless Interconnection Networks
- Final Degree: Doctor rerum naturalium (Dr. rer. nat.) awarded in June 2017

Diploma Student in Mathematics

Dresden, Germany

Technische Universität Dresden

Oct. 2004 – Dec. 2010

- Topic: Optimized Routing for InfiniBand Networks in the field of HPC systems
- Final Degree: Diplommathematiker (equivalent to a master's degree)

High School

Weißwasser, Germany

Landau-Gymnasium Weisswasser

Aug. 1995 – July 2003

- Qualification: High School Diploma (ger. Abitur)

Primary School

Weißwasser, Germany

1. Grundschule Weisswasser

Aug. 1991 – June 1995

Skills

Team/People Management	Mentoring and Supervision, Research Guidance, Interviewing, Financial Planing, Event Planing
Deep Learning and AI	PyTorch, Model training (Fugaku-LLM), AI model serving infrastructure (ollama, vllm, etc.)
Programming and HPC	C, Python, FORTRAN, Bourne Shell, C++, Java, MPI, OpenMP, Co-Design, Network Design, Routing Protocols, Network Architecture, InfiniBand, x86, aarch64, Lustre, Cell BE
Platforms	GNU/Linux, Mac OS X, QEMU
Math Packages	PyTorch, MATLAB, Octave, Maple, SPSS
Performance Analysis	Vampir/VampirTrace, Score-P, BenchIT, GNU gprof, TAU, PAPI, gem5, benchmark suites, etc.
Publishing	LaTeX
Languages	German (Fluent), English (C1 Advanced (CEFR)), Japanese (Elementary), Russian (Elementary)

Grants, Honors, and Awards

Grants

Aug.'22 – Mar.'25	**Subgroup Leader for Performance Modelling**, "次世代計算基盤に係る調査研究 (Feasibility Studies on Next-Generation Supercomputing Infrastructures)" funded by Ministry of Education, Culture, Sports, Science and Technology (MEXT), 2.430.000.000¥	*Kobe, Japan*
Apr.'19 – Mar.'24	**Principal Investigator**, "ExaPath: Hierarchical Routing for Next-Gen Supercomputers and Beyond" funded by JSPS KAKENHI Grant-in-Aid for Scientific Research (B) Grant Number JP19H04119, 13.200.000¥	*Kobe, Japan*
Apr.'18 – Mar.'19	**co-Principal Investigator**, TokyoTech/HPE collaboration: HyperX project funded by Hewlett Packard Enterprise (HPE), 200.000$	*Tokyo, Japan*

Awards

2025	**2025 RIKEN BAIHO Award for** 「富岳」で学習した大規模言語モデル「**Fugaku-LLM**」の開発・公開, Member of large domestic team (including 9 participants from RIKEN)	*Tokyo, Japan*
2025	**3rd Place (main author: Ivan R. Ivanov)**, International Symposium on Code Generation and Optimization (CGO) 2025 - ACM Student Research Competition	*Las Vegas, NV, USA*
2024	**Rusty Lusk Award for Best Paper (main author: Semih Burak)**, 31st European MPI Users' Group Meeting, EuroMPI 2024	*Perth, WA, Australia*
2018	**Best Student Paper Finalist (main author: Stacy Smith)**, International Conference for High Performance Computing, Networking, Storage and Analysis (SC '18)	*Dallas, TX, USA*
2009	**2nd Best Paper Award (main author: Jupp Mueller)**, 10th LCI International Conference on High-Performance Clustered Computing	*Boulder, CO, USA*
2008	**1st Place (Team)**, Cluster Challenge at International Conference for High Performance Computing, Networking, Storage and Analysis (SC '08)	*Austin, Texas*

Community Services and Program Committees

Organizing Committees

2024	**Conference (Co-)chair:**, Architecture & Networks track at International Conference for High Performance Computing, Networking, Storage and Analysis, SC '24	*Atlanta, GA, USA*
2026	**Chair for Birds-of-a-Feather**, SC Asia / HPC Asia Conference 2026	*Osaka, Japan*
2025	**Chair for Workshop and Tutorial**, International Symposium on Code Generation and Optimization (CGO'25)	*Las Vegas, NV, USA*
2021	**Poster (Co-)chair**, The 3rd R-CCS International Symposium (RCCS-IS3)	*Kobe, Japan*
2024	**Panel Organizer**, International Conference for High Performance Computing, Networking, Storage and Analysis, SC '24	*Atlanta, GA, USA*
2018	**Workshop Organizer**, Minisymposium at the ACM Platform for Advanced Scientific Computing Conference (PASC'25): "Empowering Interdisciplinary Collaboration through Reproducible Benchmarking"	*Tokyo, Japan*
2025	**Workshop Organizer**, Harnessing the Power of various forms of Generative AI for Science and Engineering: The Trillion Parameter Consortium (TPC; at SCA'25)	*Singapore*
2023	**Workshop Organizer**, 28th International Workshop on High-level Parallel Programming Models and Supportive Environments (HIPS 2023; at IPDPS)	*St. Petersburg, FL, USA*
2024	**Workshop Organizer**, Benchmarking in the Data Center: Expanding to the Cloud (BID'24 at ICPE 2024)	*London, UK*
2023	**Workshop Organizer**, Benchmarking in the Data Center: Expanding to the Cloud (BID'23 at PPoPP)	*Montreal, QC, Canada*
2024	**Workshop Organizer**, 16th Joint Laboratory for Extreme Scale Computing (JLESC)	*Kobe, Japan*

2023	**Workshop Organizer**, 12th Accelerated Data Analytics and Computing Institute Workshop (ADAC12)	*Kobe, Japan*
2020	**Workshop Organizer**, 2nd International Workshop on Legacy Software Refactoring for Performance (REFAC'20 at ISC'20)	*online*
2019	**Workshop Organizer**, 1st International Workshop on Legacy Software Refactoring for Performance (REFAC'19 at ISC'19)	*Frankfurt, Germany*
2018	**Workshop Organizer**, Minisymposium at the 18th SIAM Conference on Parallel Processing for Scientific Computing (SIAM PP'18): "Applied Graph Theory in Interconnection Network Design and Operation"	*Tokyo, Japan*
2024	**Birds-of-a-Feather Organizer**, Advanced Architecture "Playgrounds" - Past Lessons, Current and Future Accesses of Testbeds (at SC'24)	*Atlanta, GA, USA*
2024	**Birds-of-a-Feather Organizer**, Advanced Architecture "Playgrounds" - Past Lessons, Current and Future Accesses of Testbeds (at SC'23)	*Denver, CO, USA*

Technical Program Committee (TPC) / Reviewer

2021–'25	**TPC**, International Conference for High Performance Computing, Networking, Storage and Analysis (SC'21, SC'23, SC'24, SC'25)
2025	**TPC**, ACM International Conference on Computing Frontiers (CF'25)
2024	**TPC**, ACM International Conference on Supercomputing (ICS)
2024–'25	**TPC**, IEEE International Conference on Cluster Computing (CLUSTER'24, CLUSTER'25)
2023–'24	**TPC**, IEEE International Conference on High Performance Computing, Data and Analytics (HiPC'23, HiPC'24)
2023	**TPC**, IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID'23)
2022–'25	**TPC**, International Symposium on High-Performance Parallel and Distributed Computing (HPDC'22, HPDC'23, HPDC'25)
2022–'25	**TPC**, IEEE International Parallel & Distributed Processing Symposium (IPDPS'22, IPDPS'23, IPDPS'24, IPDPS'25
2020	**TPC**, International Conference on High Performance Computing in Asia-Pacific Region (HPC Asia'20)
2017–'24	**TPC**, International Conference on Parallel Processing (ICPP'17, ICPP'22, ICPP'24)
2024	**TPC**, Parallel AI and Systems for the Edge (PAISE'24 at IPDPS)
2023–'24	**TPC**, International Workshop on HPC Testing and Evaluation of Systems, Tools, and Software (HPCTESTS'23 at SC, HPCTESTS'24 at SC)
2023–'25	**TPC**, International Workshop on Coarse-Grained Reconfigurable Architectures for High-Performance Computing (CGRA4HPC'23 at IPDPS, CGRA4HPC'24 at IPDPS, CGRA4HPC'25 at IPDPS
2023–'25	**TPC**, International Workshop on Arm-based HPC: Practice and Experience (IWAHPCE'23 at HPCAsia, IWAHPCE'24 at HPCAsia, IWAHPCE'25 at HPCAsia
2024	**TPC**, International Workshop on Large Language Models (LLMs) and HPC (LLMxHPC'24 at CLUSTER)
2020–'22	**TPC**, EAHPC - Embracing Arm (EAHPC'20 at CLUSTER, EAHPC'21 at CLUSTER, EAHPC'22 at CLUSTER)
2020–'21	**TPC**, R-CCS International Symposium
2018–'24	**TPC**, Annual OpenFabrics Alliance Workshop (OFAWS'18–'24)
2018–'19	**TPC**, IEEE International Workshop of High-Perfomance Interconnection Networks in the Exascale and Big-Data Era (HiPINEB'18, HiPINEB'19)
2025	**TPC**, International Workshop on Foundational large Language Models Advances for HPC (LLM4HPC'25 at ISC)

Funding Agency and Conference/Journal (Sub-)Reviewer

Grant Reviewer, Office of Advanced Scientific Computing Research (ASCR), Department of Energy (DoE)
Reviewer, Computing (COMP)
Reviewer, IEEE Micro
Reviewer, IEEE Transactions on Computers (TC)
Reviewer, IEEE Transactions on Parallel and Distributed Systems (TPDS)
Reviewer, IEEE Transactions on Network and Service Management
Reviewer, Data in Brief (DIB)
Reviewer, Journal of King Saud University - Computer and Information Sciences (JKSU-CIS)
Reviewer, Journal of Parallel and Distributed Computing (JPDC)
Reviewer, Parallel Computing (PARCO)
Reviewer, Institute of Electronics, Information and Communication Engineers (IEICE)
Reviewer, ACM Symposium on High-Performance Parallel and Distributed Computing (HPDC)
Reviewer, IEEE International Conference on Parallel and Distributed Systems (ICPADS)
Reviewer, IEEE International Parallel & Distributed Processing Symposium (IPDPS)
Reviewer, IFIP International Conference on Network and Parallel Computing (NPC)

Reviewer, International Conference for High Performance Computing, Networking, Storage and Analysis
Reviewer, International Conference on Supercomputing (ICS)
Reviewer, International European Conference on Parallel and Distributed Computing (Euro-Par)
Reviewer, International Journal of High Performance Computing Applications (IJHPCA)
Reviewer, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)

Publications

[1] Faveo Hoerold et al. “RAPTOR: Numerical Profiling of Scientific Applications”. In: *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. SC ’25. IEEE Press, Nov. 2025.

[2] Daniele De Sensi et al. “Bine Trees: Enhancing Collective Operations by Optimizing Communication Locality”. In: *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. SC ’25. IEEE Press, Nov. 2025.

[3] Semih Burak et al. “Extending the SPMD IR for RMA Models and Static Data Race Detection”. In: EuroMPI/USA. Oct. 1, 2025.

[4] Gianmarco Accordi et al. “Leveraging Modern HPC CPUs Architectures for Portable Molecular Docking: Insights on Auto-Vectorization and Optimization”. In: 2025 IEEE International Conference on Cluster Computing (CLUSTER). Sept. 2025.

[5] Francesco Antici et al. “An Online Algorithm for Power Consumption Prediction of HPC Workload”. In: *Future Generation Computer Systems* (Aug. 5, 2025), p. 108064. ISSN: 0167-739X. DOI: 10.1016/j.future.2025.108064. URL: https://www.sciencedirect.com/science/article/pii/S0167739X25003590.

[6] Francesco Antici et al. “F-DATA: A Fugaku Workload Dataset for Job-centric Predictive Modelling in HPC Systems”. In: *Scientific Data* 12.1 (July 30, 2025), p. 1321. ISSN: 2052-4463. DOI: 10.1038/s41597-025-05633-1.

[7] Francesco Antici et al. “UoPC: A User-based Online Framework to Predict Job Power Consumption in HPC Systems”. In: ISC High Performance 2025. Springer, June 2025.

[8] Maximilian Sander and Jens Domke. “Compressing Large Language Models with ZFP: Lessons Learned”. In: *Proceedings of the LG-ARC’ 2025 Workshop in Conjunction with the 52st International Symposium on Computer Architecture (ISCA-52)*. LG-ARC’2025 Workshop @ISCA 2025. June 2025.

[9] Ke Fan et al. “Parameterized Hierarchical Algorithms for Non-uniform All-to-all”. In: *Proceedings of the 34th ACM International Symposium on High-Performance Parallel and Distributed Computing*. HPDC ’25. New York, NY, USA: ACM, June 20, 2025. DOI: 10.1145/3731545.3731590.

[10] Clément Gavoille et al. “Performance Projection for Design-Space Exploration on Future HPC Architectures”. In: *2025 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. Milan, Italy: IEEE Computer Society Press, June 3, 2025, pp. 261–272. DOI: 10.1109/IPDPS64566.2025.00031.

[11] Jens Domke et al. “A Unifying Framework to Enable Artificial Intelligence in High-Performance Computing Workflows”. In: *Computing in Science & Engineering* 27.01 (May 28, 2025), pp. 73–78. ISSN: 1558-366X. DOI: 10.1109/MCSE.2025.3543940.

[12] William Dawson et al. “Reducing Numerical Precision Requirements in Quantum Chemistry Calculations.” In: *Journal of chemical theory and computation* 20.24 (Dec. 7, 2024), pp. 10826–10837. URL: https://pubs.acs.org/doi/full/10.1021/acs.jctc.4c00938.

[13] Semih Burak et al. “SPMD IR: Unifying SPMD and Multi-Value IR Showcased for Static Verification of Collectives”. In: *Recent Advances in the Message Passing Interface: 31st European MPI Users’ Group Meeting, EuroMPI 2024, Perth, WA, Australia, September 25–27, 2024, Proceedings*. Berlin, Heidelberg: Springer-Verlag, Sept. 2024, pp. 3–20. ISBN: 978-3-031-73369-7. DOI: 10.1007/978-3-031-73370-3_1.

[14] Ivan R. Ivanov et al. “Automatic Parallelization and OpenMP Offloading of Fortran Array Notation”. In: *Advancing OpenMP for Future Accelerators: 20th International Workshop on OpenMP, IWOMP 2024, Perth, WA, Australia, September 23–25, 2024, Proceedings*. Berlin, Heidelberg: Springer-Verlag, Sept. 2024, pp. 197–209. ISBN: 978-3-031-72566-1. DOI: 10.1007/978-3-031-72567-8_13.

[15] Nils Blach et al. “A High-Performance Design, Implementation, Deployment, and Evaluation of the Slim Fly Network”. In: *Proceedings of the 21st USENIX Symposium on Networked Systems Design and Implementation*. NSDI’24. Santa Clara, USA: USENIX Association, Apr. 16, 2024, p. 20. ISBN: 978-1-939133-39-7. DOI: 10.5555/3691825.3691882.

[16] Ivan R. Ivanov et al. “Retargeting and Respecializing GPU Workloads for Performance Portability”. In: *2024 IEEE/ACM International Symposium on Code Generation and Optimization (CGO)*. Edinburgh, UK, Mar. 2024, pp. 119–132. DOI: 10.1109/CGO57630.2024.10444828.

[17] Jens Domke et al. “At the Locus of Performance: Quantifying the Effects of Copious 3D-Stacked Cache on HPC Workloads”. In: *ACM Trans. Archit. Code Optim.* TACO 20.4 (Dec. 14, 2023). ISSN: 1544-3566. DOI: 10.1145/3629520.

[18] Francesco Antici et al. “Augmenting ML-based Predictive Modelling with NLP to Forecast a Job’s Power Consumption”. In: *Proceedings of the SC ’23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis*. 1st International Workshop on the Environmental Sustainability of High-Performance Software (SHiPS@SC’23), Denver, CO, USA, November 12, 2023. SC-W ’23. New York, NY, USA: IEEE, Nov. 2023, pp. 1820–1830. ISBN: 9798400707858. DOI: 10.1145/3624062.3624264.

[19] Olga Pearce et al. “Towards Collaborative Continuous Benchmarking for HPC”. In: *Proceedings of the SC ’23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis*. First International Workshop on HPC Testing and Evaluation of Systems, Tools, and Software (HPCTESTS@SC’23), Denver, CO, USA, November 17, 2023. SC-W ’23. New York, NY, USA: IEEE, Nov. 2023, pp. 627–635. ISBN: 9798400707858. DOI: 10.1145/3624062.3624135.

[20] Satoshi Matsuoka et al. “Myths and Legends in High-Performance Computing”. In: *The International Journal of High Performance Computing Applications* 37.3–4 (Apr. 23, 2023), pp. 245–259. DOI: 10.1177/10943420231166608.

[21] William S. Moses et al. “High-Performance GPU-to-CPU Transpilation and Optimization via High-Level Parallel Constructs”. In: *Proceedings of the 28th ACM SIGPLAN Annual Symposium on Principles and Practice of Parallel Programming*. PPoPP ’23. New York, NY, USA: Association for Computing Machinery, Feb. 2023, pp. 119–134. ISBN: 9798400700156. DOI: 10.1145/3572848.3577475.

[22] Truong Thao Nguyen et al. “Why Globally Re-shuffle? Revisiting Data Shuffling in Large Scale Deep Learning”. In: *2022 IEEE International Parallel and Distributed Processing Symposium, IPDPS 2022, Lyon, France, May, 2022*. 36rd IEEE International Parallel and Distributed Processing Symposium. Lyon, France, May 2022.

[23] Satoshi Matsuoka et al. “Preparing for the Future—Rethinking Proxy Applications”. In: *Computing in Science & Engineering* 24.2 (Mar. 2022), pp. 85–90. DOI: 10.1109/MCSE.2022.3153105. URL: https://ieeexplore.ieee.org/document/9789513.

[24] Steven Farrell et al. “MLPerf HPC: A Holistic Benchmark Suite for Scientific Machine Learning on HPC Systems”. In: *7th IEEE/ACM Workshop on Machine Learning in High Performance Computing Environments (MLHPC@SC’21), St. Louis, MO, USA, November 15, 2021*. 7th IEEE/ACM Workshop on Machine Learning in High Performance Computing Environments (MLHPC@SC’21), St. Louis, MO, USA, November 15, 2021. IEEE, Nov. 2021.

[25] Jens Domke. “A64FX – Your Compiler You Must Decide!” In: *2021 IEEE International Conference on Cluster Computing (CLUSTER), EAHPC Workshop*. Portland, Oregon, USA: IEEE Computer Society, Sept. 2021.

[26] Jens Domke et al. “Matrix Engines for High Performance Computing: A Paragon of Performance or Grasping at Straws?” In: *2021 IEEE International Parallel and Distributed Processing Symposium, IPDPS 2021, Portland, Oregon, USA, May 17-21, 2021*. 35rd IEEE International Parallel and Distributed Processing Symposium. Portland, Oregon, USA: IEEE Press, May 2021, pp. 1056–1065. DOI: 10.1109/IPDPS49936.2021.00114.

[27] Maciej Besta et al. “High-Performance Routing With Multipathing and Path Diversity in Ethernet and HPC Networks”. In: *IEEE Transactions on Parallel and Distributed Systems* 32.4 (Apr. 1, 2021), pp. 943–959. DOI: 10.1109/TPDS.2020.3035761.

[28] Mohamed Wahib et al. “Scaling Distributed Deep Learning Workloads beyond the Memory Capacity with KARMA”. In: *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. SC ’20. IEEE Press, Nov. 2020. ISBN: 978-1-72819-998-6.

[29] Tonmoy Dey et al. “Optimizing Asynchronous Multi-Level Checkpoint/Restart Configurations with Machine Learning”. In: *2020 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*. HPS ’20. New Orleans, LA, USA, May 2020, pp. 1036–1043. DOI: 10.1109/IPDPSW50202.2020.00174.

[30] Jens Domke et al. “HyperX Topology: First At-Scale Implementation and Comparison to the Fat-Tree”. In: *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. The International Conference for High Performance Computing, Networking, Storage, and Analysis. SC ’19. New York, NY, USA: ACM, Nov. 2019, 40:1–40:23. ISBN: 978-1-4503-6229-0. DOI: 10.1145/3295500.3356140.

[31] Jens Domke et al. “The First Supercomputer with HyperX Topology: A Viable Alternative to Fat-Trees?” In: 2019 IEEE 26th Symposium on High-Performance Interconnects. HOTI 26. Piscataway, NJ, USA: IEEE Press, Aug. 2019, p. 4.

[32] Jens Domke et al. “Double-Precision FPUs in High-Performance Computing: An Embarrassment of Riches?” In: *2019 IEEE International Parallel and Distributed Processing Symposium, IPDPS 2019, Rio de Janeiro, Brazil, May 20-24, 2019*. 33rd IEEE International Parallel and Distributed Processing Symposium. Rio de Janeiro, Brazil: IEEE Press, May 2019, pp. 78–88. DOI: 10.1109/IPDPS.2019.00019.

[33] Harsh Bhatia et al. “Interactive Investigation of Traffic Congestion on Fat-Tree Networks Using TREESCOPE”. In: *Computer Graphics Forum* 37.3 (2018). Ed. by Jeffrey Heer, Heike Leitte, and Timo Ropinski, pp. 561–572. DOI: 10.1111/cgf.13442.

[34] Staci A. Smith et al. “Mitigating Inter-Job Interference Using Adaptive Flow-Aware Routing”. In: *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. SC ’18. Piscataway, NJ, USA: IEEE Press, Nov. 2018, 27:1–27:15. DOI: 10.1109/SC.2018.00030.

[35] Misbah Mubarak et al. “Toward Reliable Validation of HPC Interconnect Simulations”. In: *Proceedings of the 2017 Winter Simulation Conference*. WSC ’17. Las Vegas, NV, USA: IEEE Press, Dec. 2017, pp. 659–674.

[36] Jens Domke. “Routing on the Channel Dependency Graph: A New Approach to Deadlock-Free, Destination-Based, High-Performance Routing for Lossless Interconnection Networks”. PhD thesis. Fakultät Informatik, Professur für Rechnerarchitektur: Technische Universität Dresden, June 16, 2017.

[37] Noah Wolfe et al. “Preliminary Performance Analysis of Multi-rail Fat-tree Networks”. In: *17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*. IEEE/ACM CCGrid 2017. CCGrid ’17. Madrid, Spain: IEEE Press, May 2017, pp. 258–261. ISBN: 978-1-5090-6610-0. DOI: 10.1109/CCGRID.2017.102.

[38] Jens Domke, Torsten Hoefler, and Satoshi Matsuoka. “Routing on the Dependency Graph: A New Approach to Deadlock-Free High-Performance Routing”. In: *Proceedings of the 25th ACM International Symposium on High-Performance Parallel and Distributed Computing*. HPDC ’16. New York, NY, USA: ACM, 2016, pp. 3–14. ISBN: 978-1-4503-4314-5. DOI: 10.1145/2907294.2907313.

[39] Jens Domke and Torsten Hoefler. “Scheduling-Aware Routing for Supercomputers”. In: *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. SC ’16. Piscataway, NJ, USA: IEEE Press, Nov. 2016, 13:1–13:12. ISBN: 978-1-4673-8815-3. URL: http://dl.acm.org/citation.cfm?id=3014904.3014922.

[40] Dali Wang et al. “A Scalable Framework for the Global Offline Community Land Model Ensemble Simulation”. In: *Int. J. Comput. Sci. Eng.* 12.1 (Feb. 2016), pp. 73–85. ISSN: 1742-7185. DOI: 10.1504/IJCSE.2016.074565.

[41] Kevin Brown, Jens Domke, and Satoshi Matsuoka. “Hardware-Centric Analysis of Network Performance for MPI Applications”. In: *2015 21th IEEE International Conference on Parallel and Distributed Systems (ICPADS)*. International Conference on Parallel and Distributed Systems. Melbourne, Australia: IEEE Press, Dec. 2015, p. 8.

[42] Jens Domke, Torsten Hoefler, and Satoshi Matsuoka. “Fail-in-Place Network Design: Interaction between Topology, Routing Algorithm and Failures”. In: *Proceedings of the IEEE/ACM International Conference for High Performance Computing, Networking, Storage and Analysis (SC14)*. The International Conference for High Performance Computing Networking, Storage, and Analysis. SC ’14. New Orleans, LA, USA: IEEE Press, Nov. 2014, pp. 597–608. ISBN: 978-1-4799-5500-8. DOI: 10.1109/SC.2014.54.

[43] Jens Domke and Dali Wang. “Runtime Tracing of the Community Earth System Model: Feasibility Study and Benefits”. In: *Procedia Computer Science* 9.0 (2012), pp. 1950–1958. ISSN: 1877-0509. DOI: 10.1016/j.procs.2012.04.213. URL: http://www.sciencedirect.com/science/article/pii/S1877050912003341.

[44] Jens Domke, Torsten Hoefler, and Wolfgang E. Nagel. “Deadlock-Free Oblivious Routing for Arbitrary Topologies”. In: *Proceedings of the 25th IEEE International Parallel & Distributed Processing Symposium (IPDPS)*. Washington, DC, USA: IEEE Computer Society, May 2011, pp. 613–624. ISBN: 0-7695-4385-7.

[45] J. Mueller et al. “Cluster Challenge 2008: Optimizing Cluster Configuration and Applications to Maximize Power Efficiency”. In: *Proceedings of the 10th LCI International Conference on High-Performance Clustered Computing*. Mar. 2009.