

performance per dollar. Second, we can sacrifice perfect balance and achieve better performance per dollar for some A2A dragonfly and express mesh systems.

ACKNOWLEDGMENTS

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 (LLNL-CONF-772399). The CODES/ROSS simulation suite has been supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, under contract number DE-AC02-06CH11357.

This document was prepared as an account of work sponsored by an agency of the U.S. government. Neither the U.S. government nor Lawrence Livermore National Security, LLC (LLNS), nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. government or LLNS. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. government or LLNS, and shall not be used for advertising or product endorsement purposes.

REFERENCES

- [1] [n. d.]. Open Trace Format 2. <https://silc.zih.tu-dresden.de/otf2-current/index.html>.
- [2] [n. d.]. Slingshot: The Interconnect for the Exascale Era. <https://www.cray.com/sites/default/files/Slingshot-The-Interconnect-for-the-Exascale-Era.pdf>.
- [3] 2015. Score-P User Manual. <https://silc.zih.tu-dresden.de/scorep-current/pdf/scorep.pdf>
- [4] Bilge Acun, Nikhil Jain, Abhinav Bhatele, Misbah Mubarak, Christopher D. Carothers, and Laxmikant V. Kale. 2015. Preliminary Evaluation of a Parallel Trace Replay Tool for HPC Network Simulations. In *Proceedings of the 3rd Workshop on Parallel and Distributed Agent-Based Simulations (PADABS '15)*. LLNL-CONF-667225.
- [5] Jung Ho Ahn, Nathan Binkert, Al Davis, Moray McLaren, and Robert S. Schreiber. 2009. HyperX: Topology, Routing, and Packaging of Efficient Large-scale Networks. In *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis (SC '09)*. ACM, New York, NY, USA.
- [6] B. Arimilli, R. Arimilli, V. Chung, S. Clark, W. Denzel, B. Drerup, T. Hoefler, J. Joyner, J. Lewis, Jian Li, Nan Ni, and R. Rajamony. 2010. The PERCS High-Performance Interconnect. In *18th Annual Symposium on High Performance Interconnects (HOTI)*. 75–82.
- [7] David W. Bauer Jr., Christopher D. Carothers, and Akintayo Holder. 2009. Scalable Time Warp on Blue Gene Supercomputers. In *Proceedings of the 2009 ACM/IEEE/SCS 23rd Workshop on Principles of Advanced and Distributed Simulation (PADS '09)*. IEEE Computer Society, Washington, DC, USA.
- [8] Claude Bernard, Tom Burch, Thomas A. DeGrand, Carleton DeTar, Steven Gottlieb, Urs M. Heller, James E. Hetrick, Kostas Orginos, Bob Sugar, and Doug Toussaint. 2000. Scaling tests of the improved Kogut-Susskind quark action. *Physical Review D* 61 (2000).
- [9] Maciej Besta and Torsten Hoefler. 2014. Slim Fly: A Cost Effective Low-diameter Network Topology. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC '14)*. IEEE Press, 348–359.
- [10] Ron Brightwell and Keith Underwood. 2003. Evaluation of an eager protocol optimization for MPI. In *European Parallel Virtual Machine/Message Passing Interface Users' Group Meeting*. Springer, 327–334.
- [11] D. Chen, P. Heidelberger, C. Stunkel, Y. Sugawara, C. Minkenber, B. Prisacari, and G. Rodriguez. 2016. An Evaluation of Network Architectures for Next Generation Supercomputers. In *7th International Workshop on Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems (PMBS)*. <https://doi.org/10.1109/PMBS.2016.007>
- [12] William Dally and Brian Towles. 2003. *Principles and Practices of Interconnection Networks*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- [13] William J. Dally. 1991. Express Cubes: Improving the Performance of K-ary N-cube Interconnection Networks. *IEEE Trans. Comput.* 40, 9 (Sept. 1991), 1016–1023. <https://doi.org/10.1109/12.83652>
- [14] D. Eschweiler, M. Wagner, M. Geimer, A. Knüpfer, W. E. Nagel, and F. Wolf. 2012. Open Trace Format 2: The Next Generation of Scalable Trace Formats and Support Libraries. 22 (2012).
- [15] Greg Faanes, Abdulla Bataineh, Duncan Roweth, Tom Court, Edwin Froese, Bob Alverson, Tim Johnson, Joe Kopnick, Mike Higgins, and James Reinhard. 2012. Cray Cascade: A Scalable HPC System Based on a Dragonfly Network. In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis (SC '12)*. IEEE Computer Society Press, Los Alamitos, CA, USA.
- [16] I. Fujiwara, M. Koibuchi, H. Matsutani, and H. Casanova. 2014. Skywalk: A Topology for HPC Networks with Low-Delay Switches. In *2014 IEEE 28th International Parallel and Distributed Processing Symposium*. 263–272. <https://doi.org/10.1109/IPDPS.2014.37>
- [17] F. Gygi, E. W. Draeger, B. R. De Supinski, R. K. Yates, F. Franchetti, S. Kral, J. Lorenz, C. W. Ueberhuber, J. A. Gunnel, and J. C. Sexton. 2005. Large-Scale First-Principles Molecular Dynamics Simulations on the Blue Gene/L Platform using the Qbox Code. In *Proceedings of Supercomputing 2005* 4 (2005), 24. Conference on High Performance Computing and Networking, Gordon Bell Prize finalist.
- [18] Chao Huang, Gengbin Zheng, Sameer Kumar, and Laxmikant V. Kale. 2006. Performance Evaluation of Adaptive MPI. In *Proceedings of ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming 2006*.
- [19] Nikhil Jain, Abhinav Bhatele, Xiang Ni, Todd Gamblin, and Laxmikant V. Kale. 2017. Partitioning Low-diameter Networks to Eliminate Inter-job Interference. In *Proceedings of the IEEE International Parallel & Distributed Processing Symposium (IPDPS '17)*. IEEE Computer Society. LLNL-CONF-706801.
- [20] Nikhil Jain, Abhinav Bhatele, Samuel T. White, Todd Gamblin, and Laxmikant V. Kale. 2016. Evaluating HPC Networks via Simulation of Parallel Workloads. In *Proceedings of the ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis (SC '16)*. IEEE Computer Society. LLNL-CONF-690662.
- [21] N. Jiang, J. Balfour, D. U. Becker, B. Towles, W. J. Dally, G. Michelogiannakis, and J. Kim. 2013. A detailed and flexible cycle-accurate Network-on-Chip simulator. In *2013 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*. 86–96.
- [22] Sangeetha Abdu Jyothi, Ankit Singla, P. Brighten Godfrey, and Alexandra Kolla. 2016. Measuring and Understanding Throughput of Network Topologies. In *Supercomputing 2016 (SC '16)*. Salt Lake City, UT.
- [23] J. Kim, W. Dally, S. Scott, and D. Abts. 2009. Cost-Efficient Dragonfly Topology for Large-Scale Systems. *IEEE Micro* 29, 1 (Jan 2009), 33–40.
- [24] John Kim, William J. Dally, Steve Scott, and Dennis Abts. 2008. Technology-Driven, Highly-Scalable Dragonfly Topology. *SIGARCH Comput. Archit. News* 36 (June 2008), 77–88. Issue 3.
- [25] C.E. Leiserson. 1985. Fat-trees: Universal Networks for Hardware-Efficient Supercomputing. *IEEE Transactions on Computers* 34, 10 (October 1985).
- [26] M.Blumrich, D.Chen, P.Coteus, A.Gara, M.Giampapa, P.Heidelberger, S.Singh, B.Steinmacher-Burrow, T.Takken, and P.Vranas. 2003. Design and Analysis of the Blue Gene/L Torus Interconnection Network. *IBM Research Report* (December 2003).
- [27] D. M. Nicol, C. C. Michael, and P. Inouye. 1989. Efficient Aggregation of Multiple PLs in Distributed Memory Parallel Simulations. In *Proceedings of the 21st Conference on Winter Simulation (WSC '89)*. ACM, New York, NY, USA, 680–685. <https://doi.org/10.1145/76738.76825>
- [28] Yan Pan, Prabhat Kumar, John Kim, Gokhan Memik, Yu Zhang, and Alok Choudhary. 2009. Firefly: Illuminating Future Network-on-chip with Nanophotonics. In *Proceedings of the 36th Annual International Symposium on Computer Architecture (ISCA '09)*. ACM, New York, NY, USA, 429–440. <http://doi.acm.org/10.1145/1555754.1555808>
- [29] Bogdan Prisacari, German Rodriguez, Philip Heidelberger, Dong Chen, Cyriel Minkenber, and Torsten Hoefler. 2014. Efficient Task Placement and Routing of Nearest Neighbor Exchanges in Dragonfly Networks. In *Proceedings of the 23rd International Symposium on High-performance Parallel and Distributed Computing (HPDC '14)*. ACM, 129–140. <https://doi.org/10.1145/2600212.2600225>
- [30] C. H. Still, R. L. Berger, A. B. Langdon, D. E. Hinkel, L. J. Suter, and E. A. Williams. 2000. Filamentation and forward Brillouin scatter of entire smoothed and aberrated laser beams. *Physics of Plasmas* 7, 5 (2000), 2023–2032.
- [31] Rajeev Thakur, Rolf Rabenseifner, and William Gropp. 2005. Optimization of Collective Communication Operations in MPICH. *International Journal of High Performance Computing Applications* 19, 1 (2005), 49–66.
- [32] J. Won, G. Kim, J. Kim, T. Jiang, M. Parker, and S. Scott. 2015. Overcoming far-end congestion in large-scale networks. In *2015 IEEE 21st International Symposium on High Performance Computer Architecture (HPCA)*. 415–427.
- [33] Gengbin Zheng, Gunavardhan Kakulapati, and Laxmikant V. Kale. 2004. BigSim: A Parallel Simulator for Performance Prediction of Extremely Large Parallel Machines. In *18th International Parallel and Distributed Processing Symposium (IPDPS)*. Santa Fe, New Mexico, 78.