

we thank the anonymous reviewers for their valuable input that helped to improve the quality of the paper.

8. REFERENCES

- [1] C. Bennett, R. Grossman, and J. Seidman. MalStone: A Benchmark for Data Intensive Computing. Technical report, Open Cloud Consortium, 2009.
- [2] D. Bitton, D. J. DeWitt, and C. Turbyfill. Benchmarking Database Systems: A Systematic Approach. In *VLDB '83: Proceedings of the 9th International Conference on Very Large Data Bases*, pages 8–19, San Francisco, CA, USA, November 1983. ACM, Morgan Kaufmann Publishers Inc.
- [3] N. Bruno and S. Chaudhuri. Flexible Database Generators. In *VLDB '05: Proceedings of the 31st International Conference on Very Large Databases*, pages 1097–1107. VLDB Endowment, 2005.
- [4] Continuent. Bristlecone. <https://bristlecone.svn.sourceforge.net/svnroot/bristlecone/trunk/bristlecone/>.
- [5] B. F. Cooper, A. Silberstein, E. Tam, R. Ramakrishnan, and R. Sears. Benchmarking Cloud Serving Systems with YCSB. In *SoCC '10: Proceedings of the 1st ACM Symposium on Cloud Computing*, pages 143–154, New York, NY, USA, 2010. ACM.
- [6] DTM Database Tools. Dtm data generator. <http://www.sqledit.com/dg/>.
- [7] J. Gray, P. Sundaresan, S. Englert, K. Baclawski, and P. J. Weinberger. Quickly Generating Billion-Record Synthetic Databases. In *SIGMOD '94: Proceedings of the 1994 ACM SIGMOD International Conference on Management of Data*, pages 243–252, New York, NY, USA, 1994. ACM.
- [8] GSApps. Gs data generator. <http://www.gsapps.com/products/datagenerator/>.
- [9] J. E. Hoag and C. W. Thompson. A Parallel General-Purpose Synthetic Data Generator. *SIGMOD Record*, 36(1):19–24, 2007.
- [10] K. Houkjær, K. Torp, and R. Wind. Simple and Realistic Data Generation. In *VLDB '06: Proceedings of the 32nd international conference on Very large data bases*, pages 1243–1246. VLDB Endowment, 2006.
- [11] K. Huppler. The Art of Building a Good Benchmark. In *TPCTC '09: First TPC Technology Conference on Performance Evaluation and Benchmarking*, pages 18–30, 2009.
- [12] P. J. Lin, B. Samadi, A. Cipolone, D. R. Jeske, S. Cox, C. Rendón, D. Holt, and R. Xiao. Development of a Synthetic Data Set Generator for Building and Testing Information Discovery Systems. In *ITNG '06: Proceedings of the Third International Conference on Information Technology: New Generations*, pages 707–712, Washington, DC, USA, 2006. IEEE Computer Society.
- [13] G. Marsaglia. Xorshift RNGs. *Journal Of Statistical Software*, 8(14):1–6, 2003.
- [14] R. O. Nambiar and M. Poess. The Making of TPC-DS. In *VLDB '06: Proceedings of the 32nd International Conference on Very Large Data Bases*, pages 1049–1058, 2006.
- [15] P. E. O’Neil. The Set Query Benchmark. In J. Gray, editor, *The Benchmark Handbook for Database and Transaction Systems (2nd Edition)*. Morgan Kaufmann Publishers, 1993.
- [16] F. Panneton and P. L’ecuyer. On the Xorshift Random Number Generators. *ACM Transactions on Modeling and Computer Simulation*, 15(4):346–361, 2005.
- [17] M. Poess and C. Floyd. New TPC Benchmarks for Decision Support and Web Commerce. *SIGMOD Record*, 29(4):64–71, 2000.
- [18] M. Poess, T. Rabl, M. Frank, and M. Danisch. A PDGF Implementation for TPC-H. In *TPCTC '11: Third TPC Technology Conference on Performance Evaluation and Benchmarking*, 2011.
- [19] T. Rabl, M. Frank, H. M. Sergieh, and H. Kosch. A Data Generator for Cloud-Scale Benchmarking. In *TPCTC '10: Proceedings of the Second TPC Technology Conference on Performance Evaluation, Measurement and Characterization of Complex Systems*, pages 41–56, 2010.
- [20] T. Rabl, A. Lang, T. Hackl, B. Sick, and H. Kosch. Generating Shifting Workloads to Benchmark Adaptability in Relational Database Systems. In R. O. Nambiar and M. Poess, editors, *TPCTC '09: First TPC Technology Conference on Performance Evaluation and Benchmarking*, volume 5895 of *Lecture Notes in Computer Science*, pages 116–131. Springer, 2009.
- [21] T. Rabl and M. Poess. Parallel Data Generation for Performance Analysis of Large, Complex RDBMS. In *DBTest '11: Proceedings of the 4th International Workshop on Testing Database Systems*, page 5, 2011.
- [22] Red Gate. Sql data generator 2.0. <http://www.red-gate.com/products/sql-development/sql-data-generator/>.
- [23] J. M. Stephens and M. Poess. MUDD: a multi-dimensional data generator. In *WOSP '04: Proceedings of the 4th International Workshop on Software and Performance*, pages 104–109, New York, NY, USA, 2004. ACM.
- [24] The Transaction Performance Processing Council. Dbgen. <http://www.tpc.org/tpch/>.
- [25] L. Wyatt, B. Caufield, and D. Pol. Principles for an ETL Benchmark. In *TPC TC '09: First TPC Technology Conference on Performance Evaluation and Benchmarking*, pages 183–198, 2009.