



Figure 1: Big Data Benchmarks Classification [25]

to cloud services. It aims to collect publicly available data sets and models for the following problems: *Image classification, Object detection, Translation, Recommendation, Reinforcement Learning, Speech to text and Sentiment Analysis.*

4.4 DataBench

DataBench² [27, 37] is a three year EU-funded project that investigates existing Big Data benchmarking tools and projects, identifies the main gaps and provides a robust set of metrics to compare technical results coming from those tools. The DataBench Toolbox is a one-stop-shop for Big Data Benchmarking, offering multiple benefits for different kind of users and businesses.

4.5 ABench

ABench [28] is *Big Data Architecture Stack Benchmark* that targets the representation and comparison of different Big Data architecture patterns. The benchmark framework shall stress test the common application business requirements (e.g. retail analytics, retail operational, etc.), big data technologies functionalities and best practice implementation architectures. The benchmark framework

²www.databench.eu

should have an open source implementation and extendable design as well as easy to be setup and extend. It should include data generator, public data sets and existing benchmarks to simulate workloads that stress test the best practice Big Data architectures.

5 CONCLUSIONS

AI is pervasive in our life and there is need to understand how to benchmark systems, which are used to build machine-learning and deep-learning based solutions using emerging big data technology stacks. This tutorial covers the research questions and available benchmarks in this domain. We have discussed in detail popular benchmarks such as BigBench, BigDataBench, MLPerf and the DataBench project.

ACKNOWLEDGMENTS

This work has been partially funded by the European Commission H2020 project DataBench - Evidence Based Big Data Benchmarking to Improve Business Performance, under project No. 780966. This work expresses the opinions of the authors and not necessarily those of the European Commission. The European Commission is not liable for any use that may be made of the information contained

in this work. The authors thank all the participants in the project for discussions and common work.

REFERENCES

- [1] Faraz Ahmad. 2012. PUMA Benchmarks. <https://engineering.purdue.edu/~puma/pumabenchmarks.htm>
- [2] AsterixDB. 2020. asterixdb.apache.org.
- [3] Tadas Baltrusaitis, Chaitanya Ahuja, and Louis-Philippe Morency. 2019. Multimodal Machine Learning: A Survey and Taxonomy. *IEEE Trans. Pattern Anal. Mach. Intell.* 41, 2 (2019), 423–443.
- [4] Jilil Boukhobza, Stéphane Rubini, Renhai Chen, and Zili Shao. 2018. Emerging NVM: A Survey on Architectural Integration and Research Challenges. *ACM Trans. Design Autom. Electr. Syst.* 23, 2 (2018), 14:1–14:32.
- [5] Yanpei Chen et al. 2012. We don't know enough to make a big data benchmark suite—an academia-industry view. *Proc. of WBDB* (2012).
- [6] Chaitanya K. Baru et al. 2014. BigDataBench: A Proposed Industry Standard Performance Benchmark for Big Data. In *Proc. of the 6th TPCTC 2014, Hangzhou, China, Sept. 1-5, 2014*.
- [7] Lei Wang et al. 2014. BigDataBench: A big data benchmark suite from internet services. In *Proc. of the 20th IEEE HPCA 2014, Orlando, FL, USA, February 15-19, 2014*. IEEE.
- [8] Michael Ferdman et al. 2012. Clearing the clouds: a study of emerging scale-out workloads on modern hardware. In *Proc. of the 17th ASPLOS 2012, London, UK, March 3-7, 2012*.
- [9] Norman Jouppi et al. 2017. In-Datacenter Performance Analysis of a Tensor Processing Unit. In *Proc. of the 44th ISCA 2017, Toronto, ON, Canada, June 24-28, 2017*.
- [10] Peter Mattson et al. 2019. MLPerf Training Benchmark. *CoRR* abs/1910.01500 (2019). arXiv:1910.01500
- [11] Vijay Janapa Reddi et al. 2019. MLPerf Inference Benchmark. *CoRR* abs/1911.02549 (2019). arXiv:1911.02549
- [12] Zijian Ming et al. 2013. BDGS: A Scalable Big Data Generator Suite in Big Data Benchmarking. In *Proc. of the 2013 Workshop on Big Data Benchmarking, Xi'an, China, July 16-17, 2013 and San José, CA, USA, October 9-10, 2013*.
- [13] Flink. 2020. flink.apache.org/.
- [14] Ahmad Ghazal, Todor Ivanov, Pekka Kostamaa, Alain Crolotte, Ryan Voong, Mohammed Al-Kateb, Waleed Ghazal, and Roberto V. Zicari. 2017. BigBench V2: The New and Improved BigBench. In *ICDE 2017, San Diego, CA, USA, April 19-22*.
- [15] Ahmad Ghazal, Tilmann Rabl, Mingqing Hu, Francois Raab, Meikel Poess, Alain Crolotte, and Hans-Arno Jacobsen. 2013. BigBench: Towards An Industry Standard Benchmark for Big Data Analytics. In *SIGMOD 2013 (New York, New York, USA)*. 1197–1208.
- [16] Jim Gray. 1992. *Benchmark Handbook: For Database and Transaction Processing Systems*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- [17] Great Responsibility: The 2018 Big Data Great Power and Matt Turck AI Landscape. 2018. <http://mattturck.com/bigdata2018/>
- [18] Hadoop. 2018. hadoop.apache.org/.
- [19] Rui Han, Lizy Kurian John, and Jianfeng Zhan. 2018. Benchmarking Big Data Systems: A Review. *IEEE Trans. Services Computing* 11, 3 (2018), 580–597.
- [20] Hive. 2020. hive.apache.org/.
- [21] Han Hu, Yonggang Wen, Tat-Seng Chua, and Xuelong Li. 2014. Toward Scalable Systems for Big Data Analytics: A Technology Tutorial. *IEEE Access* 2 (2014), 652–687.
- [22] ICT, Chinese Academy of Sciences. 2015. BigDataBench 3.1. <http://prof.ict.ac.cn/BigDataBench/>
- [23] Impala. 2020. impala.apache.org/.
- [24] Intel. 2020. HiBench Suite. <https://github.com/intel-hadoop/HiBench>
- [25] Todor Ivanov. 2019. *Classifying, evaluating and advancing big data benchmarks*. Ph.D. Dissertation. Goethe University Frankfurt. <http://publikationen.uni-frankfurt.de/frontdoor/index/index/docId/51157>
- [26] Todor Ivanov, Patrick Bedué, Ahmad Ghazal, and Roberto V. Zicari. 2018. Adding Velocity to BigBench. In *Proc. of the 7th DBTest@SIGMOD 2018, Houston, TX, USA, June 15, 2018*. ACM, 6:1–6:6.
- [27] Todor Ivanov, Timo Eichhorn, Arne Berre, Tomás Pariente Lobo, Ivan Martínez Rodríguez, Ricardo Ruiz Saiz, Barbara Pernici, and Chiara Francalanci. 2019. Building the DataBench Workflow and Architecture. (2019).
- [28] Todor Ivanov and Rekha Singhal. 2018. ABench: Big Data Architecture Stack Benchmark. In *Companion of the 2018 ACM/SPEC ICPE 2018, Berlin, Germany, April 09-13, 2018*. ACM.
- [29] C. Kachris, B. Falsafi, and D. Soudris. 2018. *Hardware Accelerators in Data Centers*. Springer Int.
- [30] Alfons Kemper and Thomas Neumann. 2011. HyPer: A hybrid OLTP&OLAP main memory database system based on virtual memory snapshots. In *Proc. of the 27th ICDE 2011, April 11-16, 2011, Hannover, Germany*.
- [31] Doug Laney. 2001. 3D data management: Controlling data volume, velocity and variety. *META group research note* 6, 70 (2001), 1.
- [32] MADlib. 2020. madlib.apache.org/.
- [33] Min Li. 2015. SparkBench. <https://bitbucket.org/lm0926/sparkbench>
- [34] MLlib. 2020. spark.apache.org/mllib/.
- [35] Fatma Özcan, Yuanyuan Tian, and Pinar Tözün. 2017. Hybrid Transactional/Analytical Processing: A Survey. In *Proc. of the 2017 ACM SIGMOD 2017, Chicago, IL, USA, May 14-19, 2017*.
- [36] Muhammet Mustafa Ozdal. 2018. Emerging Accelerator Platforms for Data Centers. *IEEE Design & Test* 35, 1 (2018), 47–54.
- [37] Barbara Pernici, Chiara Francalanci, Angela Geronazzo, Polidori Lucia, Ray Stefano, Riva Leonardo, Arne Jørgen Berre, Ivanov Todor, et al. 2018. Relating Big Data Business and Technical Performance Indicators. In *Conference of the Italian Chapter of AIS*. 1–12.
- [38] Nicolas Poggi. 2018. *Microbenchmark*. Springer International Publishing, Cham.
- [39] Ivens Portugal, Paulo S. C. Alencar, and Donald D. Cowan. 2018. The use of machine learning algorithms in recommender systems: A systematic review. *Expert Syst. Appl.* 97 (2018).
- [40] Tilmann Rabl, Christoph Brücke, Philipp Härtling, Rodrigo Escobar Palacios, Hamesh Patel, Satyam Srivastava, Christoph Boden, Jens Meiners, and Sebastian Schelter. 2019. ADABench-Towards an Industry Standard Benchmark for Advanced Analytics. (2019).
- [41] Sherif Sakr, Anna Liu, and Ayman G. Fayoumi. 2013. The family of mapreduce and large-scale data processing systems. *ACM Comput. Surv.* 46, 1 (2013), 11:1–11:44.
- [42] Xuanhua Shi, Zhigao Zheng, Yongluan Zhou, Hai Jin, Ligang He, Bo Liu, and Qiang-Sheng Hua. 2018. Graph Processing on GPUs: A Survey. *ACM Comput. Surv.* 50, 6 (2018), 81:1–81:35.
- [43] Spark. 2020. spark.apache.org.
- [44] SparkSQL. 2020. spark.apache.org/sql/.
- [45] SPEC. 2019. <https://www.spec.org/>.
- [46] Tensorflow. 2020. tensorflow.org.
- [47] TPC. 2020. www.tpc.org/.
- [48] Anuj Vaishnav, Khoa Dang Pham, and Dirk Koch. 2018. A Survey on FPGA Virtualization. In *Proc. of the 28th FPL 2018, Dublin, Ireland, August 27-31, 2018*.
- [49] Kizheppatt Vipin and Suhaib A. Fahmy. 2018. FPGA Dynamic and Partial Reconfiguration: A Survey of Architectures, Methods, and Applications. *ACM Comput. Surv.* 51, 4 (2018).
- [50] Qingchen Zhang, Laurence T. Yang, Zhikui Chen, and Peng Li. 2018. A survey on deep learning for big data. *Information Fusion* 42 (2018), 146–157.