Daftar Pustaka

[1] Y. Chen, X. Qin, H. Bian, J. Chen, Z. Dong, X. Du, Y. Gao, D. Liu, J. Lu, and H. Zhang. A study of sql-on-hadoop systems. In *Workshop on Big Data Benchmarks, Performance Optimization, and Emerging Hardware*, pages 154–166. Springer, 2014.

[2] Cloudera^{TM}. Machine learning — analytics — cloud - cloudera. Accessed : June 3rd, 2018.

[3] N. Francis and S. Kurian K. Data processing for big data applications using hadoop framework. *International Journal of Advanced Reasearch in Computer and Communications Engineering*, 4(3):1–4, 2015.

[4] A. Hadoop^{TM}. Hdfs architecture guide. Accessed : June 1st, 2018.

[5] Hortonworks^{TM}. Data management platform. Accessed : June 3rd, 2018.

[6] X. Li and W. Zhou. Performance comparison of hive, impala and spark sql. In *Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2015 7th International Conference on*, volume 1, pages 418–423. IEEE, 2015.

[7] Q. Liu, Y. Fu, G. Ni, and J. Mei. Big data management performance evaluation in hadoop ecosystem. In *2017 3rd International Conference on Big Data Computing and Communications (BIGCOM)*, pages 413–421. IEEE, 2017.

[8] B. Marr. *Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results*. Wiley, Chichester, 2016.

[9] A. Mehmood, M. Iqbal, M. Khaleeq, and Y. Khaliq. Performance analysis of shared-nothing sql-on-hadoop frameworks based on columnar database systems. In *Innovative Computing Technology (INTECH), 2016 Sixth International Conference on*, pages 128–133. IEEE, 2016.

[10] A. ORC^{TM}. Apache orc - high performance columnar storage for hadoop. Accessed : August 16th, 2018.

[11] P. Pirzadeh, M. Carey, and T. Westmann. A performance study of big data analytics platforms. In *Big Data (Big Data), 2017 IEEE International Conference on*, pages 2911–2920. IEEE, 2017.

[12] X. Qin, Y. Chen, J. Chen, S. Li, J. Liu, and H. Zhang. The performance of sql-on-hadoop systems-an experimental study. In *Big Data (BigData Congress), 2017 IEEE International Congress on*, pages 464–471. IEEE, 2017.

[13] Y. Samadi, M. Zbakh, and C. Tadonki. Comparative study between hadoop and spark based on hibench benchmarks. In *Cloud Computing Technologies and Applications (CloudTech), 2016 2nd International Conference on*, pages 267–275. IEEE, 2016.

[14] TPC^{TM}. Tpc - benchmark. Accessed : June 1st, 2018.

[15] P. Vaidya and J. J. Lee. Characterization of tpc-h queries for a column-oriented database on a dual-core amd athlon processor. In *Proceedings of the 17th ACM conference on Information and knowledge management*, pages 1411–1412. ACM, 2008.