

---

## REFERENCES

- Kyong-Ha Lee, Yoon-Joon Lee, Hyunsik Choi, Yon Dohn Chung, Bongki Moon. *Parallel Data Processing with MapReduce: A Survey*. KAIST, Department of Computer Science. Korea University, Department of Computer Science and Engineering. University of Arizona, Department of Computer Science, 2011.
- Herodotos Herodotou, Shivnath Babu. *Profiling, What-if Analysis, and Cost-based Optimization of MapReduce Programs*. Duke university, 2011.
- Dominique Heger. *Hadoop Performance Tuning – A Pragmatic & Iterative Approach*. DH technologies, 2013.
- Dhruba Borthakur. *The Hadoop Distributed File System: Architecture and Design*. The apache Software Foundation, 2007.
- Konstantin Shvachko, Hairong Kuang, Sanjay Radia, Robert Chansler. *The Hadoop Distributed File System*. Sunnyvale, California USA, 2010
- Alan F. Gates, Olga Natkovich, Shubham Chopra, Pradeep Kamath, Shravan M. Narayanamurthy, Christopher Olston, Benjamin Reed, Santhosh Srinivasan, and Utkarsh Srivastava. *Building a high-level dataflow system on top of Map-Reduce: the Pig experience*. Proceedings of the VLDB Endowment 2, no. 2 (2009): 1414-1425, 2009.
- Christopher Olston, Benjamin Reed, Utkarsh Srivastava, Ravi Kumar, and Andrew Tomkins. *Pig latin: a not-so-foreign language for data processing*. Proceedings of the 2008 ACM SIGMOD international conference on Management of data, pp. 1099-1110. ACM, 2008.
- Alan F. Gates, Jianyong Dai, and Thejas Nair. *Apache Pig's Optimizer*. IEEE Data Engineering Bulletin 36, no. 1, 2013.
- Ashish Thusoo, Joydeep Sen Sarma, Namit Jain, Zheng Shao, Prasad Chakka, Suresh Anthony, Hao Liu, Pete Wyckoff, and Raghotham Murthy. *Hive: a warehousing solution over a map-reduce framework*. Proceedings of the VLDB Endowment 2, no. 2 (2009): 1626-1629.
- Ashish Thusoo, Joydeep Sen Sarma, Namit Jain, Zheng Shao, Prasad Chakka, Ning Zhang, Suresh Antony, Hao Liu, and Raghotham Murthy. *Hive-a petabyte scale data warehouse using hadoop*. In Data Engineering (ICDE), 2010 IEEE 26th International Conference on, pp. 996-1005. IEEE, 2010.

- Ronstrom, Mikael, and Lars Thalmann. *MySQL cluster architecture overview*. MySQL Technical White Paper, 2004.
- Oracle. *MySQL Cluster Evaluation Guide*. A MySQL Technical White Paper, 2013.
- <https://cwiki.apache.org/confluence/display/SQOOP/Home>
- [https://blogs.apache.org/sqoop/entry/apache\\_sqoop\\_overview](https://blogs.apache.org/sqoop/entry/apache_sqoop_overview)
- [http://sqoop.apache.org/docs/1.4.4/SqoopUserGuide.html#\\_introduction](http://sqoop.apache.org/docs/1.4.4/SqoopUserGuide.html#_introduction)
- <http://grouplens.org/datasets/movielens/>

