

Conclusion

By considering the above mentioned use case and comparing the existing system with proposed system we can say that the retrieval and processing of huge volumes of data using combiners will take less time when compared with the process without using combiners. The usage of combiner functions by the map reduce framework depends on the run time environment, if enough memory was available then the combiner functions will call ,otherwise the mapper directly sends the map function output to shuffle and sort functionality and then to reducer function and finally output phase.

This paper deals with architecture of map reduce framework and its processing explained without combiners and how the processing speed is being improved with use of combiners.

References

- [1] <https://en.wikipedia.org/wiki/MapReduce>
- [2] feng li, beng chin ooi, m. tamer özsu, sai wu, Distributed Data Management Using MapReduce, ACM, 2009
- [3] Seema Maitrey, C.K.Jha, Map Reduce: Simplified Data Analysis of Big Data, Elsevier, 2015
- [4] <http://ieeexplore.ieee.org/document/4125858/>
- [5] Mohd Rehan Ghazi, Durgaprasad Gangodkar, Hadoop, MapReduce and HDFS: A Developers Perspective, Elsevier, 2015
- [6] <http://data-flair.training/blogs/shuffling-and-sorting-in-hadoop>
- [7] Katarina Grolinger , Michael Hayes, Wilson A. Higashino, Alexandra L'Heureux, David S. Allison, Miriam A.M. Capretz, Challenges for MapReduce in Big Data, IEEE, 2014
- [8] <http://hadooptutorial.info/combiner-in-mapreduce>
- [9] Tzu-Chi-Huang, Kuo-Chih-Chu, Wei-Tsong Lee, Yu-Sheng Ho, Adaptive Combiner for MapReduce on cloud computing, ACM, 2014
- [10] <https://hadooptutorial.wikispaces.com/MapReduce>