BChain-Driven Scalable Approach to Big DataVerification of Db. Applications Processing

Abhilasha Singh[1], Akrati Sharma[2],

Email ID: abhilasha.singh@jlu.edu.in

akrati.sharma@jlu.edu.in

Abstract [Minimum 500 Words and Maximum 1000 Words]

In current world data is the key intergradient among all the organizations including IT sectors, academic fields, medical records and list goes on. Every sector deals with one common problem that is the management of such huge data generation that is commonly called as big data [1-2]. With the time data becomes historical as every day, every hour, every minute new data is generated. The main issues are arises when handling of such huge amount of data become an issue, because such data can neither be deleted as it might be useful for the organization nor be kept because it causes no memory space for newly generated data [4]. Hence to solve the issues one trending technology can be helpful that is block chain. Block chain [10-11] can be one of the promising technology which dealt easily with the problem occurred to manage big data,

2 Q's can be take care while combining these two terms together that is quality and quantity [15]. Quality in terms of efficient management of data so that without degrading the quality of data it can easily be managed without causing any issues with the newly generated data. Quantity in terms of as the name defines itself that is the generation of big data. Every day several organizations produce data in terabytes and zeta bytes. The reason behind using these two technologies together is the ones disadvantage becomes another's advantage. First one is security, such huge data generation can cause security loops. Maintain security of such big data is really a hazardous task and block chain dealt with security issues [31] quite well. The most important positive feature of block chain is decentralization [17] that means data does not belong to one single person; hence chances of data breaches are going downwards. Second one is flexibility, as big data contains every type of data like structured, unstructured, and semi-structured that include image files, video files, audio files etc. To work upon on various kinds of data at same time and same place again is not at all feasible. Here comes the working of block chain, as there is no limitation on block chain. Block chain [25] can easily worked upon variety of data.

As Big Data through block chain analysis is still a relatively young field in the collection of massive datasets recognition and gain on discoveries of certain patterns in the data. Since the certain patterns in the data are tremendously big and difficult data, it cannot be developed through the traditional data processing systems. To facilitate analyze such ever-growing amount of data to argues that block chain analysis should be pleasured as a new type of application for Big Data platforms in particular Map/Reduce,[7-8] to extract and analyze information from the block chain. Since all the information will be stored in the Block chain it will be convenient to access these details. Because critical operational data source of the design of the Block chain technology, users can view historical transactions effortlessly and need to ingest and analyze as part of their analytics.

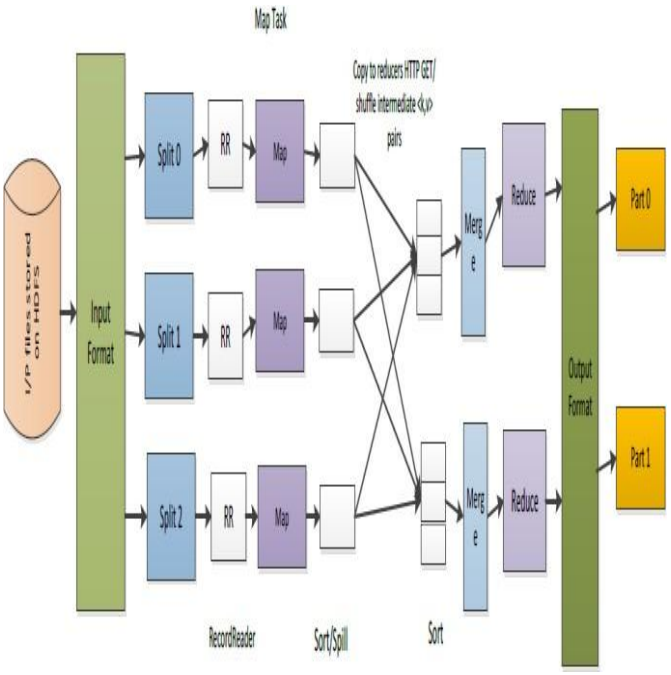

Fig.1. Data Flow of a Task supported by MapReduce

REFERENCES

[1] D Laney. 3-D data management: Controlling data volume, velocity and variety, Meta group, research note, February 2001. https://idoc.pub/documents/3d-data-management-controlling-data-volume-velocity-and-variety-546g5mg3ywn8

[2] Attila Marton, Michel Avital, and Tina Blegind Jensen. Reframing open big data. 2013. https://aisel.aisnet.org/ecis2013_cr/146/

[3] http://en.wikipedia.org/wiki/Apache Hadoop.

[4] Jeffrey Dean and Sanjay Ghemawat, MapReduce: Simplified Data Processing on Large Clusters, Google, 2004. https://www.usenix.org/legacy/events/osdi04/tech/full_papers/dean/dean.pdf

[5] M. Goudarzi, "Heterogeneous Architectures for Big Data Batch Processing in MapReduce Paradigm", IEEE Transactions on Big Data, Volume 5, Issue 1, March 2019, Pages 18 –33 https://dblp.org/db/journals/tbd/tbd5.html

[6] Gunasekaran Manogaran, Daphne Lopez and Naveen Chilamkurti, "In-Mapper Combiner based Map-Reduce Algorithm for Big Data Processing of IoT based Climate Data", Future Generation Computer Systems, Elsevier, Volume 86, September 2018, Pages 433-445 https://www.researchgate.net/publication/324638040_In-Mapper_combiner_based_Map-Reduce_algorithm_for_big_data_processing_of_IoT_based_climate_data

[7] Huan Ke, Peng Li, Song Guo, and Minyi Guo, "On Traffic- Aware Partition and Aggregation in MapReduce for Big Data Applications", IEEE Transactions on Parallel and Distributed Systems, Volume 27, Issue 3, March 2016, Pages 818 -828 https://jpinfotech.org/on-traffic-aware-partition-and-aggregation-in-mapreduce-for-big-data-applications/

[8] Marouane Birjali, Abderrahim Beni Hssane and Mohammed Erritali, "Evaluation of high level query languages based on MapReduce in Big Data", Journal of Big Data, Springer, Volume 5, Issue 36, December 2018, Pages 2-21 https://journalofbigdata.springeropen.com/articles/10.1186/s40537-018-0146-3

[9] S. K. Singh, S. Rathore, and J. H. Park, "Blockiotintelligence: A blockchain-enabled intelligent IoT architecture with artificial intelligence," Future Generation Computer Systems, vol. 110, pp. 721–743, 2020. http://jips-k.org/full-text/463

[10] W. Viriyasitavat and D. Hoonsopon, "Blockchain characteristics and consensus in modern business processes," Journal of Industrial Information Integration, vol. 13, pp. 32–39, Mar. 2019. https://www.researchgate.net/publication/326680277_Blockchain_Characteristics_and_Consensus_in_

Modern Business Processes

[11] L. Da Xu and W. Viriyasitavat, “Application of blockchain in collaborative internet-of-things services,” IEEE Transactions on Computational Social Systems, vol. 6, no. 6, pp. 1295–1305, 2019. https://www.sciencedirect.com/science/article/abs/pii/S221421262030822X

[12] C. Berg, S. Davidson, and J. Potts, Understanding the blockchain economy: An introduction to institutional cryptoeconomics. Edward Elgar Publishing, 2019. https://www.elgaronline.com/view/9781788974998/reference.xhtml

[13] Y. Yuan and F.-Y. Wang, “Blockchain and cryptocurrencies: Model, techniques, and applications,” IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 48, no. 9, pp. 1421–1428, 2018. https://www.researchgate.net/publication/326611239_Blockchain_and_Cryptocurrencies_Model_Techniques_and_Applications

[14] J. Abou Jaoude and R. G. Saade, “Blockchain applications– usage in different domains,” IEEE Access, vol. 7, pp. 45 360–45 381, 2019. https://www.researchgate.net/publication/344066677_A_Survey_on_Blockchain_for_Big_Data_Approaches_Opportunities_and_Future_Directions

[15] F. Casino, T. K. Dasaklis, and C. Patsakis, “A systematic literature review of blockchain-based applications: current status, classification and open issues,” Telematics and Informatics, vol. 36, pp. 55–81, Mar. 2019. https://www.sciencedirect.com/science/article/pii/S0736585318306324

[16] J. Zhang, S. Zhong, T. Wang, H.-C. Chao, and J. Wang, “Blockchainbased systems and applications: A survey,” Journal of Internet Technology, vol. 21, no. 1, pp. 1–14, Jan. 2020. https://pubmed.ncbi.nlm.nih.gov/31995857/

[17] H. Huang, W. Kong, S. Zhou, Z. Zheng, and S. Guo, “A survey of state-of-the-art on blockchains: Theories, modelings, and tools,” arXivpreprint arXiv: 2007.03520, 2020. https://link.springer.com/chapter/10.1007/978-3-030-75762-5_3

[18] A. K. Dey, C. G. Akcora, Y. R. Gel, and M. Kantarcioglu, “On the role of local blockchain network features in cryptocurrency price formation,” Canadian Journal of Statistics, vol. 48, no. 3, pp. 561–581, Sep. 2020. https://www.researchgate.net/publication/349652650_Knowledge_Discovery_in_Cryptocurrency_Transactions_A_Survey

[19] J. Angelis and E. R. da Silva, “Blockchain adoption: A value driver perspective,” Business Horizons, vol. 62, no. 3, pp. 307– 314, 2019. https://www.sciencedirect.com/science/article/abs/pii/S0007681318302088

[20] Y. Chen and C. Bellavitis, “Blockchain disruption and decentralized finance: The rise of decentralized business models,” Journal of Business Venturing Insights, vol. 13, p. e00151, Jun. 2020.

https://web.stevens.edu/ses/documents/fileadmin/documents/pdf/Blockchain_Disruption_and_Decentralized_Finance_The_Rise_of_Decentralized_Business_Models.pdf

[21] S. Moin, A. Karim, Z. Safdar, K. Safdar, E. Ahmed, and M. Imran, “Securing IoTs in distributed blockchain: Analysis, requirements and open issues,” Future Generation Computer Systems, vol. 100, pp. 325–343, 2019. https://www.sciencedirect.com/science/article/abs/pii/S0167739X18330851

[22] R. Zhang, R. Xue, and L. Liu, “Security and privacy on blockchain,” ACM Computing Surveys (CSUR), vol. 52, no. 3, pp. 1–34, Jul. 2019. https://dl.acm.org/doi/10.1145/3446373

[23] S. Leonardos, D. Reijsbergen, and G. Piliouras, “PREStO: A systematic framework for blockchain consensus protocols,” IEEE Transactions on Engineering Management, 2020, in press. https://deepai.org/publication/presto-a-systematic-framework-for-blockchain-consensus-protocols

[24] L. Zhang, Y. Xie, Y. Zheng, W. Xue, X. Zheng, and X. Xu, “The challenges and countermeasures of blockchain in finance and economics,” Systems Research and Behavioral Science, vol. 37, no. 4, pp. 691–698, Jul. 2020. https://www.worldscientific.com/doi/abs/10.1142/S2424862220500220

[25] A. Oussous, F.-Z. Benjelloun, A. A. Lahcen, and S. Belfkih, “Big data technologies: A survey,” Journal of King Saud University-Computer and Information Sciences, vol. 30, no. 4, pp. 431–448, Oct. 2018.

https://www.sciencedirect.com/science/article/pii/S1319157817300034

[26] Dai, Hong-Ning, Zibin Zheng, and Yan Zhang. "Blockchain for Internet of Things: A survey." IEEE Internet of Things Journal 6, no. 5 (2019): 8076-8094. https://www.researchgate.net/publication/333446530_Blockchain_for_Internet_of_Things_A_Survey

[27] Seo S, Jang I, Woo K, Kim I, Kim JS, Maeng S. HPMR: Prefetching and preshuffling in shared MapReduce computation environment. In: International Conference on Cluster Computing and Workshops. IEEE; 2009. p. 1–8.
https://www.sciencedirect.com/science/article/abs/pii/S1084804514001635

[28] Park J, Lee D, Kim B, Huh J, Maeng S. Locality-aware dynamic VM reconfiguration on MapReduce clouds. In: Proceedings of the 21st international symposium on High-Performance Parallel and Distributed Computing. New York, NY, USA: ACM; 2012. p. 27–36. ACM.
https://www.sciencedirect.com/science/article/abs/pii/S2214579615000167

[29] Sun, M., Zhuang, H., Zhou, X., Lu, K., & Li, C. (2014, August). HPSO: Prefetching based scheduling to improve data locality for MapReduce clusters. In International Conference on Algorithms and Architectures for Parallel Processing (pp. 82- 95). Springer, Cham.
https://link.springer.com/chapter/10.1007%2F978-3-319-11194-0_7

[30] Li, C., Zhang, J., Chen, Y., & Luo, Y. (2019). Data Prefetching and File Synchronizing for Performance Optimization in Hadoop-Based Hybrid Cloud. Journal of Systems and Software. Chicago.
https://www.sciencedirect.com/science/article/abs/pii/S0164121219300287

[31] X. Li, P. Jiang, T. Chen, X. Luo, and Q. Wen, “A survey on the security of blockchain systems,” Future Generation Computer Systems, vol. 107, pp. 841–853, Jun. 2020.
https://dl.acm.org/doi/abs/10.1145/3391195

[32] Y. Chen, J. Guo, C. Li, and W. Ren, “FaDe: a blockchain-based fair data exchange scheme for big data sharing,” Future Internet, vol. 11, no. 11, p. 225, 2019.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6408408/

[33] H. Hassani, X. Huang, and E. Silva, “Big-Crypto: big data, blockchain and cryptocurrency,” Big Data and Cognitive Computing, vol. 2, no. 4, p. 34, 2018.
https://www.mdpi.com/2504-2289/2/4/34/html