Reflection on the Popularity of MapReduce and Observation of Its Position in a Unified Big Data Platform

Xiongpai Qin[1,2], Biao Qin[1,2], Xiaoyong Du[1,2], and Shan Wang[1,2]

[1] MOE Key Lab of Data Engineering and Knowledge Engineering, Beijing, 100872, China
[2] School of Information, Renmin University of China, Beijing, 100872, China
{qxp1990,qinbiao,duyong,swang}@ruc.edu.cn,
qxp1990@sina.com

Abstract. In recent years MapReduce has risen to be the de-facto tool for big data processing. MapReduce is a disruptive innovation. It has changed the landscape of database market, the landscape of technologies, as well as the landscape of saying power. The article will give a reflection on the popularity of the technique and some observations of its position in a unified big data platform.

Keywords: MapReduce, Popularity, Reflection, Unified Big Data Platform.

1 The Popularity of MapReduce

MapReduce, introduced by Google in 2004[1] as a tool for huge volume of unstructured data processing, is rising to become the de-facto tool for big data processing. Industry has first recognized the value of MapReduce, and dozens of Hadoop (an open source implementation of MapReduce) based startups are launched to provide big data processing, analysis, and visualization solutions. To name a few of them, they are Cloudera, HortonWorks, MapR, Karmasphere, DataMeer, Aster Data, Greenplum, Hadapt, and Platfora. The startups share a common characteristic, they built their businesses upon the MapReduce technology. The strong wind of MapReduce swept through parallel computing community and aroused a tide of research during the period of 2006 to 2008, as well as another tide of research in database community during the period of 2009 to 2012.

The research has improved MapReduce in many aspects [2] [3], including: (1) Storage layout & data placement optimization, handling of data skew, index support, and data variety support; (2) Extension of MapReduce for stream processing, incremental processing, iterative processing; (3) two way join, multi way join, theta join optimization, parallelization of data mining/machine learning algorithms; (4) Schedule strategies for multi core CPU, GPGPU, heterogeneous environment, and cloud; (5) Easy to use interfaces for SQL query, statistical, data mining & machine learning algorithms, such as Hive, Pig, System ML, and Mahout; (6) Energy saving techniques, private and security guarantee for MapReduce.

The computing model of MapReduce is rather simple. MapReduce is a general execution engine that is ignorant of schemas and storage models, i.e. there is no

Y. Gao et al. (Eds.): WAIM 2013 Workshops, LNCS 7901, pp. 339–347, 2013.
© Springer-Verlag Berlin Heidelberg 2013

structure to the data. The runtime system automatically parallelizes computing tasks across a large cluster of commodity hardware, handles failures and manages disk I/O & network efficiency. The user only needs to provide a map function (which is applied to all input rows of the dataset to produce an intermediate output) and a reduce function (who will aggregate the intermediate results to produce the final result). However, MapReduce is not that simple. MapReduce can do analytic works beyond computing statistics from up to PB of data, or simply performing ETL (extract, transform, and load) work. Various data analytics algorithms have been migrated onto the MapReduce platform, including data summarization and reporting, multi-dimensional analysis and OLAP, data mining, machine learning, information retrieval, text mining and sentiment analysis, science data processing, as well as social network analysis and general graph analysis. The performance of the algorithms is improved, and application of the MapReduce technology has expanded to many domains. Some people think that such works are reinventing the data warehousing and BI techniques that have been created over the years, the authors believe that the so-called "reinventions" are necessary and may lead to important innovations.

With the growing influence of MapReduce, Traditional vendors in the data market noticed the popularity of MapReduce. IBM moves quickly with its *Big Insights* Plan, the plan tries to integrate DB2, Netezza, Hadoop, and SPSS into a big data analytic platform. EMC, formerly not as a database vendor, became a big player in the market overnight through acquiring Greenplum, which has combined the PostgreSQL database, shared nothing architecture and MapReduce to create the core product. TeraData acquired Aster Data to obtain its experience of MapReduce engineering as well as the analytic software package using MapReduce-style parallelism. Most funny is that, several venders who looked down on MapReduce before finally change the tough minds, Microsoft rejected MapReduce in 2009, in the year of 2012 it has closed the Dryad project (an equivalent parallel computing framework of MapReduce by Microsoft) [4] and warmly embrace Hadoop; Oracle despised MapReduce early in 2011, eventually published its *Big Plan* late in 2011, which incorporates noSQL and Hadoop technologies into the whole picture.

The penetrating power of MapReduce has changed the landscape of market, i.e. traditional venders vs. dozens of startups, the landscape of technology, i.e. RDBMS vs. noSQL/Hadoop, as well as the landscape of saying power, we see that many researchers previously not in the database community now rush into the field of data processing and analytics, and bring forth their innovative ideas, which deserve serious notice.

2 Reflection – Why MapReduce Is So Popular

With the price of the storage device going down and the capacity of it increasing, development of internet applications (including e-commerce, social network …), the requirements of science simulation and research etc., it is the first time that people collect some data sets with huge volumes. The era of Big Data is coming. To describe big data, people have agreed on three *V*s – *Volume, Variety, and Velocity*. Firstly, the volume of the data has exceeded the capability of traditional data processing tools (specifically RDBMS, relational database management systems). Secondly, people

need to handle various types of data to extract insightful information for decision making. The data is essentially multi-structured, including semi-structured data, unstructured data as well as structured data. Finally, the speed of data generation is very fast, the data emitted from sensors can be taken as an example, which needs to be handled in a timely manner.

The three Vs raise a range of challenges, among which the volume is the dominant one, followed by data variety. On one hand, handling various types of data of big volume with high velocity is beyond traditional tool's capability, on the other hand, although RDBMS technology has advanced dramatically over the past several decades, the computing paradigm are shifting from big server-based computing to distributed computing on clusters of low-cost commodity hardware for cost reason. MapReduce was born at the right time.

2.1 The Debate and the Outcome

Some prominent database researchers argue that there is nothing new about systems like MapReduce and that they're actually a step backward [5]. The argument has aroused a fierce debate. Is that true? The answer is no. The Apache Hadoop open source software project won the top prize of 2011 Media Guardian Innovation Awards [6], described by the judging panel as a "*Swiss army knife of the 21st century*".

Actually database researchers should be more open-minded. Sometimes the strong sense of glory and pride that has been brought about by the huge success of RDBMS will blind their eyes. The rising of instead of dying out of MapReduce has been beyond their expectation, and most annoying is that MapReduce begins to expand its application from web search to territories used to be occupied by relational database systems, including OLAP, data mining, machine learning, and information retrieval. Hadoop has become the winner of the ***Jim Gray prize*** for the fastest sort of a terabyte of data. In 2008, Hadoop required 209 seconds to analyze the terabyte, using a 900 node cluster. In 2009, it won the prize for an analysis that took only 62 seconds, running on 1,500 nodes. Some criticisms on MapReduce are not so rational, conversely the criticisms justify the promise of MapReduce technology.

The authors would like to clarify some wrong saying about MapReduce, including:

(1) MapReduce is Map function plus Reduce function: No, that is not true. MapReduce is a general parallel computing framework for big data processing. It includes the highly fault tolerant distributed file system, the MapReduce parallel computing model, and the running time that taking care of parallel running of applications on a large cluster. The whole MapReduce ecosystem has several more peripheral add-ons.

(2) MapReduce can only do some batch processing: No, that is not true. After Google published its Dremel [7]. Several venders are trying to provide interactive ad-hoc query processing capability onto Hadoop, including Cloudera *Impala*, HortonWorks *Stinger*, Apache *Drill*, and EMC *HAWQ*. Dremel achieves interactive ad-hoc query capability on big data by combining multi-level execution trees and columnar data layouts. It is capable of running aggregation queries over trillion-row tables in seconds. Dremel can scale to thousands of CPUs and PB of data. Now the technology has been used by thousands of users inside Google.

(3) MapReduce can only handle un-structured data, or is more suitable for un-structured data handling. No, that is not true. One example to counter argue the point is the RCFile [8] of FaceBook (together with Ohio State University and Chinese Academy of Science). RCFile is a storage layout optimization for Hadoop data blocks for high performance of structured data processing. The work of RCFile is inspired by the PAX idea from RDBMS research, it applies an elaborate structure to HDFS (Hadoop Distributed File System) blocks to achieve higher performance of data accessing while fully retaining nice properties of scalability and fault tolerance of MapReduce. In RCFile, firstly data (structured table) is horizontal partitioned into blocks. In each block, the data is further broken into columns and every column is store contiguously in the block. Compression techniques can be used to reduce space consumption because the data is stored in blocks using a columnar layout. RCFile has been deployed to production by FaceBook, and plays an important role in daily Facebook operations. From the example, we can see that, by applying some elaborate structure to HDFS blocks, MapReduce can handle structured data as well. Actually the MapReduce framework is ignorant of the underlying storage layout of data blocks, it depends on applications to interpret the data at run time, no matter what the storage layout of the data is.

2.2 Reflection – Application (User) Requirement is the King

Taking TV as an example, we would like to comment that - the user requirement is finally the king, not the technologies. People have the desire to entertain themselves with Audio and Video devices. TV is a good product for that purpose. TV technology has evolved from CRT, to LCD, and then to LED recently, the technology keep changing, but the desire of people for better playback of Audio and Video never change, and technologies just serve the purpose of entertainment and information acquirement of people. Technologies are means, not objectives.

Similarly, when the big data era is coming and RDBMS can not completely address the challenges, MapReduce can be an alternative/complementary tool to RDBMS. What people need is to handle various types of data (structured data, unstructured data, and semi-structured data), data at different stages (data in motion, data at rest, and archived data) to extract useful insights, what tools are used to do the job doesn't matter in some sense. Actually the preferred architecture model for web scale data processing of parallel computing over large scale low cost server-based clusters has set off an innovation path that outpace the traditional database market.

Recently what is a really pressing issue is to blend the two technologies of RDBMS and Hadoop for multi-structured data analytics (next section).

3 Positioning MapReduce in a Unified Big Data Platform

RDBMS has been studied for more than 40 years, optimization techniques across different levels from storage layer to execution layer have been investigated, and many of them have been implemented in RDBMS. An ecosystem (vendors, products,

tools, services…) built around RDBMS has been around for several decades. MapReduce was born for large dataset processing, it is designed with highly scalability and highly fault tolerance as top considerations. A big data analytics ecosystem built around MapReduce is emerging alongside the traditional one built around RDBMS.

When it comes to data analytics, the objectives of RDBMS and MapReduce as well as the ecosystems built around them, overlap much really, in some sense they do the same thing and MapReduce can even accomplish more works, such as graph processing, which RDBMS can not handle well. Why two techniques/ecosystems for one thing? People are envisioning convergence of the two.

While traditional venders such as IBM, Microsoft, TeraData and Oracle reestablish their positions in the big data market by adopting Hadoop, so many startups such as EMC (Greenplum), HortonWorks, Cloudera etc. will cut off a large portion of the market. The big data platforms they provide fall in three categories: (1) *Co-Exist* solutions; (2) *SQL with MapReduce Support* solutions; and (3) *MapReduce with SQL Support* solutions.

3.1 Co-exist Solution – Different Tools for Different Jobs Plus High Speed Data Integration

The big data solutions provided by IBM and Oracle fall into this category.

IBM is a fast mover in the big data era. Its Watson computer, which has won the game of Jeopardy in 2011, was powered by Hadoop technology. IBM has announced its *Big Insights* in 2010. Since IBM has so many data management and processing product lines, it is natural to combine these underlying technologies to provide a big data processing environment, which is called *Big Data Platform* by IBM [9].

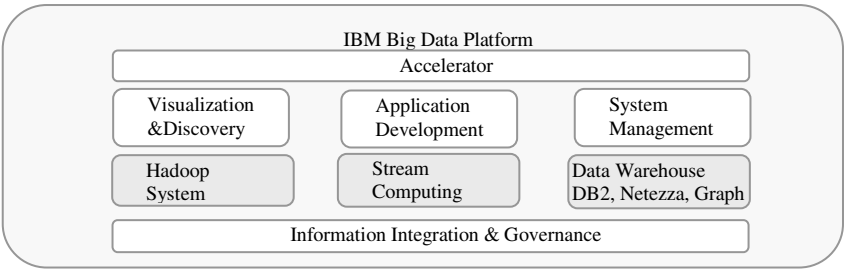

Fig. 1. The Architecture of IBM's Big Data Platform [9]

There are several components in the IBM Big Data Platform, including basic data management and processing engines such as Hadoop system, Stream system, and Data warehouse system, visualization & discovery tools, applications development tools, system management tools, as well as information integration and governance facilities. For IBM, the big data analysis is viewed as a workflow of analytic tasks instead of a single analytic task. These analytical processes can run across multiple tools in the Big Data Platform to process multi-structured data (both structured and un-structured data).

Oracle was rather reluctant to accept noSQL & Hadoop in the early of 2011. In the end of 2011, under the great pressure of EMC and TeraData, as well as many other data warehouse startups that based their business on the Hadoop technology, Oracle released its *Big Plan.* In the *Big Plan*, Oracle adopts a combination and co-exist strategy [10], which resembles IBM's Big Data Platform.

IBM and Oracle provide connectors to facilitate the process of data integration, including tools for loading data from Hadoop into database, accessing HDFS data from database, transparent access to Hadoop from statistical package etc.

3.2 SQL with MapReduce Support

Microsoft PolyBase, EMC Greenplum, TeraData Aster Data fall in this category.

In PolyBase [11], Users can create external tables on HDFS data. This allows queries to reference data stored in HDFS as if it were loaded into a relational table. Users can seamlessly perform joins between tables in PDW and data in HDFS. PolyBase use a cost based optimizer to determine on whether the job is done on PDW or on Hadoop HDFS according to some statistics gathered before. Future version of PolyBase will further utilize the huge computation and I/O power of the Hadoop cluster by offloading more work to Hadoop for processing, even for queries that only reference PDW-resident data.

EMC became a big player in the market by acquiring Greenplum. In the whole picture of EMC *Unified Analytics Platforms (UAP),* Greenplum DATABASE and Greenplum HD, two data storage and processing engines, are working in a unified and interaction manner to co-process all of user data. The Greenplum DATABASE is built upon PostgreSQL database for structured data processing. Greenplum HD is responsible for unstructured data processing with Hadoop. Actually with the *HAWQ* technique, Greenplum HD can perform SQL queries over Hadoop resident data. *Dynamic Pipelining*™ technology, together with many other implementation optimizations, deliver a 100X to 600X performance improvement over other interactive Hadoop SQL solutions such as *Impala* of Cloudera according to EMC [12].

To response to the EMC acquirement of Greenplum, TeraData acquired Aster Data to maintain its competitive edge in the market. TeraData believes that different data carries different value, comes in different formats, requires different analytics, and serves different business objectives [13]. It is reasonable to utilize multiple technologies, including: data staging, data discovery, data warehouse to processing the data. TeraData integrates open source Hadoop, Aster Data discovery platform, and TeraData data warehouse into the *TeraData Unified Data Architecture* ™, which act as the staging tool, the discovery tool, and the data warehouse respectively. The whole architecture resembles IBM's Big Data Platform.

But wait a minute, when we take a closer look at it, we can see that Aster Data database is similar to EMC *Unified Analytics Platforms (UAP)* in itself. Aster Data database is a massively parallel (MPP) analytic platform, both SQL and MapReduce analytic processing capabilities are embedded in the product through *SQL-MapReduce*® [14] technology. Aster Data database can seamlessly access multi-structured data in database table and HDFS. Actually Aster Data database can be a replacement of TeraData database, used as a standalone data processing platform.

3.3 MapReduce with SQL Support

The Apache *Drill* project, solutions of startups such as HortonWorks, Cloudera, Hadapt, Platfora, as well as next generation of Hadoop fall in this category.

In a small territory, HortonWorks and Cloudera are competitors, because they all provide enhanced Hadoop solutions for enterprise adoption. In the basic Hadoop ecosystem, there are common components such as HDFS, MapReduce, HBase, Pig, Hive, Meta Data Services, Management and Monitoring Services, as well as Data Integration Services. Besides basic components, HortonWorks and Cloudera provide add-on tools to facilitate enterprise adoption of Hadoop as the de-facto big data platform, including high available capability, management and operation tools.

As for SQL support, HortonWorks and Cloudera introduce *Stinger* and *Impala* respectively. *Stinger* and *Impala* are similar in that they try to provide the interactive SQL-like query capability over Hadoop, as a replacement or enhancement of Hive. Apache Hive and its HiveQL language have become the standard SQL interface for Hadoop since introduced by Facebook in 2007. Hive is often criticized for its low performance. But now the situation is changing.

Stinger is an initiative launched by HortonWorks, which is to make Hive much faster, so that people can run ad-hoc query on Hadoop interactively. HortonWorks plans to realize the *Stinger* in a few steps. Firstly, they are making Hive a more suitable tool for people to perform decision support queries on Hadoop by adding new features to the language and making Hive's system a model more like the standard SQL. Secondly, they are improving the Hive's query optimizer for better query execution plans, and Hive's execution engine for higher throughput of data processing. Thirdly, they develop a new columnar file format for higher performance of analytic tasks. Lastly, they are introducing a new runtime framework - *Tez*, to reduce Hive's latency and throughput constraints [15].

Cloudera tries to push Hadoop beyond a batch analytic tool to an interactive analytic tool with *Impala*. *Impala* is a real-time query engine that allows users to query data stored in HDFS and HBase in seconds via a SQL interface. Instead of using MapReduce, *Impala* uses its own processing framework to execute queries over HDFS, which results in a 10x-50x performance boost compared to Hive.

Apache *Drill* is design to scale to 10,000 servers or more and to be able to process Petabytes of data and trillions of records in seconds. The *Drill* architecture consists of four key layers [16]: (1) Query languages: This layer is responsible for parsing the user query and constructing an effective execution plan. *Drill* is designed to support SQL-like language (called DrQL) as well as other languages; (2) Low-latency distributed execution engine: The layer provides the scalability and fault tolerance needed to query Petabytes of data efficiently on a large cluster; (3) Data formats: various data formats, including column-based format, schema-less formats, and row-based formats are to be supported in this layer; (4) Data sources: Hadoop HDFS etc.

The interactive capabilities that these tools provide will refute the bias saying that "Hadoop can only do some batch processing of data".

Hadapt is the commercialized version of the HadoopDB [17] of Yale University. Hadapt combines the scalable architecture of Hadoop with a hybrid storage layer that integrates a relational data store and HDFS, by consolidating data into a single

platform Hadapt eliminates the need to move around data inside the system. The *Hadapt Interactive Query* tries to deliver interactive capability on Hadoop, like Cloudera *Impala*, Apache *Drill*, EMC Greenplum *HAWQ*, and HortonWorks *Stinger*. Hadapt also provides *Hadapt Development Kit*™ (HDK) for easier application development.

Platfora is a startup launched in the late of 2012, which deserves serious notice. After series A and series B fund securing, it has drawn in a total of more than $25.7M. The great ambition of Platfora is to make Hadoop an interactive big data platform by using in memory data processing (aggregation and cache) technology, data compression, and columnar data layouts. Platfora is a big data solution completely built upon Hadoop [18].

The current Apache Hadoop MapReduce framework has several limitations, such as the scalability limit of around 4,000 machines. Next generation of Hadoop is on the horizon. The major change to current version of Hadoop is to separate the two functions of the Job Tracker (resource management and job scheduling/monitoring) into different components. The new Resource Manager manages the global assignment of compute resources to applications, and the Application Master manages the application's scheduling and coordination [19]. The re-architecture leads to a new framework that scales out more easily. After decoupling MapReduce computing paradigm from the resource management, new application types can be plugged into Hadoop, including stream processing, graph processing, bulk synchronous processing (BSP, for graph data processing) and message passing interface (MPI, many parallel computing applications use the MPI computing model) [20]. Next generation of Hadoop will be greatly enhanced in terms of scalability, availability and performance. And it is expected to be the underpinning technology for Hadoop-based unified big data platforms, competing with RDBMS-based ones. The days RDBMS dominates the whole market has gone.

4 Conclusions

When people need to analyze big volume of multi-structured data for insightful information, Hadoop is an indispensable component in a unified big data platform. There are *Co-Exist, SQL with MapReduce Support, MapReduce with SQL Support* solutions for combining RDBMS and MapReduce. RDBMS and Hadoop are the tools to serve our goal of data management and analytics, user requirements are the ultimate forces that drive the innovation, tools in themselves cannot be the objectives. Database community should be open to incorporate innovative ideas (such as MapReduce…) from other sub-disciplines of computer science (such as parallel computing…) to tackle the challenges of big data processing.

Acknowledgements. This work is partially funded by the NSF of China under Grants No. 61170012 & 61170013, the Funds of Renmin University of China under Grant No. 10XNJ048, and EMC Global CTO Office.

References

1. Dean, J., Ghemawat, S.: MapReduce: simplified data processing on large clusters. In: Symposium on Operating Systems Design and Implementation (OSDI), pp. 137–150. USENIX Association, San Francisco (2004)
2. Lee, K.H., Lee, Y.J., Choi, H., Chung, Y.D., Moon, B.: Parallel data processing with MapReduce: a survey. SIGMOD Record 40(4), 11–20 (2011)
3. Sakr, S., Liu, A., Fayoumi, A.G.: The Family of MapReduce and Large Scale Data Processing Systems (2013), http://arxiv.org/abs/1302.2966
4. Foley, M.J.: Microsoft drops Dryad; puts its big-data bets on Hadoop (2011), http://www.zdnet.com/blog/microsoft/microsoft-drops-dryad-puts-its-big-data-bets-on-hadoop/11226
5. Kraska, T.: Finding the Needle in the Big Data Systems Haystack. IEEE Internet Computing 17(1), 84–86 (2013)
6. Winckler, M.: Apache Hadoop takes top prize at Media Guardian Innovation Awards (2011), http://www.guardian.co.uk/technology/2011/mar/25/media-guardian-innovation-awards-apache-hadoop
7. Melnik, S., Gubarev, A., Long, J.J., Romer, G., Shivakumar, S., Tolton, M., Vassilakis, T.: Dremel: Interactive Analysis of WebScale Datasets. Proceedings of the VLDB Endowment 3(1-2), 330–339 (2010)
8. He, Y., Lee, R., Huai, Y., Shao, Z., Jain, N., Zhang, X., Xu, Z.: RCFile: A Fast and Space-efficient Data Placement Structure in MapReduce-based Warehouse Systems. In: International Conference on Data Engineering (ICDE), pp. 1199–1208. IEEE Computer Society, Hannover (2011)
9. Ferguson, M.: Architecting a Big Data Platform for Analytics. A Whitepaper Prepared for IBM (2012)
10. Oracle: Oracle: Big Data for the Enterprise. Oracle White Paper (2012)
11. Dewitt, D.: Polybase: What, Why, How. SQL PASS Summit Keynote (2012)
12. EMC: Unified Analytics Platform (2013), http://www.greenplum.com/products/greenplum-uap
13. TeraData: TeraData Unified Data Architecture. TeraData Whitepaper (2012)
14. Friedman, E., Pawlowski, P., Cieslewicz, J.: SQL/MapReduce: A practical approach to self describing, polymorphic, and parallelizable user defined functions. Proceedings of the VLDB Endowment 2(2), 1402–1413 (2009)
15. Gates, A.: The Stinger Initiative: Making Apache Hive 100 Times Faster (2013), http://hortonworks.com/blog/100x-faster-hive/
16. Incubator Wiki: Drill Proposal (2013), http://wiki.apache.org/incubator/DrillProposal
17. Abouzeid, A., Bajda-Pawlikowski, K., Abadi, D., Silberschatz, A., Rasin, A.: HadoopDB: an architectural hybrid of MapReduce and DBMS technologies for analytical workloads. Proceedings of the VLDB Endowment 2(1), 922–933 (2009)
18. Platfora: Platfora Homepage (2013), http://www.platfora.com/
19. Murthy, A.C.: The Next Generation of Apache Hadoop MapReduce (2011), http://developer.yahoo.com/blogs/hadoop/posts/2011/02/mapreduce-nextgen/
20. BUSINESS WIRE: HortonWorks to Deliver Next-Generation of Apache Hadoop (2012), http://www.businesswire.com/news/home/20120119005825/en/Hortonworks-Deliver-Next-Generation-Apache-Hadoop