Mechanical Property of Composite Material Based on Map-Reduce Model

WU Jun[1,2 a] and HUANG Zhixiong[1,b]

[1]School of Materials Science and Engineering, Wuhan University of Technology, Wuhan 430070, China

[2] School of Computer Science, Hubei University of Technology, Wuhan, 430068, China

[a] wujun@whut.edu.cn[1], [b] lavazza@foxmail.com

Abstract. In this Big Data era, all the scientific research areas be needed by analyzing and mining knowledge from large-scale data sets. Mechanical Property of Composite Material is increasingly being used to improve our industrial product but how to get the advantages and challenges derived from Big Data is important part in vast amount of Composite Material data. Map-Reduce as an open source implementation would be a high reliability, high fault tolerance capability of parallel computing software framework. Map-Reduce based applications can be run in large clusters of parallel processing large data sets. this paper discusses the definition Mechanical Property of Composite Materials research by the parallel processing of Map-Reduce model. The performance this model of new material was established by using the method of Map-Reduction provided the basis for the performance optimization.

1. Introduction

With the rapid development of information technology in twenty-first Century, it is urgent to work on the new research and production for traditional composite materials into a high-tech, new technology, big data oriented information analysis model. Our paper intends to use the Big Data platform and intelligent data analysis methods. With this kind of soft power, it is easy to reduce energy consumption of composite materials, be shorten the composite material innovation cycle, explore the law of sustainable development of composite materials. This mainly for composite materials based on Hadoop platform for big data characteristic analysis, from following 6 E areas: Energy conservation、Exuberance、Easement、Environmental protection、Ecology and Economy which could work on composite materials research and development and production of integrated relational analysis and evaluation, to composite materials R & D and production in the process of integration of advanced and efficient information technology, of composite large related data analysis, mining to obtain effective data in industrial production from exploring new rules for the R & D and production of composite materials.

---

* Corresponding author: wujun@whut.edu.cn

2. Background

In the 1980s with the new material exploration and research when First‑Principle Calculations method played an important role, it is important to discover those new materials, such as High temperature superconducting material, Super hard material, Nano material, Artificial low dimensional quantum structure materials. The first principle calculation is also called as Calculated based on the design of quantum theory. And its basic method has the solid quantum theory and the quantum chemistry theory. The rise of first principle calculations based on quantum mechanical techniques takes a prominent place in computational material science with the help of supercomputers [4]. The movement towards creating a material database using simulation opens a new way of seeking and designing new materials [5–8]. Such movement is also occurring in the field of catalysts where desired chemical reactions are evaluated using first principle calculations [9-10]. With the rapid increase of supercomputer development, creating material databases using first principle calculations is gaining popularity, yet accessing and manipulating the data for predicting new undiscovered materials has yet to be realized. Especially suitable for the calculation and design of materials for atomic scale, Nano scale engineering, materials for many devices, and materials for electronic devices.

Recent studies show that the data mining and some other intelligent computing technology used in material design, test can be obtained with fewer ideal materials, to achieve a multiplier effect. For example, composite materials R & D and production, need to consider whether the integration with the surrounding environment. Through R & D and production of composite multi attribute relevance analysis, obtained the excellent lightweight, heat insulation, sound insulation, fire, earthquake and other performance, not only can make the composite towards the direction of diversification and greatly extends the application field of research and development and production of the new composite materials, to promote the technology development and innovation of the composites industry has important significance.

D'Orazioand Bimonte[11-12]tried to store big data in multidimensional array and apply the storage model to Pig(Apache,2013a,b), which is a data analysis tool an dis-based on Hadoop. Then, the storage model is proved efficient by experiment. The data storage and data model are discussed, in addition we adopt the dimension coding method and come up with dimension related algorithms and OLAP algorithm to boost the OLAP performance.

Yue[13]introduce epic, an elastic power-aware data-intensive cloud platform for supporting both data intensive analytical operations and online transactions, a storage system, supporting OLAP and OLTP through index, data partition, was introduced.

3. Map-Reduce Model

Map-Reduce Hadoop as an open source implementation is a high reliability, high fault tolerance capability of parallel computing software framework. Map-Reduce based applications can be run in large clusters of parallel processing large data sets. Simplifies concurrent programming model provides application programming interface (API) to the user through the map reduce, make not familiar with parallel computing users can easily develop map reduce application, and can reduce the amount of repeated work, map reduce execution flow chart is as follows Fig 1. [13]
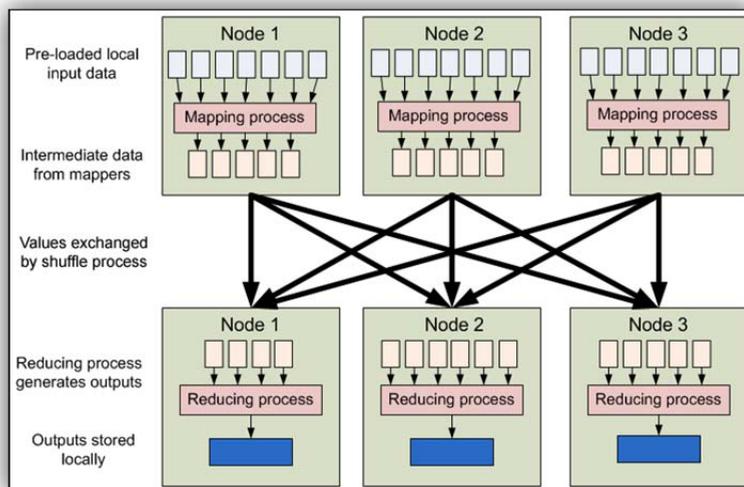
Fig 1 Map-Reduce structural work

The performance of Map-Reduce optimization progress rapidly, its performance is gradually approaching the relational database. The direction of research and two directions: theoretical focus to improve the Map-Reduce by using relational database technology and theory of performance; industry focus based on Map-Reduce platform, the application software. In the field of data warehouse, we consider the following research direction is more important and current researches are also less involved:

Multi-dimensional Data is Pre Computed.

Map-Reduce more for is a one-time operation analysis. Data analysis operation although it is difficult to predict, but traditional analysis, such as based on reports, and multidimensional data analysis still account for the majority. Therefore, Map-Reduce platform can also accelerate the speed of data analysis by making use of pre-computation. Based on storage space considerations (you can imagine, on the explosion of data computing data cube need to pay expensive storage space cost), MOLAP is not desirable, hybrid OLAP should is the Map-Reduce platform optimization of OLAP implementation scheme.

Specific research such as based on the Map-Reduce framework for efficient cube computation algorithm; II materialized view selection problem material which data; 3 different operation analysis of physicochemical methods (such as predictive analysis and operation) and how to based on materialized data for analysis of complex operations such as data access path selection problem).

Operation Analysis of the Parallel Implementation.

Data analysis is required for efficient complex statistical analysis function support. IBM will open-source statistical analysis of R software integrated into Hadoop platform and enhance the Hadoop statistical analysis function. But the more challenging problem is, such as what analysis algorithm based on Map-Reduce framework design parallelization and efficient. In particular, needs to be emphasized, in view of the enormous cost of mobile data, the algorithm should be based on the mobile computing to achieve.

Query Sharing.

Map-Reduce using step by step in physical and chemical treatment, resulting in high cost and network transmission cost of I / O. A effectively reduce the cost of the way is materialized intermediate results sharing among multiple queries, and even the original data,

by sharing costs and avoid repeated computation. Therefore, how to in multi query sharing intermediate results would be a very practical value of the research.

User Interface.

How to achieve a better display of data analysis and operation, especially the visual display of complex analysis operation.

Hadoop Reliability Research.

The current Hadoop using master-slave structure, which determines the main once a node failure, will appear the situation of failure of the whole system. Therefore, how to improve reliability of the master node does not affect the Hadoop existing implementations of the promise, will be a practical research.

Data Compression.

The execution model of Map-Reduce determines its performance depends on the I/ O and the network transmission costs. In the parallel database and the Map-Reduce performance based on the compressed data, found that the compression technology and no performance improvement of Hadoop. But the actual situation is that the compression can not only save space and save I/ O and network bandwidth, can also use the multi-core parallel computing ability of the current CPU, I/ O and CPU processing capacity balance, so as to improve the performance of parallel database using data compression. For example, the performance can often dramatically. Since then, the research literature the successful use of compression technology to enhance the performance of the Hadoop. But these studies are based on the storage model of their own, instead of the default Hadoop memory model (row store model). Therefore, Map-Reduce compression is an important question remains to be research Questions.

Multidimensional Index Research. Based on the Map-Reduce Framework to Realize the Multidimensional Index, Accelerate Multidimensional Data Retrieval Speed.

Map-Reduce mainly consists of two core operations: mapping (Map) and protocol (Reduce). The mapping (Map) method is used to get a set of key value pairs (key/value) mapping into a new set of key value pairs (key/value). Protocol (Reduce) is used to ensure that all the keys to share the same key group. By running the Map-Reduce computing framework and Hadoop distributed file system (HDFS) on the same node, the cluster can use the network bandwidth to achieve efficient task scheduling.

4. Acknowledgments

References

1. Nongnuch Artrith, Alexander Urbanb. An implementation of artificial neural-network potentials for atomistic materials simulations: Performance for TiO2. Computational Materials Science. 2016(114): 135-150.
2. Ivanovskii A L. Mechanical and electronic properties of diborides of transition 3 d –5 d metals from first principles: Toward search of novel ultra-incompressible and superhard materials. Progress in Materials Science, 2012, 57(1):184-228.

3. Yao M, Qi M, Li J, et al. A novel classification method based on the ensemble learning and feature selection for aluminophosphate structural prediction. Microporous & Mesoporous Materials, 2014, 186(3):201–206.

4. Goncharov A A, Dub S N, Agulov A V, et al. Structure, composition, and mechanical properties of thin films of transition metals diborides. Journal of Superhard Materials, 2015, 37(6):422-428.

5. Bamakan S M H, Gholami P. A Novel Feature Selection Method based on an Integrated Data Envelopment Analysis and Entropy Model . Procedia Computer Science, 2014, 31:632-638.

6. James E. Saal, Scott Kirklin.Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database,JOM,November 2013, Volume 65, Issue 11,1501-1509

7. Worku H M. Joint Modeling of Clustered Data: Prenatal Exposure to Endocrine Disrupting Compounds and their Impact on Newborns in Flanders, Belgium. Journal of Superhard Materials, 2015, 37(6):422-428.

8. Zerr A, Miehe G, Li J, et al. High-Pressure Synthesis of Tantalum Nitride Having Orthorhombic U_2S_3 Structure. Advanced Functional Materials, 2009, 19(14):2282–2288.

9. Sealy C. Synthesis of ultrahard material under no pressure : Mechanical properties. Materials Today, 2007, 10(6):8.

10. Goncharov A A, Dub S N, Agulov A V, et al. Structure, composition, and mechanical properties of thin films of transition metals diborides. Journal of Superhard Materials, 2015, 37(6):422-428.

11. D'Orazio,L.,Bimonte,S.,2010.Multidimensionalarraysforwarehousingdataonclouds,In: 3rdInternationalConferenceonDataManagementinGridandPeer-to-PeerSystems,Globe2 010.SpringerVerlag,Bilbao,Spain,September1–2,2010:26–37.

12. Taleb,A.,Eavis,T.,Tabbara,H.,2013.QueryoptimizationfortheNOXOLAPalgebra,Trans actionsonLarge-ScaleData-andKnowledge-CenteredSystemsVIII.Springer,53–88.

13. Yu,C.,Chun,C.,Fei,G.,etal.,2011.ES2:aclouddatastoragesystemforsupportingbothOLT PandOLAP,27thIEEEInternationalConferenceonDataEngineering(ICDE2011).Piscata way,NJ,USA,11–16April,2011:291–302.

14. www.google.com