National Research University Higher School of Economics

as a manuscript

Tyryshkina Yevgeniya

Research and development of a method for accelerating the operation of joining distributed datasets according to a given criterion

Dissertation summary

for the purpose of obtaining academic degree

Doctor of Philosophy in Technology

Academic supervisor:

Doctor of Technical Sciences, Professor
Vladimir Saenko

Doctor of Technical Sciences, Professor
Sergey Tumkovskiy

Moscow – 2024

Relevance of the research

Over the past two decades, the need for data storage and analysis has increased. The volume of information in the modern world is growing rapidly and poses new challenges for analytical systems. Large analytical platforms process tens of terabytes of data daily; at the same time, the complexity of applications aimed at processing big data increases, and the need for the use of special computing algorithms becomes more acute.

For a long time, calculations were accelerated by increasing processor performance. This trend towards hardware expansion ended around 2005, when hardware developers faced severe limitations on increasing processor performance. Thus, to increase the speed of calculations, data processing began to be carried out in cluster systems consisting of many computers.

This research is aimed at developing a method for accelerating the operation of joining data according to a given criterion, which is one of the difficult analytical computing problems. Joining is an expensive operation that is difficult to scale and improve efficiency in distributed databases. This operation is used when it is necessary to obtain data from two or more tables based on logical relationships between them. The join operation specifies the criterion that is needed to indicate how data from one table intersects with data from another table.

This work examines the possibility of accelerating a special case of a join operation: a left-sided join, that is, the final selection includes all values from the left data set, to which, if the join criterion is met, records from the right set are added. If a record matching the left data set is not found in the right data set, a null value is added to the result.

This development is built on the basis of HDFS, which is a file system for storing large files. The algorithm developed based on the method proposed in the work was developed in the Apache Spark framework. The choice of this framework was due to a number of advantages that it has: it provides a convenient developer interface and high-level tools for creating custom extensions and utilities; distributed

under a free open source license; used by a large developer community; demonstrates high efficiency; confirmed by independent research from different companies.

In the direction of accelerating data operations in distributed file systems, in particular accelerating join operations, active research and development of new methods are underway. Methods are being explored to speed up the join operation for certain hardware, such as FPGAs (programmable logic integrated circuits) [1]. Research is underway to speed up data joins using certain types of conditions, such as theta-join, which joins rows from multiple tables using a join criterion that includes a comparison operator greater than or less than [2]. As well as accelerating the connection operation for a specific representation of source data, for example, RDF (Resource Description Framework), which is a model for representing and exchanging data on the Internet and provides a flexible, extensible way to describe resources, their properties and relationships between them in a machine-readable format [3].

Thus, there are several areas of active research related to data connectivity. Here are some examples:

• New join algorithms (hash join, sort-merge join, index-based join and adaptive join algorithms);

• Parallel and distributed processing;

• Connection optimization;

• Fusion techniques for big data;

• Connection in streaming data;

• Optimization of connections in cloud environments;

• Joining in graph databases.

The research community is continuously working to address these challenges and improve the performance, scalability, and efficiency of join operations in a variety of data processing scenarios.

Thus, the topic of the dissertation research, devoted to the development of a method for accelerating the operation of connecting distributed data arrays according to a given criterion, is undoubtedly relevant.

Purpose and objectives of the study

The object of study is the operation of joining distributed datasets according to a given criterion.

The subject of the study is the cost of computer time to perform a data join operation in distributed computing systems.

The purpose of the research and scientific task is to reduce the cost of computer time by developing and implementing a method for accelerating the operation of joining distributed datasets according to a given criterion.

To achieve this goal, the following objectives were formulated and set:

1. Conduct a critical review and analysis of literature sources that address the design of distributed data storage architectures and parallel computing algorithms;

2. Based on the analysis of literature data, identify the limiting stages of standard algorithms for joining datasets;

3. Develop a method for accelerating the operation of joining distributed datasets according to a given criterion;

4. Based on the developed method for the operation of joining distributed datasets, create an algorithm and utility for its implementation, which allows you to speed up the process of joining them;

5. Carry out experimental studies confirming the performance and effectiveness of the developed method;

6. Implement the proposed method into the VK work cycle.

Scientific novelty

1. Two previously not fully studied limiting stages of performing the operation of joining distributed datasets according to a given criterion have been identified.

This is the sorting stage and shuffling stage, which is the process of moving data across the computing cluster;

2. A method has been developed to speed up the operation of joining datasets, which differs from the known ones in that the data of one of the joined datasets is not sorted or shuffled within the cluster;

3. To confirm the performance of the proposed method, a technique was created that differs from the known ones in that it uses the techniques of partitioning and partial transfer of datasets to the computing nodes of the cluster, and computer experiments were performed that showed the adequacy and effectiveness of the developed method for accelerating the operation of joining distributed datasets.

The practical significance of the work lies in the possibility of using the presented method for accelerating the operation of joining distributed datasets according to a given criterion for a wide range of tasks. Based on the proposed method, an algorithm and a utility for implementing this algorithm were created. It has been experimentally determined that the larger the volume of data arrays, the greater the connection acceleration compared to standard methods. It was shown that for 2Tb data the join operation was performed ~37% faster than the proposed algorithm in Spark SQL; for 7Tb data this speedup was already ~47%.

Research methods

In this dissertation, methods of object-oriented analysis and design, methodologies for modeling and designing software systems, and modular programming methods were used. To evaluate the results and reliability of the proposed method and the developed algorithm, computer experiments were carried out.

Provisions for defense

1. A method for accelerating the operation of joining datasets, which differs from the known ones in that the data of one of the joined datasets is not sorted and does not move within the cluster;

2. An algorithm and a user library developed based on the proposed method that expands the functionality of the Apache Spark software product;

3. Results of experimental studies of the operation of joining distributed datasets using standard methods of the Spark SQL library and the developed algorithm on data volumes from 100Mb to 7Tb.

Compliance with the specialty passport

The purpose of the research and scientific task is to reduce the computer time costs by developing and implementing a method for accelerating the operation of joining distributed datasets according to a given criterion and contributes to the following areas of research in the specialty “Mathematical and software of computers, complexes and computer networks”, listed in passport of this specialty of the National Research University Higher School of Economics "Informatics, Computer Science and Management":

1. database and knowledge management systems;

2. models and methods for creating programs and software systems for parallel and distributed data processing, languages and tools for parallel programming;

3. models, methods, algorithms and software infrastructure for organizing globally distributed data processing.

Approbation of work

1. Y. Tyryshkina. Understanding join strategies in distributed systems. International Seminar on Electron Devices Design and Production (SED-2021), 2021, Czech Republic, pp 1-4. doi:10.1109/SED51197.2021.9444489.

2. Y. Tyryshkina, Accelerating join of distributed datasets by a given criterion. In proceedings of 2022 IEEE Moscow Workshop on Electronic and Networking Technologies (MWENT), 2022, Russian Federation, pp 1-3. doi:10.1109/MWENT55238.2022.9802185.

3. Yevgeniya Tyryshkina, Sergey Tumkovskiy. Method for accelerating the joining of distributed datasets by a given criterion. Information and Control Systems, 2022, Russian Federation.

Personal contribution of the author to the development of the problem

The author's personal contribution consists in the formulation of research problems and their solution, preparation, implementation, calculation and analysis of experimental and theoretical data, modification and use of program code for carrying out all kinds of calculations, preparation and visualization of graphic material, preparation of the text of articles and presentation of research results in Russian and international publications and conferences. An act on the implementation of the dissertation research results into the data processing process at the VK company was received (see Appendix 1).

The reliability of the results obtained in the dissertation research is confirmed by:

• tests carried out in accordance with generally accepted standards;

• compliance of the results obtained with benchmarks of independent studies;

• compliance of the results of computer experiments and calculations with the proposed empirical models.

Degree of development of the problem

The join operation has a significant impact on analytical computing performance because it uses a lot of memory, network, and disk resources. Spark SQL uses the Catalyst Query Optimizer module to solve the pressing problem of

speeding up join operations. However, the join operation methods implemented in this module have limiting stages that slow down the query execution time. This is the sorting stage and shuffling stage, which is the process of moving data across the computing cluster.

The proposed method for speeding up the operation of joining datasets differs from the known ones in that the data of one of the merged datasets is not sorted and does not move within the cluster. The developed method is aimed at eliminating the key shortcomings of existing approaches to performing the operation of connecting data arrays. The developed custom library can be used as an extension of the functionality of the Apache Spark framework and can be used for many analytical purposes.

List of published articles on the topic of the dissertation

1. Y. Tyryshkina. Understanding join strategies in distributed systems. International Seminar on Electron Devices Design and Production (SED-2021), 2021, Czech Republic, pp 1-4. doi:10.1109/SED51197.2021.9444489.

2. Y. Tyryshkina, Accelerating join of distributed datasets by a given criterion. In proceedings of 2022 IEEE Moscow Workshop on Electronic and Networking Technologies (MWENT), 2022, Russian Federation, pp 1-3. doi:10.1109/MWENT55238.2022.9802185.

3. Yevgeniya Tyryshkina, Sergey Tumkovskiy. Method for accelerating the joining of distributed datasets by a given criterion. Information and Control Systems, 2022, Russian Federation.

Conclusion

The paper shows that the development of fast and reliable methods for processing big data in distributed systems is an urgent task, and the existing methods are not effective enough. To improve the efficiency of existing work methods:

1. The stages of the operation of joining distributed datasets have been identified, which determine the speed of the operation, which includes sorting and shuffling data across the computing cluster;

2. A method has been developed to speed up the joining of datasets, due to the fact that the data of one of the joined datasets is not sorted or shuffled within the cluster, which gives a benefit, since they are the most time-consuming;

3. A methodology for experimental research of the developed method has been created, using the techniques of partitioning and partial transfer of datasets to the computing nodes of the cluster;

4. Computer experiments were performed that confirmed the effectiveness of the developed method for accelerating the operation of joining distributed datasets in volumes from 100Mb to 7Tb.

Further development of the results of the work performed involves their application to different types of data joins, the construction of optimal algorithms for distributing datasets into partitions in order to minimize their uneven division.

Appendix 1. Implementation Act.

ООО «ВК»

Адрес 125167, г. Москва, Ленинградский проспект, д. 39, строение 79

ОГРН 1027739850962
ИНН 7743001840
КПП 997750001
Тел. +7 (495)-725-63-57
+7 (495)-725-63-59

АКТ о внедрении

Утилита для реализации алгоритма ускорения операции соединения распределенных массивов данных по заданному критерию разработанная в рамках кандидатской работы аспиранткой Национального исследовательского университета «Высшая школа экономики» Тырышкиной Евгенией Сергеевной соответствуют всем требованиям и принята для опытной эксплуатации.

Руководитель направления разработки _____ Швеенков Д.Б

Директор по персоналу _____ Шокина Л.В.

ВК
отдел кадров

ООО «ВК»