

Федеральное государственное автономное образовательное учреждение
высшего образования
«Национальный исследовательский университет
«Высшая школа экономики»

на правах рукописи

Тырышкина Евгения Сергеевна

**Исследование и разработка метода ускорения операции соединения
распределенных массивов данных по заданному критерию**

РЕЗЮМЕ ДИССЕРТАЦИИ

на соискание учёной степени кандидата технических наук

Научный руководитель:

доктор технических наук, профессор
Саенко Владимир Степанович

доктор технических наук, профессор
Тумковский Сергей Ростиславович

Москва – 2024

Актуальность темы исследования

За последние два десятилетия потребность в хранении и анализе данных выросла. Объем информации в современном мире стремительно растет и ставит новые задачи перед аналитическими системами. Крупные аналитические платформы обрабатывают десятки терабайт данных ежедневно, вместе с этим повышается и сложность приложений, направленных на обработку больших данных, обостряется необходимость в применении специальных вычислительных алгоритмов.

Долгое время скорость вычислений увеличивалась за счет увеличения производительности процессоров. Эта тенденция к аппаратной экспансии закончилась примерно в 2005 году, когда разработчики аппаратного обеспечения столкнулись с жесткими ограничениями повышению производительности процессоров. Так, для повышения производительности обработка данных стала производиться в кластерных системах, состоящих из множества вычислительных машин.

Данное исследование направлено на разработку метода ускорения операции соединения (слияния) данных по заданному критерию, которая представляет собой одну из тяжелых аналитических вычислительных задач. Соединение — дорогостоящая операция, которую трудно масштабировать и повышать эффективность в распределенных базах данных. Данная операция применяется тогда, когда необходимо получить данные из двух или нескольких таблиц на основе логических связей между ними. В операции соединения задается критерий, который необходим, чтобы указать, как данные из одной таблицы пересекаются с данными из другой таблицы.

В данной работе рассматривается возможность ускорения частного случая операции соединения: левостороннего соединения, то есть в итоговую выборку попадают все значения из левого набора данных, к которым при соответствии критерию соединения добавляются записи из правого набора.

Если соответствующая левому набору данных запись не найдена в правом наборе данных, к результату добавляется нулевое значение.

Данная разработка построена на базе HDFS, что является файловой системой для хранения файлов больших размеров. Алгоритм, разработанный на основе предлагаемого в работе метода, был разработан во фреймворке Apache Spark. Выбор данного фреймворка был обусловлен рядом преимуществ, которыми он обладает: предоставляет удобный интерфейс разработчика и высокоуровневые инструменты для создания собственных расширений и утилит; распространяется под свободной лицензией с открытым исходным кодом; используется большим сообществом разработчиков; демонстрирует высокую эффективность; подтвержденную независимыми исследованиями разных компаний.

В направлении ускорения операций с данными в распределенных файловых системах, в частности ускорение операции соединения, ведутся активные исследования и разработка новых методов. Исследуются способы ускорения операции соединения для определенного аппаратного обеспечения, например, ПЛИС (программируемых логических интегральных схем)[1]. Ведутся исследования, направленные на ускорение соединения данных по определенным типам условий, таким как *theta-join*, соединяющим строки из нескольких таблиц по критерию соединения, включающему оператор сравнения больше или меньше [2]. А также ускорение операции соединения для специфичного представления исходных данных, например, RDF (Resource Description Framework), который представляет собой модель представления и обмена данными в Интернете и обеспечивает гибкий, расширяемый способ описания ресурсов, их свойств и отношений между ними в машинно-читаемом формате [3].

Таким образом, существует несколько областей активных исследований, связанных с соединением данных. Вот некоторые примеры:

- Новые алгоритмы соединения (хэш-соединение, соединение сортировкой-слиянием, соединение на основе индексов и адаптивные алгоритмы соединения);

- Параллельная и распределенная обработка;
- Оптимизация соединения;
- Техники объединения для больших данных;
- Соединение в потоковых данных;
- Оптимизация соединения в облачных средах;
- Соединение в графовых базах данных.

Исследовательское сообщество непрерывно работает над решением этих задач и улучшением производительности, масштабируемости и эффективности операций соединения в различных сценариях обработки данных.

Таким образом, тематика диссертационного исследования, посвященная разработке метода ускорения операции соединения распределенных массивов данных по заданному критерию, несомненно, является актуальной.

Цель и задачи исследования

Объект исследования – операция соединения распределенных массивов данных по заданному критерию.

Предмет исследования – затраты машинного времени на выполнение операции соединения данных в распределенных вычислительных системах.

Цель исследования и научная задача заключается в снижении затрат машинного времени за счет разработки и внедрения метода ускорения операции соединения распределенных массивов данных по заданному критерию.

Для достижения этой цели были сформулированы и поставлены **задачи:**

1. Провести критический обзор и анализ литературных источников, в которых рассматриваются вопросы проектирования архитектур распределенных хранилищ данных и алгоритмы параллельных вычислений;
2. На основе анализа литературных данных выявить лимитирующие стадии стандартных алгоритмов соединения массивов данных;
3. Разработать метод ускорения операции соединения распределенных массивов данных по заданному критерию;
4. На основе разработанного метода операции соединения распределенных массивов данных создать алгоритм и утилиту для его реализации, позволяющую ускорить процесс их соединения;
5. Выполнить экспериментальные исследования, подтверждающие работоспособность и эффективность разработанного метода;
6. Внедрить разработанный метод в рабочий цикл VK.

Научная новизна

1. Выявлены две ранее до конца не изученные лимитирующие стадии выполнения операции соединения распределенных массивов данных по заданному критерию. Это стадия сортировки и стадия перемещения данных по вычислительному кластеру;
2. Разработан метод ускорения операции соединения массивов данных отличающийся от известных тем, что данные одного из объединяемых датасетов не сортируются и не перемещаются внутри кластера;
3. Для подтверждения работоспособности разработанного метода создана методика, отличающаяся от известных тем, что использует приемы партиционирования и частичной передачи наборов на вычислительные узлы кластера, и выполнены компьютерные эксперименты, показавшие адекватность и эффективность разработанного метода ускорения операции соединения распределенных массивов данных.

Практическая значимость работы заключается в возможности применения разработанного метода ускорения операции соединения распределенных массивов данных по заданному критерию для широкого круга задач. На основе разработанного метода создан алгоритм и утилита для реализации этого алгоритма. Экспериментально определено, что чем больше объем массивов данных тем большее ускорение соединения относительно стандартных методов. Показано, что для данных размером 2Тб операция соединения была выполнена на ~ 37% быстрее, чем предложенный алгоритм в Spark SQL, для данных размером 7Тб это ускорение составило уже ~ 47%.

Методы исследования

В данной диссертационной работе применялись методы объектно-ориентированного анализа и проектирования, методологии моделирования и проектирования программных систем, методы модульного программирования. Для оценки результатов и достоверности предложенного метода и разработанного алгоритма были проведены компьютерные эксперименты.

Положения, выносимые на защиту

1. Метод ускорения операции соединения массивов данных, отличающийся от известных тем, что данные одного из объединяемых датасетов не сортируются и не перемещаются внутри кластера;
2. Разработанный на основе предложенного метода алгоритм и пользовательская библиотека, расширяющая функционал программного продукта Apache Spark;
3. Результаты экспериментальных исследований работы операции соединения распределенных массивов данных стандартными методами

библиотеки Spark SQL и разработанного алгоритма на объемах данных от 100Мб до 7Тб.

Соответствие паспорту специальности

Цель исследования и научная задача заключается в снижении затрат машинного времени за счет разработки и внедрения метода ускорения операции соединения распределенных массивов данных по заданному критерию и вносит вклад в следующие области исследований специальности «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей», перечисленные в паспорте этой специальности НИУ ВШЭ «Информатика, вычислительная техника и управление»:

1. системы управления базами данных и знаний;
2. модели и методы создания программ и программных систем для параллельной и распределенной обработки данных, языки и инструментальные средства параллельного программирования;
3. модели, методы, алгоритмы и программная инфраструктура для организации глобально распределенной обработки данных.

Апробация работы

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.

З. Е. С. Тырышкина, С. Р. Тумковский. Метод ускорения соединения распределенных наборов данных по заданному критерию. Информационно-управляющие системы, 2022, Российская Федерация.

Личный вклад автора в разработку проблемы

Личный вклад автора состоит в формулировке и постановке задач исследования и их решении, подготовке, реализации, расчёте и анализе экспериментальных и теоретических данных, модификации и использовании программного кода для проведения всевозможных расчётов, подготовки и визуализации графического материала, подготовки текста статей и представлении результатов исследований в российских и международных изданиях и конференциях. Получен акт о внедрении результатов диссертационного исследования в процесс обработки данных в компании «VK» (см. Приложение 1)

Достоверность полученных результатов в диссертационном исследовании подтверждаются:

- тестами, проведенными в соответствии с общепринятыми стандартами;
- соответствием полученных результатов бенчмаркам независимых исследований;
- соответствием результатов компьютерного эксперимента и расчетов предложенным эмпирическим моделям.

Степень разработанности проблемы

Операция соединения сильно влияет на производительность аналитических вычислений, так как использует много ресурсов памяти, сетевых ресурсов и дисковых ресурсов. В Spark SQL для решения актуальной проблемы ускорения операции соединения применяется модуль «Оптимизатор запросов Catalyst». Однако реализованные в данном модуле

методы выполнения операции соединения имеют лимитирующие стадии, замедляющие время выполнения запросов. Это стадия сортировки и стадия перемещения данных по вычислительному кластеру.

Предложенный метод ускорения операции соединения массивов данных отличается от известных тем, что данные одного из объединяемых датасетов не сортируются и не перемещаются внутри кластера. Разработанный метод направлен на устранение ключевых недостатков существующих подходов к выполнению операции соединения массивов данных. Разработанная пользовательская библиотека может использоваться в качестве расширения функционала фреймворка Apache Spark и применения во многих аналитических целях.

Список опубликованных статей по теме диссертации.

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. Е. С. Тырышкина, С. Р. Тумковский. Метод ускорения соединения распределенных наборов данных по заданному критерию. Информационно-управляющие системы, 2022, Российская Федерация.

Заключение

В работе показано, что разработка быстрых и надежных методов обработки больших данных в распределенных системах является актуальной

задачей, а имеющиеся методы недостаточно эффективны. Для повышения эффективности существующих методов в работе:

1. Выявлены стадии выполнения операции соединения распределенных массивов данных, определяющих скорость выполнения операции, к которым относится сортировка и перемещение данных по вычислительному кластеру;

2. Разработан метод ускорения соединения массивов данных, за счет того, что данные одного из объединяемых датасетов не сортируются и не перемещаются внутри кластера, что и дает выигрыш, поскольку именно они являются наиболее затратными по времени;

3. Создана методика экспериментального исследования разработанного метода, использующая приемы партиционирования и частичной передачи наборов на вычислительные узлы кластера;

4. Выполнены компьютерные эксперименты, подтвердившие эффективность разработанного метода ускорения операции соединения распределенных массивов данных на объемах от 100Мб до 7Тб.

Дальнейшее развитие результатов выполненной работы предполагает их применение на разные типы соединений данных, построению оптимальных алгоритмов распределения наборов данных по партициям с целью минимизации их неравномерного разделения.

Приложение 1. Акт о внедрении.



ООО «БК»

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 _____  Швенков Д.Б.

Директор по персоналу _____  Улохина И.В.



ООО «БК»