УДК 00.004.75

НАГРУЗОЧНОЕ ТЕСТИРОВАНИЕ КЛАСТЕРНОЙ СИСТЕМЫ НА ОСНОВЕ ПОИСКА ОРТОГОНАЛЬНЫХ ЛАТИНСКИХ КВАДРАТОВ

Н.С. Ганенко[1], К.В. Михайлова[2]

Самарский университет, г. Самара
[1] e-mail: nik.ganenko@mail.ru
[2] e-mail: kristinamihaylova19@gmail.com

Аннотация

В работе рассматривается подход к нагрузочному тестированию кластерных систем с использованием сложной вычислительной задачи. В качестве тестовой нагрузки применяется задача поиска ортогональных пар диагональных латинских квадратов. Эксперименты проводились с помощью облачной платформы Everest и файлового хранилища «Сергей Королёв». В статье показана зависимость ускорения от количества разбиения, проанализировано влияние разбиения данных на эффективность параллельных вычислений.

Ключевые слова: распределённые вычисления, нагрузочное тестирование, параллельные вычисления, кластерные системы.

Введение

В этой работе рассматривается подход к нагрузочному тестированию кластерной системы, который основан на использовании сложных комбинаторных задач. Для оценки работы был использован вычислительно-нагруженный тест – поиск ортогональных пар диагональных латинских квадратов. Целью работы является исследование методов нагрузочного тестирования кластерной системы с использованием интенсивной вычислительной нагрузки и параллелизма.

1. Описание и анализ предметной области

Латинский квадрат – это таблица размером n×n, где элементы матрицы заполняются элементами множества {0, …, n - 1} так, что элементы встречаются ровно один раз в каждой строке и столбце. Частным случаем этого квадрата является диагональный латинский квадрат, у которого каждый элемент не повторяется по горизонтали и вертикали и на главной и побочной диагоналях [1].

2. Организация вычислительного эксперимента

В работе использовалась облачная платформа для распределенных вычислений Everest [2]. Дополнительная документация по организации вычислений, окружениям, задачам и использованию вычислительных ресурсов представлена в проекте Templet [3].

В работе проводится сравнение между облачной платформой Everest и файловым хранилищем виртуальных машин «Сергей Королёв».

В экспериментах использовалась база из 171200 диагональных латинских квадратов, разбитая на 4-20 частей [4]. Также на платформе Everest необходимо

подготовить приложения для 2 видов тестирования и сформировать конфигурацию из 10 вычислительных узлов.

Для эксперимента берутся диагональные латинские квадраты порядка 7, так как корректность результатов проверяется сравнением с известным числом решений. Данная задача ресурсоёмкая, поэтому для проведения таких подсчетов необходима распределенная система. Для сравнения берется метод непосредственной передачи через систему управления распределенными вычислениями. Так как грид-система передает и формирует задачи автоматически, то сравнение проводится между использованием распределённой файлового хранилища и через грид-систему.

3. Результаты экспериментов

В ходе работы был проведён ряд вычислительных экспериментов, измерялись следующие характеристики:

1) время последовательного выполнения вычислений;

2) время параллельного выполнения вычислений;

3) ускорение вычислений.

Общее количество задач в эксперименте включает n задач первого типа и $\frac{n(n-1)}{2}$ задач второго типа. Тогда для расчета необходимого количества ресурсов, воспользуемся следующей формулой (1):

$$n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2} = N, \quad (1)$$

где n – количество файлов;

N – количество вычислительных ресурсов.

С помощью формулы (1) при фиксированном числе вычислительных ресурсов (10) определяем общее количество вычислительных задач. При этом увеличение количества задач влияет на степень загрузки системы и эффективность параллельного выполнения.

На графиках 1-2 представлены результаты сравнения последовательного и параллельного времени выполнения через Everest и файловое хранилище «Сергей Королёв».

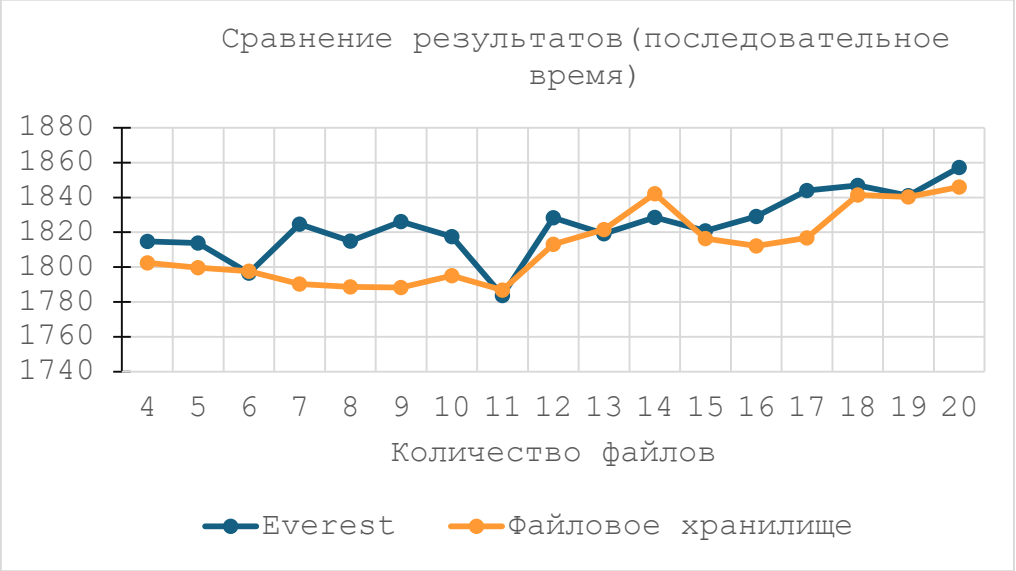

График 1 – График зависимости последовательного времени от разбиения

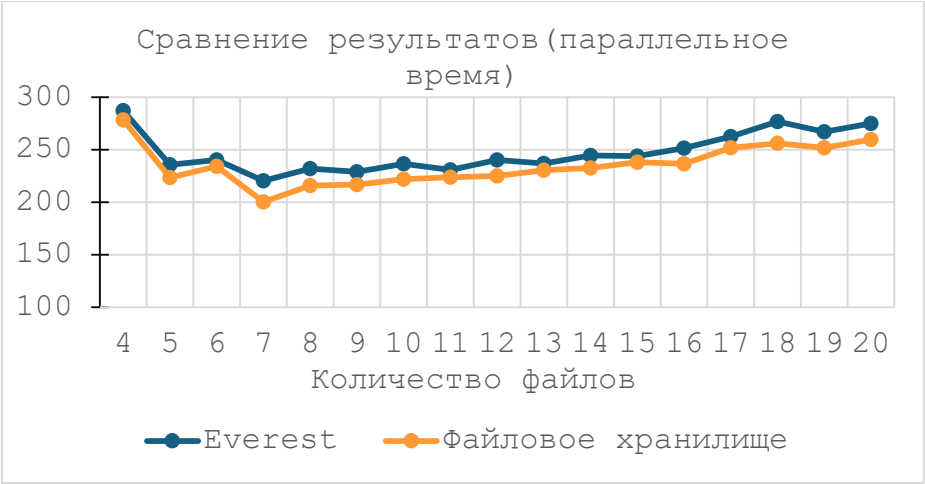

График 2 – График зависимости параллельного времени от разбиения

При малом количестве вычислительных задач наблюдается недостаточная загрузка вычислительных узлов. С увеличением числа разбиений общее количество задач возрастает квадратично, что способствует более равномерному распределению нагрузки и повышает эффективность параллельных вычислений. При этом, если число вычислительных задач слишком велико, то возрастают накладные расходы, связанные с передачей данных, и уменьшается эффективность вычислений.

Экспериментальными результатами было установлено, что максимальные показатели достигаются при n ≈ 7.

Ускорение определялось по формуле (2) как отношение времени последовательного выполнения к времени параллельного выполнения:

$$a = \frac{T_s}{T_p}, \tag{2}$$

где Ts – время последовательного выполнения;

Tp – время параллельного выполнения.

На графике 3 представлена зависимость ускорения от разбиения (количество файлов) двух экспериментов.

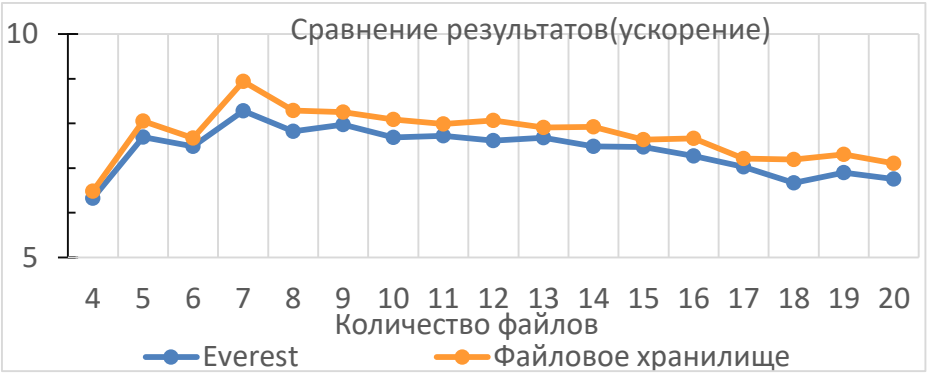

График 3 – График зависимости ускорения от разбиения (сравнение результатов)

По графику видно, что при малом числе разбиений количество задач недостаточно для эффективного использования всех ресурсов системы, при этом происходит простой вычислительных узлов. Увеличение числа файлов приводит к квадратичному росту количества задач. По графику наблюдается нелинейная зависимость, характер которой определяется увеличением степени параллелизма и ростом накладных расходов в распределённой системе.

Наблюдаемое поведение ускорения объясняется увеличением степени параллелизма при росте числа задач и последующим увеличением накладных расходов при дальнейшем увеличении.

Заключение

По результатам экспериментов установлено, что зависимость ускорения от числа разбиений имеет нелинейных характер и определяется степенью параллелизма и накладными расходами. Было выявлено, что максимальное ускорение достигается при $n \approx 7$. При этом числе происходит наиболее эффективное использование вычислительных ресурсов при заданной конфигурации. Сравнение платформ показало, что использование файлового хранилища «Сергей Королёв» обеспечивает большее ускорение, чем платформа Everest. Использование этой платформы целесообразно при автоматизации распределённых вычислений и удобного управления ресурсами.

Список использованных источников

1. Использование графических процессоров для поиска пар ортогональных диагональных латинских квадратов порядка 10. – URL: https://2016.russianscdays.org/files/pdf16/180.pdf (дата обращения: 20.03.2026).

2. Sukhoroslov, O. Supporting Efficient Execution of Workflows on Everest Platform / O. Sukhoroslov // Russian Supercomputing Days. – Springer, Cham, 2019. – P. 713–724. (дата обращения: 20.03.2026).

3. Проект Templet. – URL: https://templet.ssau.ru/wiki/ (дата обращения: 20.03.2026).

4. Учет алгоритмических особенностей задачи при генерации диагональных латинских квадратов / Э.И. Ватутин, А.Д. Журавлев, О.С. Заикин, В.С. Титов // Известия ЮЗГУ. – 2016. – № 2 (65). – P. 46-59. (дата обращения: 28.03.2026).

LOAD TESTING OF A CLUSTER SYSTEM BASED ON ORTHOGONAL LATIN SQUARES SEARCH

N.S. Ganenko[1], K.V. Mihaqlova[2]

Samara University, Samara
[1] e-mail: nik.ganenko@mail.ru
[2] e-mail: kristinamihaylova19@gmail.com

Abstract

The paper discusses an approach to load testing of cluster systems using a complex computational task. The task of finding orthogonal pairs of diagonal Latin squares is used as a test load. The experiments were conducted using the Everest cloud platform and the Sergey Korolev file storage. The paper shows the dependence of the acceleration on the number of partitions and analyzes the impact of data partitioning on the efficiency of parallel computing.

Keywords: distributed computing, load testing, parallel computing, and cluster systems.