

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
Информатики и радиоэлектроники

УДК 004.75

Гацура
Вадим Дмитриевич

Распределенная система решения задач

АВТОРЕФЕРАТ

на соискание степени магистра информатики и вычислительной техники
по специальности 1-40 81 02 Технологии виртуализации и облачных
вычислений

Научный руководитель
Матвейчук Наталья Михайловна
кандидат физико-математических
наук, доцент

Минск 2020

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Объект и предмет исследования

Объектом исследования являются системы планирования решения задач в распределенных системах.

Предметом исследования выступают методы планирования задач, которые используются в распределенных системах.

Цель и задачи исследования

Целью исследования является разработка и реализация алгоритма планирования распределенных задач в условиях неопределенности длительностей, который обеспечит оптимальное или приближенное к оптимальному время выполнения задач.

Задачи исследования

Для достижения поставленной цели были сформулированы следующие задачи:

- Изучение существующих алгоритмов планирования для распределенного решения задач.
- Разработка математической модели планирования выполнения задач в распределенной системе.
- Реализация алгоритма планирования для распределенного решения задач.
- Проведение экспериментов по составлению расписания на основе предложенного метода для распределенного решения задач.

Связь с реальным сектором экономики

На основе разработанной системы планирования можно создать программные модули для популярных распределенных вычислительных систем, что обеспечит распространение системы среди широкого круга лиц, а также позволит распространять систему на соответствующих площадках.

Апробация диссертации

Результаты исследований по теме диссертации были представлены в виде доклада «Распределенная система решения задач» на 56-ой научной конференции аспирантов, магистрантов и студентов БГУИР в 2020 году.

Публикация результатов исследований

На основе полученных в ходе исследований результатов была опубликована статья «Two machine job-shop scheduling problem to minimize the makespan with uncertain job duration» в международном научном журнале «Algorithms» [1]. Также результаты исследований были опубликованы в виде тезисов доклада на 56-ой научной конференции аспирантов, магистрантов и студентов БГУИР [2].

ВВЕДЕНИЕ

В последнее время широкое распространение получили совершенно новые виды распределенных систем: кластерные системы, грид-системы и облачные системы. Это связано с увеличением числа решаемых прикладных задач и значительным возрастанием нагрузок на вычислительные системы. В связи с этим все более актуальными становятся вопросы оптимизации распределенных вычислений на основе рационального планирования работы как отдельного узла, так и целой распределенной сети.

Значительное число исследований показало, что современные вычислительные нагрузки высоко изменчивы, и характеризуются особым распределением: множество небольших задач (или запросов) с малыми ресурсными требованиями и сравнительно небольшое множество крупных задач с непропорционально высокими требованиями к потребляемым ресурсам. Интеллектуальное планирование распределения задач по узлам обработки, обеспечивающее эффективное использование ресурсов становится чрезвычайно актуальным при таких характеристиках задач.

На сегодняшний день существует множество решений по составлению расписаний для решения задач в распределенной сети, но все они не учитывают множество факторов, а именно неопределенность фактической длительности решения подзадач. Учитывая этот фактор при составлении расписаний решения подзадач в распределенной системе, можно добиться построения оптимального или приближенного к оптимальному расписания. Необходимость учета ограничений на используемые ресурсы при планировании распределенной задачи приводит к математическим задачам, рассматриваемым в теории расписаний.

Актуальность данной работы заключается в отсутствии алгоритма составления расписаний решения задач, который бы учитывал неопределенность фактической длительности решения подзадач, что позволяет обеспечить максимальное использование ресурсов.

Методы планирования задач, которые используются в распределенных системах являются предметом данной диссертации, цель которой заключается в разработке алгоритма планирования задач в условиях неопределенности длительностей.

Компоненты данной системы, как и система в целом могут быть интегрированы в планировщики распределенных систем.

СОДЕРЖАНИЕ РАБОТЫ

Общий объем магистерской диссертации составляет 73 страниц, включая 15 иллюстраций, 7 таблиц, библиографический список из 31 наименований, 1 приложение.

Во **введении** описывается текущее состояние области исследования, дается обоснование актуальности работы, описываются прикладные задачи, в которых может быть использована разработанная система. Также в разделе приводится краткий обзор проблематики задачи и современного состояния исследований по улучшению алгоритмов планирования в распределенных системах.

В **общей характеристике работы** сформулированы цель и задачи исследования, даны сведения об объекте и предмете исследования, приведены апробации и публикации результатов.

В **первой главе** произведен анализ литературы по данному направлению. В первом разделе дана характеристика распределенных вычислительных систем и приведена краткая историческая справка по их развитию. Во втором разделе приведена классификация распределенных систем с их кратким описанием, а также рассмотрена архитектура каждого из видов. Помимо этого, во втором разделе рассмотрены отличительные черты каждой из систем. В третьем разделе рассмотрены планировщики заданий, дисциплины обслуживания, их преимущества и недостатки. Также в данном разделе рассматриваются различного рода подходы к назначению узлов или ресурсов на задачи. В четвертом разделе приведены описания основных методов планирования, стратегии выбора ресурсов, их описание и отличительные черты. Также в данном разделе проводится оценка рассмотренных стратегий и обосновывается выбор стратегии, используемой в работе.

Во **второй главе** приводится описание теоретической базы исследования. В первом разделе рассматривается теория расписаний и её применимость к данному исследованию. Во втором разделе описаны обслуживающие системы flow-shop и job-shop. В третьем разделе приведено описание алгоритма Джонсона. В четвертом разделе описана теория по поиску оптимального по быстродействию расписания для задачи обслуживания двумя узлами множества требований с различными маршрутами обслуживания требований и неопределенными длительностями операций. В данном разделе ставится задача в формальном виде и приводятся основные определения. В пятом разделе описывается построение оптимального расписания для поставленной задачи из предыдущего раздела.

Третья глава посвящена проведению экспериментов разработанной системы планирования задач. В первом разделе представлено описание

основных реализованных алгоритмов. Во втором разделе методика проведения экспериментов для этапа *offline*, описание и анализ полученных результатов. В третьем разделе представлена методика проведения экспериментов для этапа *online*, а также описание и анализ полученных результатов.

В **заключении** приводится краткий обзор результатов, полученных на каждом из этапов исследования, приводится обоснование выбранных методов, дается критический анализ разработанной системы. Также в заключении указан список рекомендаций по дальнейшему продолжению исследования: на каких моментах стоит сфокусироваться, что можно улучшить в реализации алгоритма, – обозначена область применения полученных результатов.

Библиотека БГУИР

ЗАКЛЮЧЕНИЕ

В ходе исследования был проведен анализ существующих видов планирования, хронологический обзор основных этапов в развитии распределенных систем, а также их систем планирования, рассмотрены основные дисциплины обслуживания и стратегии планирования. Результаты исследования были систематизированы и на основании этого было принято решение в разрабатываемой системе совместить статическую и динамическую стратегию планирования для минимизации недостатков каждой из стратегий, что также позволило объединить их преимущества. Также было принято во внимание гетерогенная природа вычислительных узлов в распределенной системе, что также получило отражение в данной работе.

Для более полного понимания объекта исследования был произведен анализ обслуживающих систем flow-shop и job-shop, а также изучен алгоритм Джонсона. Полученные знания получили отражение в выбранной модели планирования. Для построения расписания был использован алгоритм поиска оптимального по быстродействию расписания для задачи обслуживания двумя узлами множества требований с различными маршрутами обслуживания и неопределенными длительностями операций.

При выборе инструментов для создания системы было принято решение использовать язык программирования C# как инструмент, предоставляющий оптимальный баланс между возможностями и уровнем контроля, а также получивший широкую популярность в последнее время, что также позволит провести интеграцию разработанной системы в популярные распределенные системы.

Использованная комбинация алгоритмов и инструментов позволила создать эффективную систему планирования. Разработанная система имеет довольно высокий процент построенных оптимальных расписаний для задач общего назначения ввиду неопределенностей длительностей.

Данная система будет использована в качестве основы для будущих исследований. В дальнейших работах планируется провести ряд экспериментов на более обширном классе задач, а также реализовать модуль для интеграции с популярными распределенными системами.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1. *Yuri N. Sotskov* Two machine job-shop scheduling problem to minimize the makespan with uncertain job duration / Yuri N. Sotskov, Natalja M. Matsveichuk, Vadzim D. Hatura // *Algorithms* 2020, 13, 4, pp. 95-139; doi:10.3390/a13010004.
2. *Гацура В.Д.* Распределенная система решения задач / В.Д. Гацура // Компьютерные системы и сети: 56-я научная конференция аспирантов, магистрантов и студентов, Минск, 21-24 апреля 2020 г. – Минск: Белорусский государственный университет информатики и радиоэлектроники, 2020. – С. 15-17.

Библиотека БГУИР