

УТВЕРЖДАЮ

Проректор ФГБОУ ВО
«МГУ имени М.В. Ломоносова»,
доктор физ.-мат. наук,
профессор



А.А. Федянин
А.А. Федянин

19 мая 2022 г.

ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный университет имени М.В. Ломоносова» на диссертационную работу АЛААСАМА Амира Басима Абдуламира «Модели, методы и алгоритмы обработки потоков данных в туманных вычислительных средах», представленную на соискание ученой степени кандидата физико-математических наук по специальности 05.13.11 – математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей.

Актуальность темы исследования. В настоящее время в различных предметных областях имеет место устойчивый рост устройств и систем интернета вещей и технологий цифровых двойников. В подобных приложениях требуется сбор, передача и анализа потоков данных от систем цифровых двойников в режимах, близких к реальному времени. Однако применение облачных технологий для настройки и актуализации виртуального состояния цифровых двойников не позволяет обеспечить требуемые характеристики предоставляемых вычислительных ресурсов как по времени задержки, так и по местоположению сервисов обработки данных. Возможным решением этой проблемы может служить применение модели туманных вычислений, которая расширяет концепцию облачных вычислений, предоставляя вычислительные ресурсы ближе к источникам данных. В силу указанных обстоятельств рассматриваемая в диссертации тема разработки моделей, методов и алгоритмов обработки потоков данных в туманных вычислительных средах, несомненно, является актуальной.

Целью исследования, проведенного А.Б.А. Алаасамом, была разработка новой концепции организации потоков работ, включая математическую модель, методы и алгоритмы, позволяющей организовать эффективную обработку потоков данных в туманных вычислительных средах. Для достижения этой цели соискатель формулирует и решает следующие **задачи**:

- 1) анализ известных концепций и принципов обработки потоковых данных, используемых для реализации приложений в туманных вычислительных средах;
- 2) разработка новой математической модели организации потоков работ, ориентированной на эффективную обработку потоков данных в туманных вычислительных средах;
- 3) разработка алгоритма преобразования монолитных приложений потоков работ в наборы независимых потоков работ, поддерживающих поточную обработку данных;
- 4) разработка комплекса программных компонентов и утилит для поддержки обработки потоков данных в туманных вычислительных средах посредством потоков работ;
- 5) проведение вычислительных экспериментов для оценки эффективности предложенной концепции и разработанного программного обеспечения.

Все поставленные задачи были успешно решены и соискателем получены следующие основные **научные результаты**.

1. Разработана новая концепция микро-потоков работ, ориентированная на организацию обработки данных в туманных вычислительных средах, позволяющая значительно уменьшить время задержки получения результата при обработке потоков данных.
2. Разработан алгоритм рефакторинга потоков работ, позволяющий преобразовать монолитный поток в набор микро-потоков, допускающих параллельное выполнение.
3. Выполнены проектирование и реализация комплекса вычислительных акторов и программных утилит, обеспечивающих функционирование

микро-потоков работ на базе платформы управления потоками работ Kerler и платформы обработки потоков данных Apache Kafka.

4. С помощью разработанного комплекса созданы микро-потоки работ, обеспечивающие поддержку типовых задач обработки данных, на базе которых проведены вычислительные эксперименты, подтверждающие эффективность предложенных подходов.

Научная новизна результатов заключается в следующем. Разработана новая концепция организации потоков работ (названная соискателем «концепция микро-потоков работ») которая включает в себя модель, методы и алгоритмы, обеспечивающие эффективную обработку потоков данных в туманных вычислительных средах, и позволяет существенно (на порядки) уменьшить время задержки получения результата.

Теоретическая значимость исследования, проведенного соискателем, состоит в том, что разработанная концепция микро-потоков работ включает в себя формальную модель организации обработки данных и алгоритм организации рефакторинга монолитных потоков работ в наборы слабосвязанных микро-потоков работ.

Практическая ценность диссертационного исследования заключается в реализации набора программных акторов и набора утилит, обеспечивающих интеграцию системы управления потоками работ Kerler и платформы обработки потоков данных Apache Kafka для реализации обработки потоков данных в виде микро-потоков.

Обоснованность и достоверность результатов диссертационного исследования подтверждается данными вычислительных экспериментов, проведенных в строгом соответствии с общепринятыми стандартами.

Соответствие специальности. Основные задачи, решенные в диссертации, и вынесенные на защиту результаты включают в себя разработку новой концепции организации потоков работ, в т.ч. математическую модель, методы и алгоритмы, позволяющие организовать эффективную обработку потоков данных в туманных вычислительных средах, что соответствует специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей.

Содержание текста диссертации

Диссертация состоит из введения, четырех глав, заключения, библиографии и двух приложений. В приложения вынесены основные аббревиатуры и обозначения, используемые в диссертации. Объем диссертации составляет 147 страниц, объем библиографии – 150 наименований.

Во введении приводится обоснование актуальности темы и степень ее разработанности; формулируются цели и задачи исследования; раскрываются новизна, теоретическая и практическая значимость полученных результатов; формулируется методологическая основа диссертационного исследования; дается обзор содержания диссертации.

В первой главе рассматриваются понятия модели туманных вычислений и обработки потоков данных в туманных вычислительных средах в контексте систем индустриального интернета вещей и цифровых двойников. Дан обзор методов декомпозиции потоков работ и подходов к проектированию программных систем, ориентированных на обработку потоков данных в туманных средах.

Во второй главе представлена математическая модель микро-потоков работ для обработки потоков данных в распределенных вычислительных средах, таких как туманные вычислительные системы. Модель включает в себя алгоритм рефакторинга зависимостей в монолитном потоке в набор автономных микро-потоков работ. Данное разделение поддерживает независимость реализации, исполнения, разработки, сопровождения и кроссплатформенного развертывания микро-потоков работ на независимых вычислительных узлах.

В третьей главе представлено описание программных компонентов, реализованных для поддержки реализации и тестирования модели микро-потоков работ.

В четвертой главе представлены результаты тестирования модели микро-потоков работ.

В заключении в краткой форме излагаются итоги выполненного диссертационного исследования, представляются отличия диссертационной работы от ранее выполненных родственных работ других авторов, даются ре-

комендации по использованию полученных результатов и рассматриваются перспективы дальнейшего развития темы.

Публикации и апробация результатов. По теме диссертации опубликовано 11 работ. Из них 5 статей опубликовано в журналах, индексируемых в Scopus и Web of Science, и 6 статей – в журналах, включенных ВАК в перечень изданий, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени доктора и кандидата наук. По теме исследования соискателем зарегистрирована одна программа для ЭВМ. Публикации соискателя с достаточной полнотой отражают основные результаты диссертации. Результаты исследования прошли апробацию на четырех международных научных конференциях.

Автореферат диссертации достаточно полно и корректно отражает ее содержание.

Рекомендации по использованию результатов диссертации

Результаты диссертации могут быть использованы в разработке программных комплексов, обеспечивающих эффективную распределенную обработку больших объемов данных, а также для создания учебных курсов для бакалавров и магистров по тематике распределенной обработки данных.

Замечания по диссертации

По диссертационной работе имеются следующие замечания, которые, тем не менее, не снижают значимости полученных результатов и оценки общего высокого научного уровня диссертационной работы А.Б.А. Алаасама.

1. Предложенный соискателем алгоритм рефакторинга монолитных потоков работ использует представление направленного ациклического графа потока работ в виде матрицы смежности, однако в тексте диссертации отсутствует обоснование выбора такого представления графа.
2. В разделе 2.2. используется нестандартная терминология теории графов: используются термины «выходные» и «входные» ребра вместо «исходящие» и «входящие» соответственно.
3. Соискатель применил систему Apache Kafka для реализации очереди передачи сообщений между распределенными микро-потоками, однако из текста диссертации неясно, может ли для этих целей быть использована

иная система, реализующая очередь сообщений.

4. В тексте диссертации имеются опечатки и не унифицирован стиль представления терминов на английском языке. Примеры опечаток: на стр. 109 вместо словосочетания «монолитный поток данных» необходимо «монолитный поток работ»; на рис. 42 используется написание «микросток», тогда как в остальном тексте используется написание через дефиса. При вводе новых терминов на английском языке для отделения термин и его аббревиатура разделяются то запятой, то тире: например, «Internet of things, IoT», «Digital Twin, DT» на стр. 4, 22 и «Event-Driven Architecture – EDA», «Scientific Workflow – SWF» на стр. 7, 27.

Заключение

Диссертационная работа А.Б.А. Алаасама представляет собой самостоятельно выполненную и завершенную научно-квалификационную работу, в которой автором разработана новая концепция организации потоков работ в туманных вычислительных средах, имеющая существенное значения для технологий распределенной обработки больших массивов данных. Диссертационная работа в полной мере соответствует требованиям Положения о порядке присуждения ученых степеней, включая п. 9, а ее автор, Амир Басим Абдуламир Алаасам, достоин присуждения ему ученой степени кандидата физико-математических наук по специальности 05.13.11 Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей.


Диссертация и отзыв обсуждены и одобрены на заседании научного семинара Научно-исследовательского вычислительного центра МГУ имени М.В. Ломоносова, протокол №2 от 05.04.2022.

Директор Научно-исследовательского
вычислительного центра МГУ
имени М.В. Ломоносова,
член-корреспондент РАН,
доктор физ.-мат. наук, профессор



Вл.В. Воеводин

УЧЕНЫЙ СЕКРЕТАРЬ
НИИЦ МГУ ИМ. МВ ЛОМОНОСОВА

 В.В. СУВОРОВ
«__» _____ 20__ г.

СВЕДЕНИЯ О ЛИЦАХ, УТВЕРДИВШИХ И ПОДГОТОВИВШИХ ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

на диссертацию Алаасама А.Б.А. «Модели, методы и алгоритмы обработки потоков данных в туманных вычислительных средах» на соискание ученой степени кандидата физико-математических наук по специальности 05.13.11 – математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей

Наименование организации	Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный университет имени М.В. Ломоносова»	
Ведомственная принадлежность	Правительство Российской Федерации	
Почтовый адрес, телефон, адрес электронной почты, адрес официального сайта в сети «Интернет»	119991, Москва, Ленинские горы, д. 1 ФГБОУ ВО МГУ им. М.В. Ломоносова +7 (495) 939-10-00 info@rector.msu.ru https://www.msu.ru/	
Сведения о лице, утвердившем отзыв	ФИО	Федянин Андрей Анатольевич
	Ученая степень (с указанием шифра специальности)	д.ф.-м.н., 01.04.04
	Должность	Проректор ФГБОУ ВО «МГУ имени М.В. Ломоносова»
Сведения о лице, подготовившем отзыв	ФИО	Воеводин Владимир Валентинович
	Ученая степень (с указанием шифра специальности)	д.ф.-м.н., 05.13.11
	Должность	Директор научно-исследовательского вычислительного центра МГУ имени М.В. Ломоносова

**Список основных работ сотрудников ведущей организации
по теме диссертации в рецензируемых научных изданиях**

1. Suplatov, D., Shegay, M., Sharapova, Y., Timokhin, I., Popova, N., Voevodin, V., Švedas, V. Co-designing HPC-systems by computing capabilities and management flexibility to accommodate bioinformatic workflows at different complexity levels // *Journal of Supercomputing*, 77 (11), pp. 12382-12398. 2021. DOI: 10.1007/s11227-021-03691-x.
2. Afanasyev, I.V., Lichmanov, D.I., Rudyak, V.Y., Voevodin, V.V. Efficient Implementation of Liquid Crystal Simulation Software on Modern HPC Platforms // *Supercomputing Frontiers and Innovations*, 8 (3), pp. 104-125. 2021. DOI: 10.14529/jsfi210306.
3. Antonov, A.S., Volkov, N.I. Information Graph Visualization Using Algo-View Software Tool // *Lobachevskii Journal of Mathematics*, 41 (8), pp. 1427-1434. 2020 г. DOI: 10.1134/S199508022008003X
4. Nikitenko, D., Antonov, A., Zheltkov, A., Voevodin, V. Describing HPC System Architecture for Understanding Its Capabilities // *Communications in Computer and Information Science*, 1331, pp. 425-435. 2020 г. DOI: 10.1007/978-3-030-64616-5_37
5. Shegay, M.V., Suplatov, D.A., Popova, N.N., Švedas, V.K., Voevodin, V.V. ParMATT: Parallel multiple alignment of protein 3D-structures with translations and twists for distributed-memory systems // *Bioinformatics*, 35 (21), pp. 4456-4458. 2019. DOI: 10.1093/bioinformatics/btz224
6. Suplatov, D., Sharapova, Y., Shegay, M., Popova, N., Fesko, K., Voevodin, V., Švedas, V. High-Performance Hybrid Computing for Bioinformatic Analysis of Protein Super-families // *Communications in Computer and Information Science*, 1129 CCIS, pp. 249-264. 2019. DOI: 10.1007/978-3-030-36592-9_21
7. Nikitenko, D., Zhumatiy, S., Paokin, A., Voevodin, V., Voevodin, V. Evolution of the Octoshell HPC Center Management System // *Communications in Computer and Information Science*, 1063, pp. 19-33. 2019. DOI: 10.1007/978-3-030-28163-2_2
8. Valkov, P., Kazmina, K., Antonov, A. Using Empirical Data for Scalability Analysis of Parallel Applications // *Communications in Computer and Information Science*, 1063, pp. 58-73. 2019. DOI: 10.1007/978-3-030-28163-2_5
9. Nasonov, D., Butakov, N., Melnik, M., Visheratin, A., Linev, A., Shvets, P., Sobolev, S., Mukhina, K. The multi-level adaptive approach for efficient execution of multi-scale distributed applications with dynamic workload // *Communications in Computer and Information Science*, 965, pp. 675-686. 2019. DOI: 10.1007/978-3-030-05807-4_58

10. Suplatov, D., Popova, N., Zhumatiy, S., Voevodin, V., Švedas, V. Parallel workflow manager for non-parallel bioinformatic applications to solve large-scale biological problems on a supercomputer // Journal of Bioinformatics and Computational Biology, 14 (2), Article1641008, 2016. DOI: 10.1142/S0219720016410080

«19» мая 2022 г.

Данные верны:

Проректор

МГУ имени М.В. Ломоносова



Федянин А.А.

