

## **ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА**

**кандидата технических наук Сухорослова Олега Викторовича 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.* Сегодняшнее информационное общество переживает взрывной рост устройств и систем интернета вещей и связанных с ними технологий цифровых двойников. В силу этого существует необходимость сбора, передачи и анализа потоков данных от систем цифровых двойников в режиме реального времени. Применение облачных технологий для настройки и актуализации виртуального состояния цифровых двойников не позволяет обеспечить требуемые характеристики предоставляемых вычислительных ресурсов с точки зрения времени задержки и местоположения служб обработки данных. Поэтому разработка моделей, методов и алгоритмов обработки потоков данных в туманных вычислительных средах является *актуальной задачей*.

*Целью исследования* была разработка новой концепции организации потоков работ, в т.ч. математической модели, методов и алгоритмов, которые обеспечивают эффективную обработку потоков данных в туманных вычислительных средах. Для достижения указанной цели необходимо было решить следующие *задачи*:

- 1) проанализировать известные концепции обработки потоковых данных, используемые для реализации приложений в туманных вычислительных средах;
- 2) разработать новую математическую модель организации потоков работ, ориентированную на эффективную обработку потоков данных в туманных вычислительных средах;
- 3) разработать алгоритм трансформации монолитных приложений потоков работ в наборы независимых потоков работ;

- 4) разработать комплекс программных компонентов и утилит для поддержки обработки потоков данных в туманных вычислительных средах;
- 5) провести вычислительные эксперименты для оценки эффективности предложенной концепции и разработанного программного обеспечения.

*Основные результаты исследования* заключаются в следующем.

1. Разработана новая концепция микро-потоков работ, ориентированная на организацию обработки данных в туманных вычислительных средах, позволяющая уменьшить время 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 Kepler и платформы обработки потоков данных Apache Kafka.
4. С помощью разработанного комплекса созданы примеры микро-потоков работ для решения типовых задач обработки данных, на базе которых проведены вычислительные эксперименты, подтверждающие эффективность предложенных подходов.

*Соответствие специальности.* Диссертационное исследование и его результаты соответствуют специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей.

*Содержание текста диссертации.* Диссертация состоит из введения, четырех глав, заключения, списка литературы и двух приложений, в которые вынесены основные аббревиатуры и обозначения диссертации соответственно. Объем диссертации – 147 страниц, объем библиографии – 150 наименований.

*Во введении* обоснованы актуальность темы и степень ее разработанности, сформулированы цель и задачи исследования, раскрыты

новизна, теоретическая и практическая значимость результатов, приведено краткое содержание диссертации. *В первой главе* рассмотрен теоретический базис обработки потоков данных в туманных вычислительных средах применительно к системам интернета вещей, цифровой индустрии и цифровым двойникам, а также приведен обзор работ по тематике исследования. *Во второй главе* описана математическая модель микро-потоков работ для обработки потоков данных в туманных вычислительных средах. В модель входит алгоритм преобразования зависимостей в монолитном потоке в набор автономных микро-потоков работ. *В третьей главе* описана разработанная соискателем программная реализация модели микро-потоков работ. *В четвертой главе* представлены результаты вычислительных экспериментов, всесторонне исследующих эффективность предложенной модели микро-потоков работ. В *заключении* изложены итоги исследования, показаны отличия данной диссертационной работы от работ других авторов, даны рекомендации по использованию полученных результатов и перспективы развития темы исследований.

*Автореферат диссертации* достаточно полно и корректно отражает содержание текста диссертации.

*Обоснованность и достоверность результатов диссертации* подтверждается данными вычислительных экспериментов, которые проведены в строгом соответствии с общепринятыми стандартами.

*Научная новизна результатов диссертации* заключается в следующем. Разработана новая концепция организации потоков работ (получившая название «концепции микро-потоков работ»), которая включает в себя модель, методы и алгоритмы, позволяющие эффективно обрабатывать потоки данных в туманных вычислительных средах, сокращая при этом время получения результата.

*Теоретическая значимость исследования* заключается в том, что предложенная соискателем концепция микро-потоков работ включает в себя формальную модель организации обработки данных и алгоритм трансформации монолитного потока работ в набор слабосвязанных микро-потоков работ. *Практическая ценность исследования* заключается в

реализации соискателем программной поддержки интеграции системы управления потоками работ Kerler и платформы обработки потоков данных Apache Kafka для обработки данных в виде микро-потоков.

*Публикации и апробация результатов.* По теме исследования соискатель опубликовал 5 статей в журналах индексов Scopus и Web of Science, а также 6 статей – в журналах из Перечня ВАК, и зарегистрировал одну программу для ЭВМ. В публикациях полно отражены основные результаты исследования. По результатам диссертации соискатель сделал четыре доклада на международных научных конференциях.

Имеются следующие *замечания по диссертации*, не снижающие ее высокий научный уровень и значимость результатов, полученных соискателем.

1. Не аргументировано утверждение из раздела 1.3.7: «Однако системы управления потоками данных обычно не накладывают ограничений на организацию процесса обработки и преобразования, реализуемых компонентами обработки данных. Это может привести к сложностям при проектировании и модернизации сложных конфигураций систем обработки данных». Например, можно было привести иллюстрирующий данные сложности пример. В целом, в работе приведено мало примеров реальных приложений, демонстрирующих преимущества предложенной концепции микро-потоков работ.
2. В разделе 1.4 не упоминается родственный предлагаемому подходу класс слабосвязанных научных потоков работ — т. н. композитных приложений, изначально создаваемых из независимых вычислительных пакетов или сервисов. В результате может сложиться представление, что все научные потоки работ являются монолитными и сильносвязанными.
3. За рамками работы остались важные с точки зрения практического применения предложенного подхода задачи о разбиении исходного потока работ на подпотоки (алгоритм из раздела 2.3 предполагает, что разбиение уже задано) и распределении соответствующих подпотокам вычислительных задач по узлам туманной вычислительной системы

(включая распараллеливание по данным) для организации максимально эффективной обработки данных при заданных ограничениях.

4. Не обоснован выбор системы управления потоками работ Kerler для реализации предложенного подхода. Указано только, что данная система является одной из наиболее популярных. Судя по официальному сайту, система давно не развивается, последняя версия была опубликована в 2015 году.
5. Некоторые ссылки на источники оформлены некорректно, например «Ян Фостер (Ian Foster) [19]» (в авторах статьи нет Фостера), «DEBS 2012 Grand Challenge [153]» (источник отсутствует в списке).

На основании анализа текстов диссертации и автореферата можно сделать следующее *заключение*. Диссертация А.Б.А. Алаасама является самостоятельно выполненной и законченной научно-квалификационной работой, в которой решена задача организации потоков работ в туманных вычислительных средах, имеющая существенное значение для технологий распределенной обработки данных. Диссертационная работа в полной мере отвечает требованиям Положения о порядке присуждения ученых степеней, включая п. 9, а ее автор, Амир Басим Абдуламир Алаасам, заслуживает присуждения ему ученой степени кандидата физико-математических наук по специальности 05.13.11 – математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей.

Старший научный сотрудник  
ФГБУН «Институт проблем передачи  
информации им. А.А. Харкевича РАН»,  
кандидат техн. наук

О.В. Сухорослов

Адрес организации: 127994, Москва, Большой Каретный пер., д. 19, стр. 1  
Телефон: +7 (495) 650-42-25  
Адрес электронной почты: [sukhoroslov@iitp.ru](mailto:sukhoroslov@iitp.ru)

«25» мая 2022 г.

подпись удостоверяю  
ведущий специалист  
по кадрам  
С.Г. Гиминская  
26.05.2022

## СВЕДЕНИЯ ОБ ОФИЦИАЛЬНОМ ОППОНЕНТЕ


диссертации 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

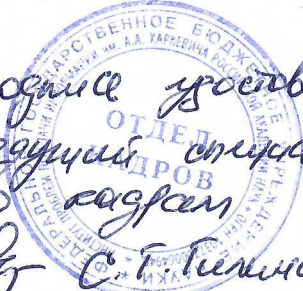
<b>Фамилия, имя, отчество</b>	Сухорослов Олег Викторович
<b>Ученая степень (с указанием номера и шифра специальности)</b>	Кандидат технических наук, 05.13.18
<b>Ученое звание</b>	-
<b>Организация основного места работы</b>	Федеральное государственное бюджетное учреждение науки «Институт проблем передачи информации им. А.А. Харкевича РАН»
<b>Ведомственная принадлежность</b>	Российская Академия наук
<b>Занимаемая должность</b>	Старший научный сотрудник
<b>Почтовый адрес</b>	127994, Москва, Большой Каретный пер., д. 19, стр. 1. ФГБУН Институт проблем передачи информации им. А.А. Харкевича РАН
<b>Телефон</b>	+7 (495) 650-42-25
<b>Адрес электронной почты</b>	sukhoroslov@iitp.ru

### Список основных публикаций по теме диссертации в рецензируемых научных изданиях

1. Sukhoroslov, O. Toward efficient execution of data-intensive workflows // Journal of Supercomputing, 77 (8), pp. 7989-8012. 2021. DOI: 10.1007/s11227-020-03612-4
2. Afanasiev, A.P., Krivonozhko, V.E., Lychev, A.V., Sukhoroslov, O.V. Multidimensional frontier visualization based on optimization methods using parallel computations // Journal of Global Optimization, 76 (3), pp. 563-574. 2020. DOI: 10.1007/s10898-019-00812-y
3. Sukhoroslov, O., Nazarenko, A., Aleksandrov, R. An experimental study of scheduling algorithms for many-task applications // Journal of Supercomputing, 75 (12), pp. 7857-7871. 2019. DOI: 10.1007/s11227-018-2553-9

4. Rumyantsev, A., Sukhoroslov, O., Eparskaya, A., Blanzieri, E., Cavecchia, V. Parameter sweep experiments in hybrid computing systems with r language // International Journal of Innovative Technology and Exploring Engineering, 8 (7), pp. 590-596. 2019
5. Sukhoroslov, O. Supporting Efficient Execution of Workflows on Everest Platform // Communications in Computer and Information Science, 1129 CCIS, pp. 713-724. 2019. DOI: 10.1007/978-3-030-36592-9\_58
6. Sukhoroslov, O. Building web-based services for practical exercises in parallel and dis-tributed computing // Journal of Parallel and Distributed Computing, 118, pp. 177-188. 2018. DOI: 10.1016/j.jpdc.2018.02.024
7. Volkov, S., Sukhoroslov, O. Simplifying the Use of Clouds for Scientific Computing with Everest // Procedia Computer Science, 119, pp. 112-120. 2018. DOI: 10.1016/j.procs.2017.11.167
8. Nazarenko, A., Sukhoroslov, O. Using simulation to improve workflow scheduling in heterogeneous computing systems // Communications in Computer and Information Science, 793, pp. 407-417. 2017. DOI: 10.1007/978-3-319-71255-0\_33
9. Sukhoroslov, O. Integration of Everest platform with BOINC-based desktop grids // CEUR Workshop Proceedings, 1973, pp. 102-107. 2017
10. Nazarenko, A., Sukhoroslov, O. An experimental study of workflow scheduling algorithms for heterogeneous systems // Lecture Notes in Computer Science (including sub-series Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10421 LNCS, pp. 327-341. 2017. DOI: 10.1007/978-3-319-62932-2\_32

 / Сухорослов О. В. /

  
подпись удостоверяю  
ведущий специалист  
по кадрам  
Иль С. Г. Тимина  
26.05.2022