

Метод построения цифровых двойников для решения задач эффективного управления и развития распределенных центров сбора, хранения и обработки данных

Д. И. Пряхина*, В. В. Кореньков, В. В. Трофимов

Международная межправительственная организация Объединенный институт ядерных исследований, г. Дубна, Российская Федерация

Адрес: 141980, Российская Федерация, Московская область, г. Дубна, ул. Жолио-Кюри, д. 6

* pryahinad@jinr.ru

Аннотация

Современные научные эксперименты класса «мегасайенс» не могут обойтись без распределенных центров сбора, хранения и обработки данных. Для проектирования и развития таких сложных систем предлагается использовать цифровые двойники. В статье сформулирован метод построения цифровых двойников распределенных центров сбора, хранения и обработки данных, который включает в себя главным образом описание структуры сложной системы, формирование ее виртуального образа и моделирование. Ключевым этапом метода является моделирование, которое должно учитывать характеристики потоков данных и задач, а также вероятности возможных изменений в процессах функционирования системы. Помимо этого, важными этапами являются процессы описания системы и представления результатов работы цифрового двойника. Особое внимание в статье уделено описанию основных алгоритмов для реализации метода построения цифровых двойников, к которым относится алгоритм работы ядра моделирования и модуля для взаимодействия пользователя с цифровым двойником. Разработанный метод ориентирован на построение цифровых двойников для решения задач эффективного управления и развития распределенных центров сбора, хранения и обработки данных, например, по части определения конфигурации требуемого оборудования среди множества вариантов путем тестирования различных сценариев масштабирования.

Ключевые слова: построение цифрового двойника, распределенный центр хранения и обработки данных, моделирование, возможности для пользователя

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Для цитирования: Пряхина Д. И., Кореньков В. В., Трофимов В. В. Метод построения цифровых двойников для решения задач эффективного управления и развития распределенных центров сбора, хранения и обработки данных // Современные информационные технологии и ИТ-образование. 2023. Т. 19, № 2. С. 272-281. doi: <https://doi.org/10.25559/SITITO.019.202302.272-281>

© Пряхина Д. И., Кореньков В. В., Трофимов В. В., 2023



Контент доступен под лицензией Creative Commons Attribution 4.0 License.
The content is available under Creative Commons Attribution 4.0 License.



Method for Constructing Digital Twins to Solve Problems of Effective Management and Development of Distributed Centers for Collecting, Storing and Processing Data

D. I. Pryahina*, V. V. Korenkov, V. V. Trofimov

Joint Institute for Nuclear Research, Dubna, Russian Federation

Address: 6 Joliot-Curie St., Dubna 141980, Moscow region, Russian Federation

* pryahinad@jinr.ru

Abstract

Modern scientific experiments of the “megascience” class cannot do without distributed centers for collecting, storing and processing data. It is proposed to use digital twins for the design and development of such complex systems. The article formulates a method for constructing digital twins of distributed centers for collecting, storing and processing data, which mainly includes a description of the structure of a complex system, the formation of its virtual image and modeling. The key stage of the method is modeling, which should take into account the characteristics of data flows and tasks, as well as the probability of possible changes in the system activity. In addition, the processes of describing the system and presenting the results of the digital twin are important steps. The article pays special attention to the description of the main algorithms for implementing the method of constructing digital twins, which include the algorithm for the modeling core and the module for user interaction with the digital twin. The developed method is focused on building digital twins to solve problems of effective management and development of distributed centers for collecting, storing and processing data, for example, in determining the configuration of the required equipment among many options by testing various scaling scenarios.

Keywords: building a digital twin, distributed data storage and processing center, modeling, user capabilities

Conflict of interests: The authors declare no conflict of interest.

For citation: Pryahina D.I., Korenkov V.V., Trofimov V.V. Method for Constructing Digital Twins to Solve Problems of Effective Management and Development of Distributed Centers for Collecting, Storing and Processing Data. *Modern Information Technologies and IT-Education*. 2023;19(2):272-281. doi: <https://doi.org/10.25559/SITITO.019.202302.272-281>



Введение

В настоящее время цифровые двойники (ЦД) применяются во многих областях науки и деятельности человека для разных целей [1-5]. Например, ЦД центров хранения и обработки данных (ЦОД) используются для проектирования объектов нового поколения, устранения неполадок на существующих объектах, ускорения процессов планирования мощностей, а также позволяют увеличить время безотказной работы за счет своевременного информирования о возможных проблемах и нагрузках на компоненты ЦОД, повысить эффективность ЦОД с помощью функции анализа текущего потребления ресурсов и оценки влияния допустимых изменений и т. д. [6]. Такие ЦД являются виртуальными копиями исключительно инженерной инфраструктуры ЦОД, которые территориально расположены в одном физическом пространстве.

Для решения научных задач, хранения и последующей обработки огромного количества данных, полученных с установок экспериментов класса «мегасайенс», используются распределенные центры сбора, хранения и обработки данных (РЦОД), которые имеют сложную географически распределенную инфраструктуру и включают множество ресурсов разных типов [7, 8]. Конечно, распределенные системы также нуждаются в ЦД, который поможет при проектировании, масштабировании и совершенствовании РЦОД. При построении и развитии РЦОД для научных задач очень важно учитывать характеристики потоков данных для хранения и потоков задач, которые обрабатывают эти данные. Это позволит проводить исследования, направленные на оптимизацию стратегии управления потоками задач, анализ используемых ресурсов, оценку необходимого количества ресурсов для конкретных задач согласно предъявленным требованиям и критериям оценки результативности РЦОД. Но на сегодняшний день не существует ЦД, с помощью которых можно описать архитектуру существующего или проектируемого распределенного центра хранения и обработки данных с учетом процессов, происходящих в системе с потоками данных и задач. В связи с этим возникает актуальная задача разработки программного комплекса для создания цифровых двойников с целью решения задач эффективного управления и развития распределенных центров сбора, хранения и обработки данных [9].

Конечно, в первую очередь необходимо разработать метод построения цифровых двойников РЦОД, который в отличие от существующих выполняет моделирование распределенных центров с учетом характеристик потоков данных и задач для хранения и обработки, а также вероятностей появления изменений в процессах, происходящих в РЦОД [10-17].

Метод построения цифровых двойников распределенных центров сбора, хранения и обработки данных

Процесс построения ЦД, как правило, реализуется поэтапно. На первом этапе осуществляется сбор данных о физическом объекте (системе), его составных частях, параметрах функционирования и т. д. [18, 19]. На втором этапе из полученных данных извлекается информация, необходимая для построения виртуального образа рассматриваемого объекта (системы) [20]. При

построении непосредственно виртуального образа на третьем этапе используются методы моделирования. Четвертый этап, как правило, посвящен проверке получившихся моделей, верификации и валидации с физическим прототипом [21]. Это очень важный этап, так как от качества, адекватности, эффективности и надежности построенной модели зависит дальнейшее изучение рассматриваемого физического объекта (системы). При необходимости проводится корректировка модели. Последний, пятый этап построения ЦД предусматривает создание модуля демонстрации результатов работы ЦД, которые должны быть открытыми, доступными и представлены в понятном виде (диаграммы, графики, схемы и т. д.) [22, 23]. Таким образом, основой ЦД являются модели, а непосредственное функционирование ЦД обеспечивается инструментами, которые создаются при помощи различных информационных технологий.

В рамках рассматриваемой задачи по созданию ЦД, который поможет качественно проводить исследование, связанные с созданием и развитием современных РЦОД с учетом характеристик потоков данных и задач, целесообразно разработать уникальное ядро. Уникальность ядра заключается в его универсальности и применимости для моделирования процессов передачи, хранения и обработки данных любых РЦОД. При построении ядра ЦД, в котором будет реализован процесс моделирования РЦОД, необходимо учитывать современные тенденции построения распределенных систем и придерживаться принципа разработки универсального программного пакета, который не нужно изменять под каждую моделируемую инфраструктуру. При моделировании необходимо также учитывать вероятности возможных изменений в процессах функционирования системы в зависимости от типа оборудования (например, потери данных, изменение производительности). Эта особенность позволит применять ЦД для решения широкого класса задач, проводить исследования, связанные с проектированием, масштабированием, повышением производительности РЦОД, анализом эффективности и надежности различных процессов, происходящих в РЦОД в период эксплуатации. В ЦД должна быть предусмотрена функциональная возможность, позволяющая выбирать несколько потенциально конкурентоспособных вариантов из множества допустимых конфигураций оборудования для РЦОД.

Конечно, цифровые двойники РЦОД не могут быть полностью автономными и требуют вмешательства человека, например, для описания моделируемой инфраструктуры, при тестировании различных модификаций распределенной системы, при поиске конфигураций оборудования согласно имеющимся требованиям. Во время проведения экспериментов и работы ЦД структура моделируемой распределенной системы может быть изменена, например, может потребоваться добавление или удаление вычислительных компонентов, хранилищ данных, линий связи между компонентами РЦОД. Это значит, что разрабатываемый ЦД должен быть оснащен хорошо продуманным, удобным и доступным интерфейсом, где пользователь сможет не только задавать требуемые параметры РЦОД, но и получать результаты работы ЦД в понятном и структурированном виде, в том числе на графиках.

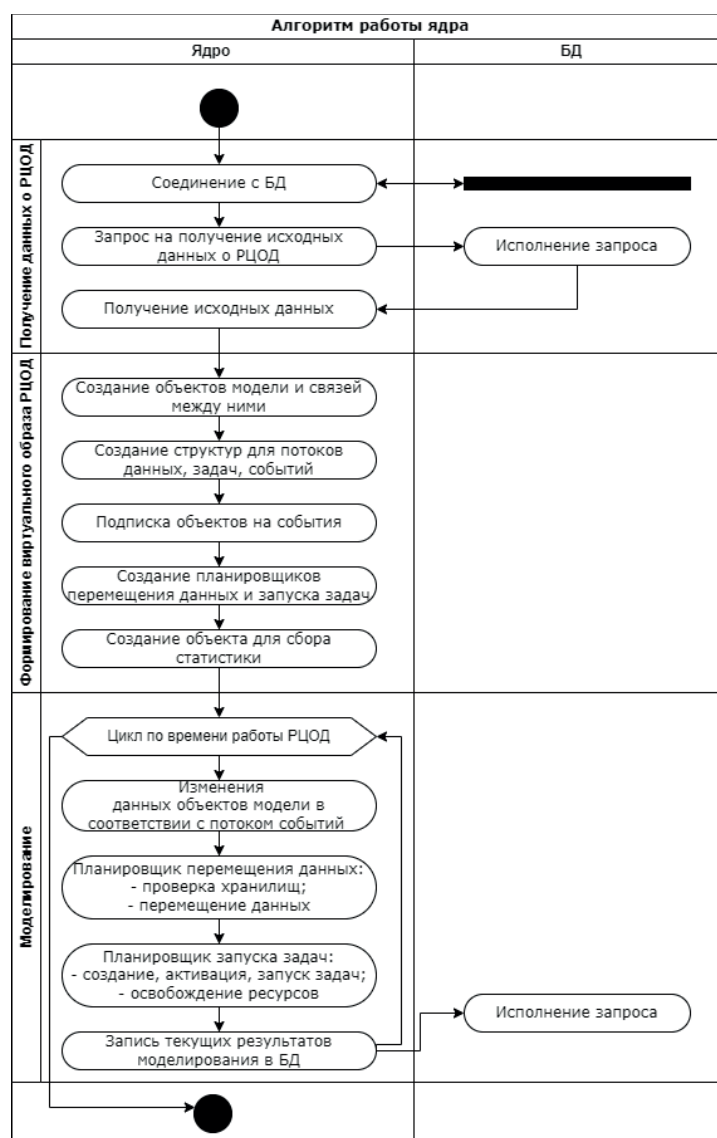
Таким образом, можно сформулировать метод построения цифровых двойников РЦОД. Метод состоит из следующих этапов.



1. Получение данных о РЦОД: параметрах оборудования, потоках данных и задач, целях построения ЦД, вероятностных характеристиках процессов, происходящих в РЦОД.
 2. Описание структуры РЦОД и связей между компонентами.
 3. Формирование виртуального образа РЦОД на основе полученных данных и описания его структуры.
 4. Моделирование РЦОД.
 5. Графическое представление результатов работы ЦД.
- Для реализации каждого этапа метода необходимо разработать алгоритмы.

Алгоритмы для реализации метода построения цифровых двойников распределенного центра сбора, хранения и обработки данных

Ключевым элементом в методе построения цифровых двойников РЦОД является ядро, или программа моделирования распределенных систем с учетом параметров потоков данных и задач для обработки, а также вероятностных характеристик процессов, происходящих в РЦОД. Поэтому, прежде всего необходимо описать алгоритм работы этой программы моделирования (см. рис. 1).



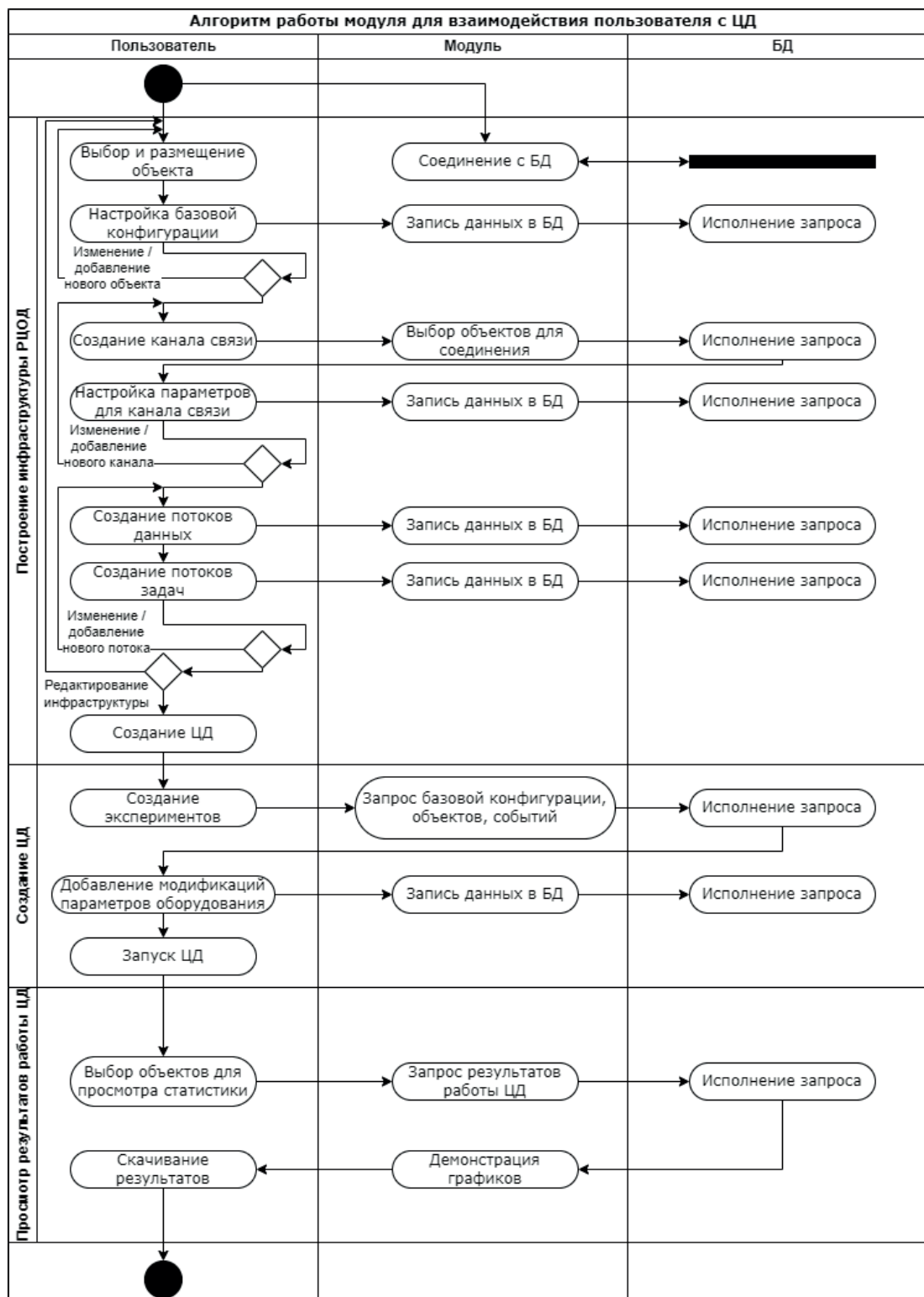
Р и с. 1. Алгоритм работы ядра цифрового двойника РЦОД

F i g. 1. Algorithm for the operation of the core of the digital twin of the distributed data center

Источник: здесь и далее в статье все рисунки составлены авторами.

Source: Hereinafter in this article all figures were made by the authors.





Р и с. 2. Алгоритм работы модуля для взаимодействия пользователя с цифровым двойником РЦОД

F i g. 2. Algorithm of operation of the module for user interaction with the digital twin of the distributed data center



В отличие от существующих средств и программных пакетов¹ [24, 25] ядро должно быть универсальным, так как разработанный алгоритм будет применяться для моделирования процессов передачи, хранения и обработки данных любых РЦОД. Для этого необходимо придерживаться принципа разработки универсального программного пакета, который не нужно изменять под каждую моделируемую инфраструктуру. В связи с этим все входные параметры, которые описывают структуру РЦОД, конфигурации входящего в его состав оборудования, характеристики потоков данных и задач должны храниться в базе данных (БД). Помимо этого, в БД должна содержаться информация о возможных событиях, происходящих в РЦОД. Под событием подразумевается, например, плановая остановка оборудования, его временное прекращение работы, изменение производительности и т. д. В БД должны быть описаны сценарии моделирования и модификации параметров оборудования РЦОД, а также критерии сбора статистики о конкретных объектах в процессе работы будущего ЦД. Таким образом, первым этапом алгоритма является настройка соединения с БД и получение всей необходимой информации о моделируемой системе.

Второй этап алгоритма — формирование виртуального образа рассматриваемого РЦОД. На этом этапе по исходным данным, полученным из БД, осуществляется создание объектов будущей модели и связей между ними. Далее определяются структуры для потоков данных и задач, вырабатывается поток событий. Кроме того, на втором этапе создаются объекты, которые должны отвечать за планирование процессов перемещения данных и задач между элементами РЦОД, а также объекты, которые будут собирать и записывать в БД статистическую информацию о процессах, происходящих в различных элементах РЦОД на протяжении третьего этапа алгоритма — моделирования.

Моделирование проводится по постоянным временным единицам, минимальное значение которых составляет 1 секунду. Для каждого построенного на предыдущем этапе алгоритма объекта модели в соответствии с потоком событий производится изменения в данных, которые соответствуют временной единице работы РЦОД. Начало и продолжительность каждого события определяется согласно одному из распределений вероятностей. Планировщики перемещения данных и управления задачами являются ключевыми объектами, которые описывают процессы, происходящие в РЦОД. В каждую единицу времени планировщик перемещения данных проверяет все хранилища на наличие новых данных. В случае их обнаружения запускается процесс передачи данных из текущего хранилища в пункты назначения согласно описанию структуры моделируемого РЦОД. Задача второго планировщика заключается в управлении потоком задач для обработки имеющихся данных. В каждую единицу времени планировщик создает некоторое количество задач, активирует ранее созданные задачи, т. е. посылает в очередь на исполнение на вычислительные компоненты. Для каждого класса задач на вычислительных компонентах имеются индивидуальные пилоты, которые кон-

тролируют процесс запуска задач. Под пилотом понимается алгоритм, который анализирует количество свободных ядер на вычислительном компоненте, выделенных под определенный тип задач. Если доступны свободные ресурсы, то пилот берет задачу из соответствующей очереди и отправляет ее на исполнение. Далее задача удаляется из очереди, занимает вычислительные ресурсы (слоты) и начинает исполняться, т. е. обрабатывать входной файл при его наличии в хранилище. Каждый слот включает в себя некоторое подмножество ядер. В результате выполнения задачи выходной файл записывается в хранилище, с которым связан вычислительный ресурс. После завершения работы каждой задачи планировщик освобождает вычислительные ресурсы. На протяжении всего этапа моделирования осуществляется запись данных в БД, где отражена статистическая информация о всех происходящих процессах. Алгоритм работы модуля для взаимодействия пользователя с ЦД представлен на диаграмме на рисунке 2. Взаимодействие пользователя с ЦД начинается с построения инфраструктуры РЦОД. Пользователь должен иметь возможность выбрать из списка объект, соответствующий оборудованию РЦОД, расположить его в некоторой области и настроить базовую конфигурацию этого устройства. После этого вся информация о добавленном в инфраструктуру объекте сохраняется в БД. После того, как все необходимые объекты будут добавлены, пользователь должен настроить каналы связи между ними. Под каналами связи понимаются логические соединения между компонентами РЦОД, по которым передаются данные. Для настройки каналов следует выбрать объекты, между которыми осуществляется соединение, и задать требуемые параметры. Информация о каналах связи также добавляется в БД. Далее требуется создать потоки данных для обработки на РЦОД и потоки задач, которые будут обрабатывать эти данные. Пользователь может редактировать инфраструктуру РЦОД, базовую конфигурацию оборудования, изменять параметры потоков данных и задач до перехода на следующий этап, который заключается непосредственно в создании ЦД.

ЦД могут создаваться для разных задач, например, для проектирования РЦОД; для проверки работоспособности существующего РЦОД с текущими параметрами оборудования и поиска проблемных мест его функционирования; для масштабирования РЦОД и поиска конфигурации оборудования, которая будет удовлетворять некоторым требованиям, и т. д. В связи с этим пользователю необходимо создать эксперименты, где ЦД будет применяться для решения конкретной задачи. При добавлении эксперимента важно указать объекты и события, информация о которых будет сохраняться в БД. В каждом эксперименте будет доступ к базовой конфигурации оборудования, которая создавалась при построении инфраструктуры РЦОД. При этом пользователь может добавить дополнительные модификации параметров оборудования, если это необходимо для решения поставленной задачи. После настройки экспериментов осуществляется запуск ЦД. Если в эксперименте существует несколько модификаций, то ЦД для каждой модификации запускаются параллельно.

¹ Проблемы моделирования GRID-систем и их реализации / О. И. Самоваров, Н. Н. Кузюрин, Д. А. Грушин [и др.] // Научный сервис в сети Интернет: решение больших задач : Труды Всероссийской научной конференции, Новороссийск, 22-28 сентября 2008 года. Новороссийск : Изд-во Московского университета, 2008. С. 83-88. EDN: SYIPZT



В процессе или по завершении работы ЦД пользователю доступны результаты для просмотра и дальнейшего анализа. Пользователю необходимо выбрать тип интересующего оборудования, после чего будут построены интерактивные графики по данным из БД, которые записывались в процессе работы ЦД. Например, графики могут отражать объемы данных и распределение различных типов файлов в хранилищах, использование вычислительных компонентов, нагрузку на каналы связи во время передачи данных и т. д. Интерактивные графики подразумевают возможность масштабирования и выбора данных для просмотра. Также пользователь может скачать результаты работы ЦД для какой-либо модификации в виде изображений с графиками, которые отражают изменения, происходящие в РЦОД.

Заключение

Распределенные центры сбора, хранения и обработки данных, которые используются в крупных научных проектах класса «мегасайенс», нуждаются в создании ЦД для решения задач эффективного управления и развития РЦОД, оптимизации стратегии управления потоками данных и задач, анализа используемых ресурсов. В статье сформулирован метод построения

цифровых двойников РЦОД, который включает в себя несколько этапов: получение данных о распределенной системе, описание структуры системы и связей между компонентами, формирование виртуального образа РЦОД, моделирование, представление результатов работы ЦД. Созданные ЦД на основе разработанного метода позволят проводить исследования эффективности и надежности функционирования РЦОД, проверять различные сценарии масштабирования, учитывая требования к потокам данных и задач.

Для успешной реализации метода разработаны алгоритмы ядра программы моделирования и модуля для взаимодействия пользователя с ЦД. В настоящее время описанные алгоритмы находятся на стадии разработки и тестирования, а в дальнейшем будут объединены в программный комплекс для создания цифровых двойников РЦОД. В состав программного комплекса будет входить ядро моделирования, веб-сервис для взаимодействия пользователя с ЦД и база данных, которая играет ключевую роль при разработке всех алгоритмов и содержит информацию об архитектуре РЦОД, параметрах оборудования, входящего в его состав, характеристиках потоков данных и задач, событиях, происходящих в РЦОД, сценариях возможного масштабирования системы, результатах работы ЦД.

Список использованных источников

- [1] Внедрение цифровых двойников как одно из ключевых направлений цифровизации производства / Н. В. Курганова, М. А. Филин, Д. С. Черняев [и др.] // International Journal of Open Information Technologies. 2019. Т. 7, № 5. С. 105-115. EDN: YLCQWI
- [2] Kukushkin K., Ryabov Yu., Borovkov A. Digital Twins: A Systematic Literature Review Based on Data Analysis and Topic Modeling // Data. 2022. Vol. 7, issue 12. Article number: 173. <https://doi.org/10.3390/data7120173>
- [3] Barricelli B. R., Casiraghi E., Fogli D. A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications // IEEE Access. 2019. Vol. 7. P. 167653-167671. <https://doi.org/10.1109/ACCESS.2019.2953499>
- [4] Wu H., Ji P., Ma H., Xing L. A Comprehensive Review of Digital Twin from the Perspective of Total Process: Data, Models, Networks and Applications // Sensors. 2023. Vol. 23, issue 19. Article number: 8306. <https://doi.org/10.3390/s23198306>
- [5] Digital twins: State of the art theory and practice, challenges, and open research questions / A. Sharma [et al.] // Journal of Industrial Information Integration. 2022. Vol. 30. Article number: 100383. <https://doi.org/10.1016/j.jii.2022.100383>
- [6] Quantifying data center performance / K. Nemati [et al.] // 2018 34th Thermal Measurement, Modeling & Management Symposium (SEMI-THERM). San Jose, CA, USA, 2018. P. 141-147. <https://doi.org/10.1109/SEMI-THERM.2018.8357365>
- [7] Кореньков В. В. Тенденции и перспективы развития распределенных вычислений и аналитики больших данных для поддержки проектов класса мегасайенс // Ядерная физика. 2020. Т. 83, № 6. С. 534-538. <https://doi.org/10.31857/S0044002720050153>
- [8] Digital twin-driven product design, manufacturing and service with big data / F. Tao [et al.] // The International Journal of Advanced Manufacturing Technology. 2018. Vol. 94, issue 9-12. P. 3563-3576. <https://doi.org/10.1007/s00170-017-0233-1>
- [9] Пряхина Д. И., Кореньков В. В. Актуальность создания цифрового двойника для управления распределенными центрами сбора, хранения и обработки данных // Современные информационные технологии и ИТ-образование. 2023. Т. 19, № 2. С. 262-271. doi: <https://doi.org/10.25559/SITITO.019.202302.262-271>
- [10] Моделирование грид и облачных сервисов как средство повышения эффективности их разработки / В. В. Кореньков [и др.] // CEUR Workshop Proceedings. 2014. Т. 1297. С. 13-19. URL: https://ceur-ws.org/Vol-1297/13-19_paper-4.pdf (дата обращения: 08.09.2023).
- [11] Service monitoring system for JINR Tier-1 / I. Kadochnikov [et al.] // EPJ Web of Conferences. 2019. Vol. 214. Article number: 08016. <https://doi.org/10.1051/epjconf/201921408016>
- [12] Моделирование грид и облачных сервисов как важный этап их разработки / В. В. Кореньков [и др.] // Системы и средства информатики. 2015. Т. 25, № 1. С. 4-19. <https://doi.org/10.14357/08696527150101>
- [13] Нечаевский А. В., Пряхина Д. И., Ужинский А. В. Разработка веб-сервиса для моделирования систем хранения и обработки данных физических экспериментов // Системный анализ в науке и образовании. 2015. № 4. С. 28-35. EDN: WMEKGF
- [14] Web-Service Development of the Grid-Cloud Simulation Tools / V. V. Korenkov [et al.] // Procedia Computer Science. 2015. Vol. 66. P. 533-539. <https://doi.org/10.1016/j.procs.2015.11.060>



- [15] Simulation concept of NICA-MPD-SPD Tier0-Tier1 computing facilities / V. V. Korenkov [et al.] // *Physics of Particles and Nuclei Letters*. 2016. Vol. 13, no. 5. P. 693-699. <https://doi.org/10.1134/S1547477116050290>
- [16] Моделирование межпроцессорного взаимодействия при выполнении MPI-приложений в облаке / Н. А. Кутовский // *Компьютерные исследования и моделирование*. 2017. Т. 9, № 6. С. 955-963. <https://doi.org/10.20537/2076-7633-2017-9-6-955-963>
- [17] Simulation approach for improving the computing network topology and performance of the China IHEP Data Center / A. Nechaevskiy [et al.] // *EPJ Web of Conferences*. 2019. Vol. 214. Article number: 08018. <https://doi.org/10.1051/epjconf/201921408018>
- [18] Batty M. Digital twins // *Environment and Planning B: Urban Analytics and City Science*. 2018. Vol. 45, issue 5. P. 817-820. <https://doi.org/10.1177/2399808318796416>
- [19] Grieves M., Vickers J. Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems // *Transdisciplinary Perspectives on Complex Systems*; ed. by J. Kahlen, S. Flumerfelt, A. Alves. Cham : Springer, 2017. P. 85-13. https://doi.org/10.1007/978-3-319-38756-7_4
- [20] Boschert S., Rosen R. Digital Twin – The Simulation Aspect // *Mechatronic Futures*; ed. by P. Hehenberger, D. Bradley. Cham : Springer, 2016. P. 59-74. https://doi.org/10.1007/978-3-319-32156-1_5
- [21] Cimino C., Negri E., Fumagalli L. Review of digital twin applications in manufacturing // *Computers in Industry*. 2019. Vol. 113. Article number: 103130. <https://doi.org/10.1016/j.compind.2019.103130>
- [22] Virtualization of Production Using Digital Twin Technology / I. Halenar, M. Juhas, D. Juhasova, D. Borkin // 2019 20th International Carpathian Control Conference (ICCC). Krakow-Wieliczka, Poland : IEEE Computer Society, 2019. P. 1-5. <https://doi.org/10.1109/CarpathianCC.2019.8765940>
- [23] Каменщиков М. А., Корниенко В. Н. Grid и технология открытых систем // *Информационные технологии и вычислительные системы*. 2003. № 3. С. 45-50. URL: <https://www.mathnet.ru/links/0814f2883c17cc2147b89da94571bfe1/itvs697.pdf> (дата обращения: 08.09.2023).
- [24] Кореньков В. В., Нечаевский А. В. Пакеты моделирования DataGrid // *Системный анализ в науке и образовании*. 2009. № 1. С. 21-35. EDN: KNNWPL
- [25] Синтез процессов моделирования и мониторинга для развития систем хранения и обработки больших массивов данных в физических экспериментах / В. В. Кореньков [и др.] // *Компьютерные исследования и моделирование*. 2015. Т. 7, № 3. С. 691-698. <https://doi.org/10.20537/2076-7633-2015-7-3-691-698>

Поступила 08.09.2023; одобрена после рецензирования 04.10.2023; принята к публикации 12.10.2023.

Об авторах:

Пряхина Дарья Игоревна, научный сотрудник Лаборатории информационных технологий имени М.Г. Мещерякова, Международная межправительственная организация Объединенный институт ядерных исследований (141980, Российская Федерация, Московская область, г. Дубна, ул. Жолио-Кюри, д. 6), **ORCID:** <https://orcid.org/0000-0002-7523-5459>, pryahinad@jinr.ru

Кореньков Владимир Васильевич, научный руководитель Лаборатории информационных технологий имени М.Г. Мещерякова, Международная межправительственная организация Объединенный институт ядерных исследований (141980, Российская Федерация, Московская область, г. Дубна, ул. Жолио-Кюри, д. 6), доктор технических наук, профессор, **ORCID:** <https://orcid.org/0000-0002-2342-7862>, korenkov@jinr.ru

Трофимов Владимир Валентинович, ведущий программист Лаборатории информационных технологий имени М.Г. Мещерякова, Международная межправительственная организация Объединенный институт ядерных исследований (141980, Российская Федерация, Московская область, г. Дубна, ул. Жолио-Кюри, д. 6), **ORCID:** <https://orcid.org/0009-0001-6531-5615>, tvv@jinr.ru

Все авторы прочитали и одобрили окончательный вариант рукописи.

References

- [1] Kurganova N., Filin M., Cherniaev D., Shaklein A., Namiot D. Digital twins' introduction as one of the major directions of industrial digitalization. *International Journal of Open Information Technologies*. 2019;7(5):105-115. (In Russ., abstract in Eng.) EDN: YLCQWI
- [2] Kukushkin K., Ryabov Yu., Borovkov A. Digital Twins: A Systematic Literature Review Based on Data Analysis and Topic Modeling. *Data*. 2022;7(12):173. <https://doi.org/10.3390/data7120173>
- [3] Barricelli B.R., Casiraghi E., Fogli D. A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications. *IEEE Access*. 2019;7:167653-167671. <https://doi.org/10.1109/ACCESS.2019.2953499>
- [4] Wu H., Ji P., Ma H., Xing L. A Comprehensive Review of Digital Twin from the Perspective of Total Process: Data, Models, Networks and Applications. *Sensors*. 2023;23(19):8306. <https://doi.org/10.3390/s23198306>
- [5] Sharma A., Kosasih E., Zhang J., Brintrup A., Calinescu A. Digital twins: State of the art theory and practice, challenges, and open research questions. *Journal of Industrial Information Integration*. 2022;30:100383. <https://doi.org/10.1016/j.jii.2022.100383>



- [6] Nemati K., Zabalegui A., Bana M., Seymour M.J. Quantifying data center performance. In: 2018 34th Thermal Measurement, Modeling & Management Symposium (SEMI-THERM). San Jose, CA, USA; 2018. p. 141-147. <https://doi.org/10.1109/SEMI-THERM.2018.8357365>
- [7] Korenkov V.V. Trends and prospects for the development of distributed computing and big data analytics to support Megascience Projects. *Nuclear Physics*. 2020;83(6):534-538. (In Russ., abstract in Eng.) <https://doi.org/10.31857/S0044002720050153>
- [8] Tao F., Cheng J., Qi Q., Zhang M., Zhang H., Sui F. Digital twin-driven product design, manufacturing and service with big data // The International Journal of Advanced Manufacturing Technology. 2018. Vol. 94, issue 9-12. P. 3563-3576. <https://doi.org/10.1007/s00170-017-0233-1>
- [9] Priakhina D.I., Korenkov V.V. The Relevance of Creating a Digital Twin for Managing Distributed Centers for Collecting, Storing and Processing Data. *Modern Information Technologies and IT-Education*. 2023;19(2):262-271. doi: <https://doi.org/10.25559/SITI-T0.019.202302.262-271>
- [10] Korenkov V.V., Nechaevskiy A.V., Ososkov G.A., Pryakhina D.I., Trofimov V.V., Uzhinskiy A.V. Simulation of Grid and Cloud Services as the Means of the Efficiency Improvement of Their Development. *CEUR Workshop Proceedings*. 2014;1297:13-19. Available at: https://ceur-ws.org/Vol-1297/13-19_paper-4.pdf (accessed 08.09.2023). (In Russ., abstract in Eng.)
- [11] Kadochnikov I., Korenkov V., Mitsyn V., Pelevanyuk I., Strizh T. Service monitoring system for JINR Tier-1. *EPJ Web of Conferences*. 2019;214:08016. <https://doi.org/10.1051/epjconf/201921408016>
- [12] Korenkov V.V., Nechaevskiy A.V., Ososkov G.A., Pryakhina D.I., Trofimov V.V., Uzhinskiy A.V. Grid and cloud services simulation as an important step of their development. *Sistemy i Sredstva Informatiki = Systems and Means of Informatics*. 2015;25(1):4-19. (In Russ., abstract in Eng.) <https://doi.org/10.14357/08696527150101>
- [13] Nechaevskiy A.V., Pryakhina D.I., Uzhinskiy A.V. Web-service development for the physical experiments data storage and processing simulation. *Sistemnyy analiz v nauke i obrazovanii = System Analysis in Science and Education*. 2015;(4):28-35. (In Russ., abstract in Eng.) EDN: WMEKGF
- [14] Korenkov V., Nechaevskiy A., Ososkov G., Pryakhina D., Trofimov V., Uzhinskiy A., Balashov N. Web-Service Development of the Grid-Cloud Simulation Tools. *Procedia Computer Science*. 2015;66:533-539. <https://doi.org/10.1016/j.procs.2015.11.060>
- [15] Korenkov V., Nechaevskiy A., Ososkov G., Pryakhina D., Trofimov V., Uzhinskiy A. Simulation concept of NICA-MPD-SPD Tier0-Tier1 computing facilities. *Physics of Particles and Nuclei Letters*. 2016;13(5):693-699. <https://doi.org/10.1134/S1547477116050290>
- [16] Kutovskiy N.A., Nechaevskiy A.V., Ososkov G.A., Pryakhina D.I., Trofimov V.V. Simulation of interprocessor interactions for MPI-applications in the cloud infrastructure. *Computer Research and Modeling*. 2017;9(6):955-963. (In Russ., abstract in Eng.) <https://doi.org/10.20537/2076-7633-2017-9-6-955-963>
- [17] Nechaevskiy A., Ososkov G., Pryakhina D., Trofimov V., Li W. Simulation approach for improving the computing network topology and performance of the China IHEP Data Center. *EPJ Web of Conferences*. 2019;214:08018. <https://doi.org/10.1051/epjconf/201921408018>
- [18] Batty M. Digital twins. *Environment and Planning B: Urban Analytics and City Science*. 2018;45(5):817-820. <https://doi.org/10.1177/2399808318796416>
- [19] Grieves M., Vickers J. Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. In: Kahlen J., Flumerfelt S., Alves A. (eds.) *Transdisciplinary Perspectives on Complex Systems*. Cham: Springer; 2017. p. 85-113. https://doi.org/10.1007/978-3-319-38756-7_4
- [20] Boschert S., Rosen R. Digital Twin – The Simulation Aspect. In: Hehenberger P., Bradley D. *Mechatronic Futures*. Cham: Springer; 2016. p. 59-74. https://doi.org/10.1007/978-3-319-32156-1_5
- [21] Cimino C., Negri E., Fumagalli L. Review of digital twin applications in manufacturing. *Computers in Industry*. 2019;113:103130. <https://doi.org/10.1016/j.compind.2019.103130>
- [22] Halenar I., Juhas M., Juhasova D., Borkin D. Virtualization of Production Using Digital Twin Technology. In: 2019 20th International Carpathian Control Conference (ICCC). Krakow-Wieliczka, Poland: IEEE Computer Society; 2019. p. 1-5. <https://doi.org/10.1109/ICarpathianCC.2019.8765940>
- [23] Kamenshchikov M.A., Kornienko V.N. *Grid i tehnologiya otkrytykh sistem* [Grid and open systems technology]. *Journal of Information Technologies and Computation Systems*. 2003;(3):45-50. Available at: <https://www.mathnet.ru/links/0814f2883c17cc2147b-89da94571bfe1/itvs697.pdf> (accessed 08.09.2023). (In Russ., abstract in Eng.)
- [24] Korenkov V.V., Nechaevskiy A.V. DataGrid simulation packages. *Sistemnyy analiz v nauke i obrazovanii = System Analysis in Science and Education*. 2009;(1):21-35. (In Russ., abstract in Eng.) EDN: KNNWPL
- [25] Korenkov V.V., Nechaevskiy A.V., Ososkov G.A., Pryakhina D.I., Trofimov V.V., Uzhinskiy A.V. Synthesis of the simulation and monitoring processes for the development of big data storage and processing facilities in physical experiments. *Computer Research and Modeling*. 2015;7(3):691-698. (In Russ., abstract in Eng.) <https://doi.org/10.20537/2076-7633-2015-7-3-691-698>

Submitted 08.09.2023; approved after reviewing 04.10.2023; accepted for publication 12.10.2023.



About the authors:

Daria I. Priakhina, Research Fellow of the Mescheryakov Laboratory of Information Technologies, Joint Institute for Nuclear Research (6 Joliot-Curie St., Dubna 141980, Moscow region, Russian Federation), **ORCID:** <https://orcid.org/0000-0002-7523-5459>, pryahinad@jinr.ru

Vladimir V. Korenkov, MLIT Scientific Leader, Joint Institute for Nuclear Research (6 Joliot-Curie St., Dubna 141980, Moscow region, Russian Federation), Dr. Sci. (Tech.), Professor, **ORCID:** <https://orcid.org/0000-0002-2342-7862>, korenkov@jinr.ru

Vladimir V. Trofimov, Senior software developer of the Mescheryakov Laboratory of Information Technologies, Joint Institute for Nuclear Research (6 Joliot-Curie St., Dubna 141980, Moscow region, Russian Federation), **ORCID:** <https://orcid.org/0009-0001-6531-5615>, tvv@jinr.ru

All authors have read and approved the final manuscript.

