Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего образования
«Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ __РОБОТОТЕХНИКА И КОМПЛЕКСНАЯ АВТОМАТИЗАЦИЯ_____

КАФЕДРА ___СИСТЕМЫ АВТОМАТИЗИРОВАННОГО ПРОЕКТИРОВАНИЯ_____

ОТЧЕТ ПО ПРОИЗВОДСТВЕННОЙ ПРАКТИКЕ

Студент_____Голубев Владислав Олгович_____
фамилия, имя, отчество

Группа___РК6-83_____

Тип практики _____Преддипломная_____

Название предприятия____НИИ АПП МГТУ им. Н.Э.Баумана_____

Студент _____28.05.20 г. ___**Голубев В. О.**_____
подпись, дата *фамилия, и.о.*

Руководитель практики _____ ___**Соколов А. П.**____
подпись, дата *фамилия, и.о.*

Оценка _____

2020 г.

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Московский государственный технический университет имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

Кафедра «Системы автоматизированного проектирования» (РК6)

З А Д А Н И Е
на прохождение производственной практики

на предприятии_____г. Москва, НИИ АПП МГТУ им. Н.Э. Баумана_____

Студент_____Голбуев Владислав Олегович____РК6-83_____
(фамилия, имя, отчество; инициалы; индекс группы)

Во время прохождения производственной практики студент должен:

1. Провести обзор литературы по теме: "Методы визуализации и организации процессов решения сложных вычислительных задач" (не менее 15 источников);
2. Оформить отчет о результатах прохождения практики.

Дата выдачи задания « 19 » __мая___ 2020 г.

Руководитель практики от кафедры_____ / _____
(подпись, дата)

Студент _____**28.05.20**_ / _____**Голубев В.О.**_____
(подпись, дата) (Фамилия И.О.)

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ

При проведении современных исследований возникает необходимость автоматизировать процессы решения сложных вычислительных задач. Достижение подобной цели не представляется возможным без формально определенного метода организации процессов в автоматизированной системе. Проектирование, создание и сопровождение подобных систем является трудоемкой задачей, для решения которой применяют инструментальные средства и среды разработки автоматизированных систем (CASE-системы) [1]

Для эффективного взаимодействия с подобной системой исследователю необходим графический пользовательский интерфейс, в котором модель организации вычислений может быть представлена визуально. Большие перспективы в автоматизации процесса решения сложных задач открывает перед исследователем возможность прямого взаимодействия с вычислительной моделью: остановка вычислительного процесса на определенном этапе, изучение обрабатываемых данных, просмотр истории изменения обрабатываемых данных, возврат к определенному этапу вычислений, ввод дополнительных параметров на определенной стадии вычислений и т.д. О необходимости подобных решений писал B.McCormic [2]

Ниже представлен обзор основных работ, в которых рассматриваются методы визуализации и организации сложных вычислительных задач.

1. АНАЛИЗ РЕЗУЛЬТАТОВ ПО ТЕМЕ ИССЛЕДОВАНИЯ

1.1. Результаты поиска источников литературы

Основная идея решения сложных вычислительных задач сводится к их декомпозиции на более мелкие процессы [1].

Методики используемые в CASE предполагают организацию процесса решения в виде иерархической структуры, которую можно представить в виде диаграмм. Например, диаграммы потоков данных (Data Flow Diogram), граф-схемы, диаграммы перехода состояний. Такое описание позволяет выделять структурные единицы приложения в виде функций или подпрограмм и связывать их между собой в определенной последовательности. Продолжением идей структурного описания проектируемого приложения можно считать объектный подход, в котором стали применять методы объектно-ориентированного программирования, что в последствии привело к методологии Component Based Design [3]. Эта технология основывалась на унификации интерфейса для интеграции программного обеспечения в сложную систему. Языком описания для CBD является UML [4]. Для визуализации процессов используются диаграммы последовательностей, диаграммы состояний, диаграммы деятельности.

Метод организации и выполнения процессов основанный на CBD описан Robert D. Bjornson и Stephen B. Weston [5]. В идее метода лежит организация вычислительный процессов, которые состоят из готовых компонентов, в виде диаграммы потока данных. Для создания и редактирования диаграммы предусмотрен графический интерфейс, в котором можно определить связи между процессами посредством указания точек входа и выхода обрабатываемых данных.

Метод разработки программного обеспечения при помощи проектирования моделей, не зависящих от предметной области запатентовали David TalbyScott и David McMaster [6]. Для описания организации процессов составляется UML-модель, которая интерпретируется для последующей генерации исходного кода

на основе шаблонов. Подобная программная разработка может применяться для создания программного обеспечения для реализации сложных вычислительных методов решения задач.

Подход к построению моделей обработки данных вычислительных экспериментов предложили российские исследователи Леонтьев, Тарасов, Харитонов [7] . Для построения моделей используются сети Петри. Модель вычислительного эксперимента состоит из моделей вычислительного и управляющего процессов. Построение этих моделей происходит раздельно. Модель вычислительного процесса строится автоматически на основе дерева событий. Модель управляющего процесса строится на определении шаблонов реакции на события. Подобный подход позволяет построить гибкую структуру вычислительного эксперимента и получать сообщения о статусе эксперимента от управляющего процесса.

Решение сложных вычислительных задач обусловлено запуском решения на производительных вычислительных машинах. Обычно к методам высокопроизводительных вычислений относят:

1. параллельные вычисления;
2. распределенные вычисления:
   - GRID
   - Облачные вычисления (PaaS, IaaS, SaaS)

GRID-системы применяются учеными для вычислительных задач, когда производительности кластеров из мощных серверов недостаточно [8].

Алекперов Р.К. [9] в своем исследовании рассмотрел организацию распределенных вычислительных сред на основе GRID-технологии и проанализировал специфические особенности этих задач. В своей работе он привел практические примеры, когда распределенные вычисления помогли решать задачи, требующие десятки лет вычислений, за несколько недель.

В работе ученых Farkas Z. И Kacsuk P. описана разработка графического пользовательского интерфейса для взаимодействия с различными видами GRID-сетей [10]. В системе предусмотрен сервис для мониторинга выполнения

задач в рамках архитектуры и визуализация статуса выполнения. Графический интерфейс представляет из себя поток работ (workflow), который графом описывает структуру GRID-сети, позволяет взаимодействовать с вычислительной системой и визуализировать состояние выполняемых задач. В системе существует Workflow Editor, который позволяет менять структуру вычислительной системы. Взаимодействия с различными видами GRID-сетей реализуется через промежуточный слой. Для каждой архитектуры написаны скрипты, которые переводят действия пользователя в графическом интерфейсе на язык конкретной архитектуры, а также обратно интерпретируют ответы от GRID-систем в workflow.

На сегодняшний день для проведения сложных вычислений многие ученные, исследовательские лаборатории и организации пользуются облачными платформами, которые предоставляют свои вычислительные мощности под любые нужны. Большинство облачных платформ предоставляет пользователю графический интерфейс (workflow) для организации процесса вычислений в виде диаграмм потока данных или граф-схемы [11], [12]. Подобные методы реализуются за счет использования промежуточного программного обеспечения, которое связывает готовые программные компоненты или набором библиотек, с помощью которых можно создавать пользовательские вычислительные процессы и исполнять их в облаке.

1.2. Современные методы организации процессов решения сложных вычислительных задач

В настоящее время организация сложных вычислений все чаще проводится с использованием облачных технологий, поскольку подобных подход снимает большую часть ответственности по инфраструктурному обеспечению вычислительной системы и обходится дешевле, чем использование или разработка GRID-сетей. Существующие облачные вычислительные платформы в основном узко специализированны на решении задач машинного обучения [13]. Существует открытая разработка системы машинного обучения

TensorFlow от компании Google, вычисления в которой выражаются в виде потоков данных через граф состояний [14]. Среди закрытых разработок стоит упомянуть проект вычислительной платформы российской компании Яндекс - Нирвана [15], которая позволяет производить вычисления не завязанные на конкретной предметной области. Подобный подход позволяет использовать вычислительную платформу для широкого класса задач. Среди российских разработок стоит отметить разработанную в МГТУ им. Н.Э.Баумана технологию GBSE, позволяющую упростить процессы проектирования, разработки, тестирования, сопровождения программных реализаций сложных вычислительных методов [16].

1.3. Анализ существующих методов визуализации и организации процессов

Все найденные методы были основаны на организации процессов решения в виде графов, сетей Петри или потоков данных. Визуализировать процессы с подобной структурой удобно с помощью соответствующих диаграмм. Для визуализации процесса решения в рассмотренных методах были предусмотрены механизмы отслеживания состояния и управления исполнением процессов (workflow monitor), в которых можно:

- увидеть детали выполнения процессов;
- увидеть историю выполнения процессов;
- остановить, прекратить или возобновить исполнение решения.

Подобные механизмы реализуются с помощью создания параллельного управляющего процесса, который может обмениваться сообщениями с исполняющим процессом [7], [10]. Такой подход позволяет разделить обязанности и не усложнять логику работы исполняющего процесса.

4. ЗАКЛЮЧЕНИЕ

1. Был проведен обзор научно-технических публикаций и патентов по теме исследования;
2. Приведены современные перспективные разработки в области организации процессов сложных вычислительных задач;
3. Проанализированы существующие методы, выделены их достоинства;
4. Предложен метод визуализации и организации процессов сложных вычислительных задач в рамках технологии GBSE.

СПИСОК ЛИТЕРАТУРЫ

1. Норенков И.П Основы автоматизированного проектирования, М: Изд-во МГТУ им. Н. Э. Баумана, 2006. - 448 с.
2. McCormic B. Visualization in Scientific Computing // New York: Computer Graphics. 1987, 81 p.
3. Mcllroy M. Mass produced software components : NATO Software Engineering Conference (7-11 Oct. 1969), Garmisch, Germany: Scientific Affairs Division, 1969,79 p.
4. ISO/IEC 19501:2005 - Information technology - Open Distributed Processing - Unified Modeling Language (UML) Version 1.4.2., Iso.org, 2005. - 456 с.
5. US 2004/0056908 Al - Turbo Worx, Inc. (New Haven, CT, US) - Method and system for dataflow creation and execution (2004).
6. US 8,949,772 Bl - Amazon Technologies, Inc., (Reno, NV, US) - Dynamic model based software application development (2015).
7. Леонтьев Д.В., Тарасов В.Г., Харитонов И.Д. Модель обработки данных вычислительных экспериментов // Научный сервис в сети Интернет: труды XX Всероссийской научной конференции. М.: ИПМ им. М.В.Келдыша, 2018. - С. 373–386.

8. Zurek R.W, Martin L.J. GridPP: development of the UK computing grid for particle physics // New York: Journal of Physics G: Nuclear and Particle Physics, 2006, p. 1–20.
9. Алекперов А. Организация распределенных вычислений на базе GRID-технологии // Баку: Искусственный интелект. - 2011. - С. 6-14.
10. Farkas Z., Kacsuk P. P-GRADE Portal: A generic workflow system to support user communities // Elsevier. - Volume 27 (Future Generation Computer Systems), 2010, p. 455–465.
11. Godec, P., Pančur, M., Democratized image analytics by visual programming through integration of deep models and small-scale machine learning // Nat Commun. - Volume 10, 2019, C. 33-40.
12. Kudryavtsev A.O., Koshelev V.K., Izbyshev A.O. Development and implementation of the cloud system for solving problems of high // Proceedings of the Institute for System Programming of Russian Academy of Sciences, Volume 24, 2011, p. 13-33.
13. Janez Demsar, Tomaz Curk, Ales Erjavec Orange: Data Mining Toolbox in Python // Journal of Machine Learning Research, 2013 (14), C. 2349–2353.
14. Официальный сайт проекта TensorFlow [Электронный ресурс]. – Режим доступа: https://www.tensorflow.org.
15. Познаём Нирвану – универсальную вычислительную платформу Яндекса [Электронный ресурс]. –Режим доступа: https://habr.com/ru/company/yandex/blog/351016/
16. Соколов А.П., Першин А.Ю. Графоориентированный программный каркас для реализации сложных вычислительных методов. Программирование. – Т.47, No5– 2019, с. 43-55.