

Тема работы. Создание алгоритмов распределения вычислений на гетерогенных системах с несколькими независимыми устройствами для класса методов PCI SPH

Состав коллектива:

1. Хайрулин Сергей Сергеевич, младший научный сотрудник ИСИ СО РАН, старший преподаватель НГУ.
2. Пальянов Андрей Юрьевич, Директор ИСИ СО РАН, с.н.с ИСИ СО РАН, д. ф.-м. н.

Научное содержание работы:

1. Постановка задачи.

Разработка алгоритмов и программных технологий, повышающих эффективность процессов обработки данных в вычислительных машинах и комплексах для семейства алгоритмов моделирования динамики несжимаемой жидкости PCI SPH - за счет использования в параллельном режиме всех доступных вычислительных узлов, интеграции разработанных методов в единую программную систему.

2. Современное состояние проблемы (на момент начала работы).

В рамках проекта OpenWorm, целью которого является полномасштабное моделирование нематоды *C. elegans*, различные системы живого организма моделируются отдельно, и при этом осуществляется обмен данными между ними в процессе симуляции. Так, модель мышечной системы нематоды и ее гидростатический скелет представлены как модель, описанная численным методом PCI SPH - predictor corrector smoothed particle hydrodynamics (модификация метода SPH). Несмотря на большую популярность и гибкость, которую предоставляет метод, в отличие от методов конечных элементов, он обладает таким значительным недостатком, как низкая производительность. В представленной работе предлагается ряд алгоритмов, с помощью которых предполагается увеличить производительность численных методов PCI SPH для задач связанных с моделированием гидродинамики и механики биологических систем и процессов. В то же время полностью Лагранжевые алгоритмы такие как SPH система координат связана с движением жидкости в пространстве (сетка движется вместе с жидкостью) и отвечает фиксированным точкам в среде. Таким образом сплошная среда представляется как дискретное множество частиц. Информация в среде передается исключительно через частицы, что делает трудным распараллеливание вычислений на разных физических узлах кластера или между GPU (графический сопроцессор) в пределах одного вычислительного узла или устройства. Для численных моделей основанных на системе частиц, несмотря на то, что каждый метод хорошо распараллеливаются по данным, представляется довольно сложным распределение вычислений на системах с не общей памятью - такие как узлы кластера общающиеся по сети или различные вычислительные устройства в

рамках одной машины, но независимыми модулями оперативной памяти (GPU). Это в частности обусловлено тем, что частицы динамичны и не привязаны к определенным позициям, как узлы расчетных сеток. Тем ни менее, в работах [Dominguez2013; Verma2017; Verma2018] предложены попытки преодолеть этот недостаток за счет логического разделения моделируемого пространства на не пересекающиеся статичные или динамичные подпространства - домены. Таким образом частицы можно кластеризовать по пространственному признаку, в зависимости от текущей позиции. При этом подразумевается, что каждый домен обрабатывается отдельным устройством-вычислителем. В нашей интерпретации метода PCI SPH мы также основывались на этом подходе ниже в таблице 2 представлено сравнение между нашим и аналогичными подходами. Сравнение представлено в таблице ниже

Источник Доступность (Open source)	Технология вычислений	Алгоритм	Доступность (Open source)
[Dominguez2013]	CUDA + MPI	SPH	-
[Verma2017]	CUDA	SPH	-
[Verma2018]	CUDA	PCI SPH	-
Sibernetic	OpenCL	PCI SPH	+

1. Caballero A, Mao W, Liang L, Oshinski J, Primiano C, McKay R, Kodali S, Sun W. Modeling Left Ventricular Blood Flow Using Smoothed Particle Hydrodynamics. *Cardiovasc Eng Technol*. 2017 Dec;8(4):465-479. doi: 10.1007/s13239-017-0324-z. Epub 2017 Jul 25. PMID: 28744784; PMCID: PMC5709227.
2. Horvath, Christopher & Solenthaler, Barbara. (2013). Mass Preserving Multi-Scale SPH Pixar Technical Memo #13-04.
3. Verma, Kevin & Peng, Chong & Szewc, K. & Wille, Robert. (2018). A Multi-GPU PCISPH Implementation with Efficient Memory Transfers. 10.1109/HPEC.2018.8547542.
4. Domínguez, José & Crespo, Alejandro & Valdez-Balderas, D. & Rogers, Benedict & Gómez-Gesteira, M.. (2013). New multi-GPU implementation for smoothed particle hydrodynamics on heterogeneous clusters. *Computer Physics Communications*. 184. 1848–1860. 10.1016/j.cpc.2013.03.008.

5. K. Verma, K. Szewc and R. Wille, "Advanced load balancing for SPH simulations on multi-GPU architectures," 2017 IEEE High Performance Extreme Computing Conference (HPEC), 2017, pp. 1-7, doi: 10.1109/HPEC.2017.8091093.
6. B. Solenthaler, Predictive-Corrective Incompressible SPH. ACM Transactions on Graphics (Proceedings of SIGGRAPH), 28(3), 2009.
7. Palyanov Andrey, Khayrulin Sergey and Larson Stephen D. 2018 Three-dimensional simulation of the Caenorhabditis elegans body and muscle cells in liquid and gel environments for behavioral analysis Phil. Trans. R. Soc.
8. Palyanov, Andrey Yu et al. "Application of smoothed particle hydrodynamics to modeling mechanisms of biological tissue." Adv. Eng. Softw. 98 (2016): 1-11.

3. Подробное описание работы, включая используемые алгоритмы.

При разработке метода, его тестировании и верификации возникает необходимость проведения большого числа расчетов. Кроме того для проверки алгоритма распределения нагрузки вычислений и синхронизации данных необходима проверка его на машинах, имеющих несколько вычислительных узлов (GPU). В основе предлагаемого нами алгоритма лежит идея распределения данных по доменам, при соблюдении условия, что все данные в пределах каждого домена могут обрабатываться независимо. При этом расчеты не производятся для частиц из смежных доменов. В силу специфики задачи данные об изменениях позиций частиц своевременно синхронизируются между устройствами. Предполагается, что параллельные вычисления для каждого домена будут производиться на различных GPU одновременно, кроме того каждый контроллер каждого решателя, запускается в отдельном потоке. Таким образом одна итерация симуляции предполагает несколько стадий:

- Формирование списка соседей для каждой частицы.
- Расчёт изменения физических величин и сглаживание флуктуаций плотности (PCI SPH).
- Численное интегрирование (Leapfrog, Semi-implicit Euler)
- Синхронизация данных
 - Сортировка, 1 поток qsort, параллельная — модификация поразрядной сортировки
 - Обновление данных

4. Полученные результаты.

Результат работы выполненной в рамках реализации методов для распараллеливания вычислений алгоритмов PCI SPH проведенной в 2020 году были

представлены в отчете опубликованном на сайте вычислительного центра НГУ (http://nusc.nsu.ru/wiki/doku.php/reports/sskhayrulin/20200625_sskhayrulin). Однако несмотря на то, что приведенные числа показали, что предложенные методы значительно ускоряют алгоритм. В результате анализа было выявлено, что в программном коде были недочеты, которые ограничивали возможность симуляции моделей с большим количеством частиц от 20 до 100 миллионов. В этом отчете представлены промежуточные расчеты полученные после модификации программного кода с учетом исправленных недостатков.

5. Иллюстрации, визуализация результатов (опционально).

На иллюстрациях ниже представлены результаты для 3 конфигураций системы, различающихся количеством частиц: 9 миллионов, 14 миллионов и 34 миллиона частиц.

Один тест - это один прогон программы моделирующей эволюцию системы частиц в течении фиксированного времени. Результатом является среднее время выполнения одной итерации симуляции. Для каждого теста варьировались настройки вычислительного кластера, а именно количество вычислительных узлов (GPU) также выключался/выключался режим параллельной сортировки.

~9 миллионов частиц

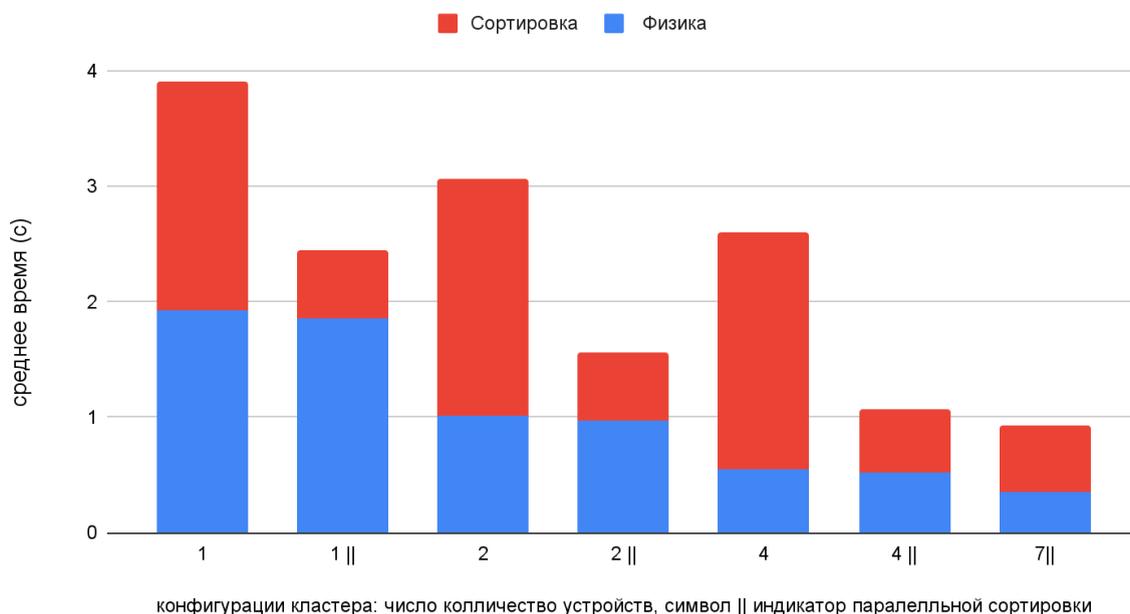


Рис. 1: Результаты для конфигурации из ~9 миллионов частиц.

~14 миллионов частиц

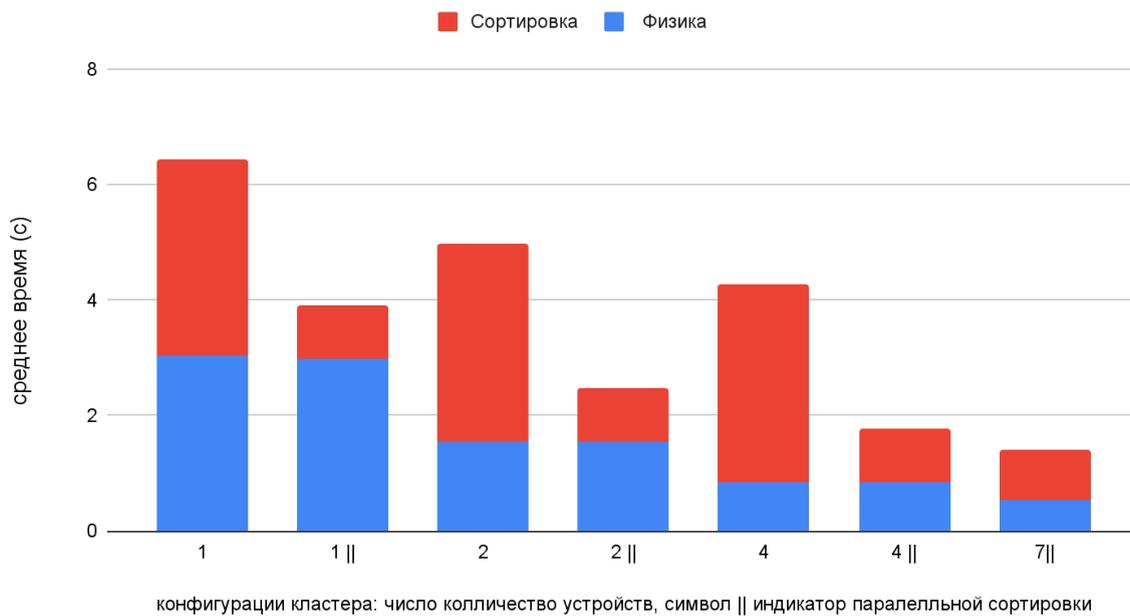


Рис. 2: Результаты для конфигурации из ~14 миллионов частиц.

~34 миллионов частиц

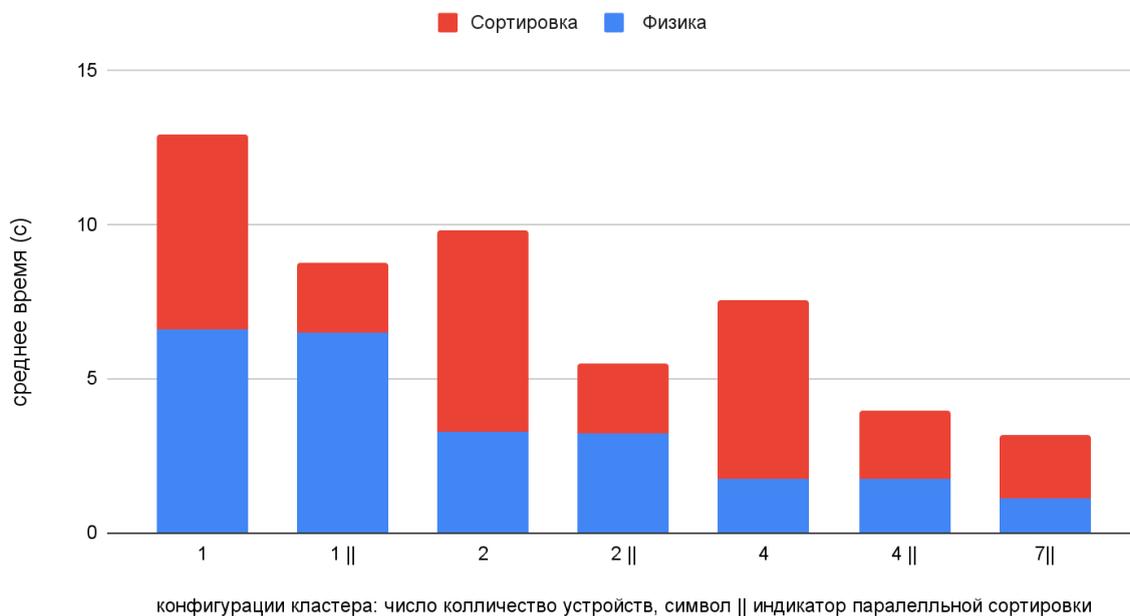


Рис. 3: Результаты для конфигурации из ~34 миллионов частиц.

Как показывают полученные результаты алгоритм распараллеливания вычислений на несколько вычислительных узлов позволяют в значительной мере ускорить вычислительный процесс для параллельных реализации алгоритмов класса PCI SPH.

6. Эффект от использования кластера в достижении целей работы

Использование вычислительного кластера дало возможность реализовать и отладить группу алгоритмов, которые помогают уменьшить время вычислений для алгоритмов PCI SPH. Расчеты проводились на различных конфигурациях вычислительного кластера. Строго говоря, осуществление расчетов, позволяющих судить о корректности и эффективности предложенного алгоритма без специализированного оборудования, на персональных компьютерах фактически является невозможным.

Благодаря качественной документации получилось разобраться с организацией работы на кластере за короткое время. Большая благодарность технической поддержке за своевременную помощь при возникновении проблем.

7. Список публикаций

1. Khayrulin S. Distributed calculations on multiple independent devices for PCISPH method. // Bulletin of the Novosibirsk Computing Center. № 45, 2021. p. 59-68. DOI: 10.31144/bncc.cs.2542-1972.2021.n45.p59-68
2. S.S. Khayrulin. Algorithms of distributed computation on heterogeneous systems with many devices for PCI SPH methods. // Proc. Marchuk Scientific Readings October 19-23, 2020, Akademgorodok, Novosibirsk, Russia. pp 183-184.