



Цифровой блок НИУ ВШЭ
Отдел суперкомпьютерного моделирования

Москва,
23 сентября 2024 г.

Экономия вычислительных ресурсов суперкомпьютера при моделировании материалов и веществ

Начальник отдела суперкомпьютерного моделирования
Костенецкий Павел Сергеевич
к.ф.-м.н., доцент.



Характеристики суперкомпьютера *sHARISMa* НИУ ВШЭ (*Computer of HSE for Artificial Intelligence and Supercomputer Modelling*)

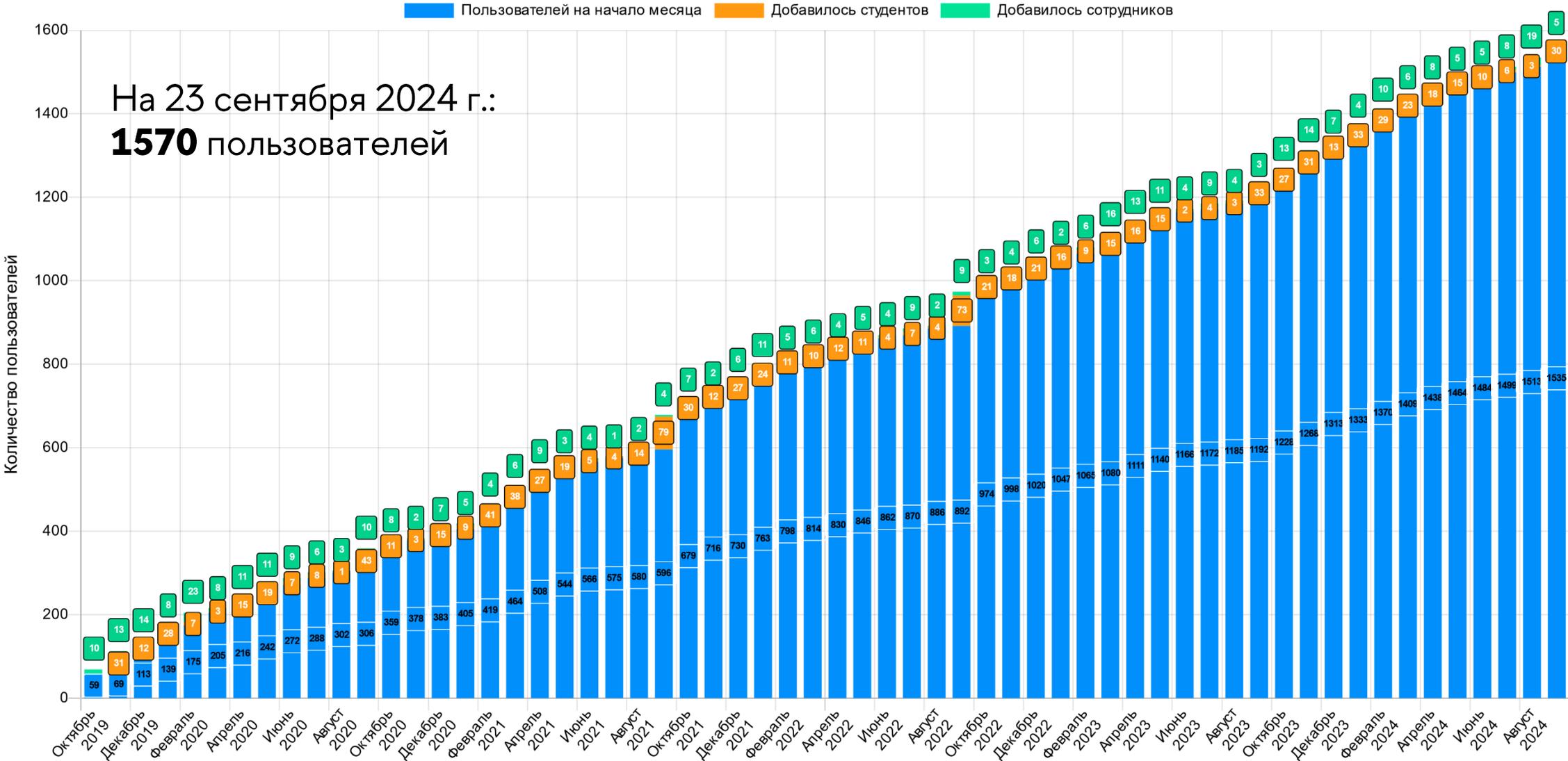
2

- **10 место в ТОП 50 СНГ**
- Пиковая производительность: **2.12 Петафлопс**
(2.12 квадриллиона операций в секунду над числами с двойной точностью)
- LINPACK-производительность: **927.4 Терафлопс**
- **47** вычислительных узлов
 - **6** узлов с **1 ТБ** ОЗУ, **8 GPU A100 80 ГБ SXM**
 - **10** узлов с **1,5 ТБ** ОЗУ, **4 GPU V100 32 ГБ**
 - **19** узлов с **768 ГБ** ОЗУ, **4 GPU V100 32 ГБ**
 - **11** узлов с **384 ГБ** ОЗУ для расчётов на CPU
 - **1** узел с **512 ГБ** ОЗУ, **2 GPU H100 80 ГБ**
- **2** управляющих узла
- **48 GPU NVIDIA Tesla A100 80 ГБ**
- **116 GPU NVIDIA Tesla V100 32 ГБ**
- **2 GPU NVIDIA H100 80 ГБ**
- **2616** ядер центральных процессоров
- Оперативная память: **40,3 ТБ RAM + 7.5 ТБ GPU Memory**
- Дисковая память: **1,26 ПБ**
 - параллельная СХД на базе Lustre **848 ТБ**
 - локальные диски **128 ТБ**
 - сервер резервного копирования **280 ТБ**
- Коммуникационная сеть: **2 x InfiniBand EDR**
(**2x100 Гбит/с**, топология **FatTree**)





Прирост количества пользователей



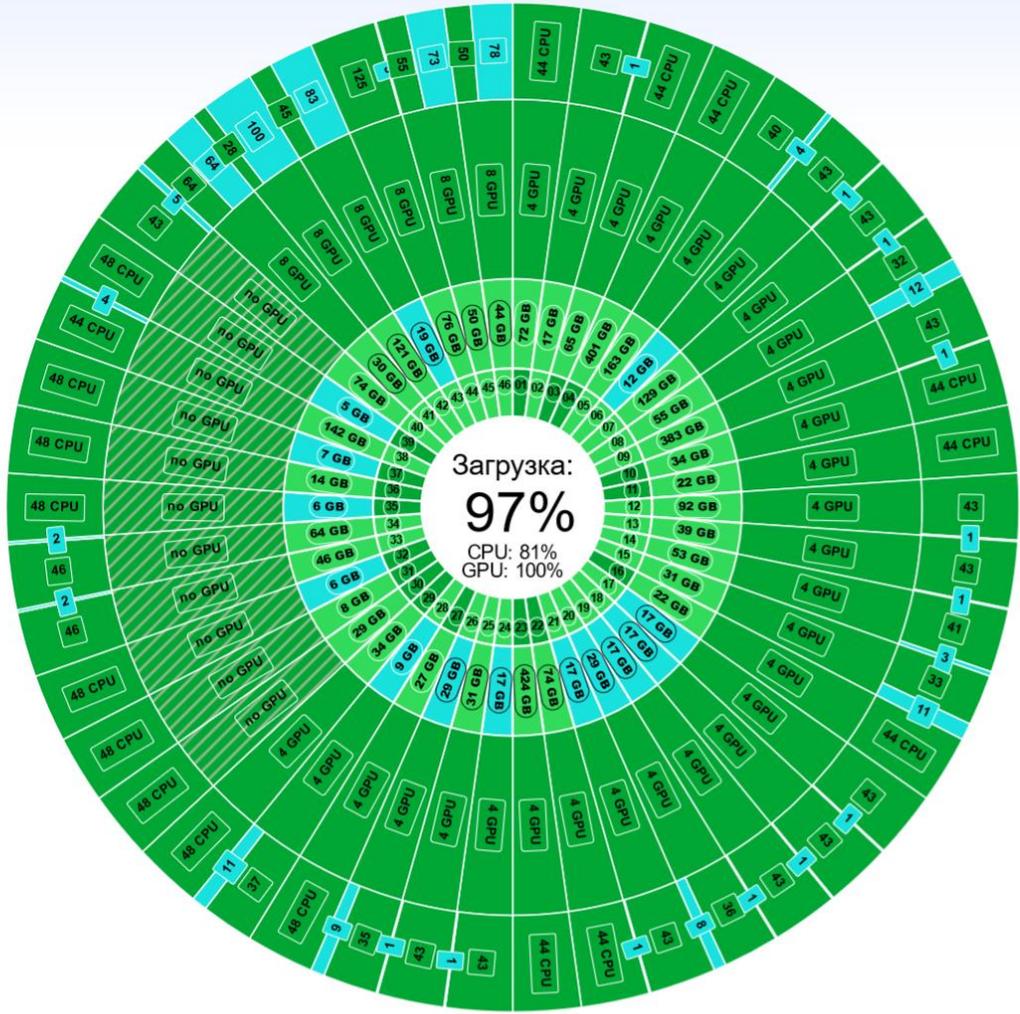
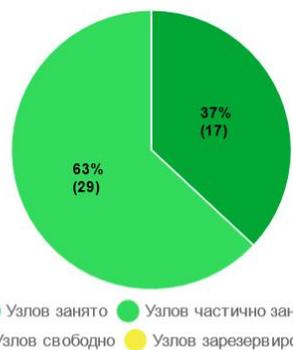


Загрузка суперкомпьютера



Загрузка суперкомпьютерного комплекса НИУ ВШЭ

16:56:21



● Занят ● Частично занят ● Заблокирован ● Свободен ● В резервации ● Отключен

Сейчас считают 36 чел.
 Задач ожидает: 10 (CPU: 134 GPU: 6)
 Задач выполняется: 330 (CPU: 2100 GPU: 164)



■ Загрузка CPU ■ Загрузка GPU



Фильтрация потока задач

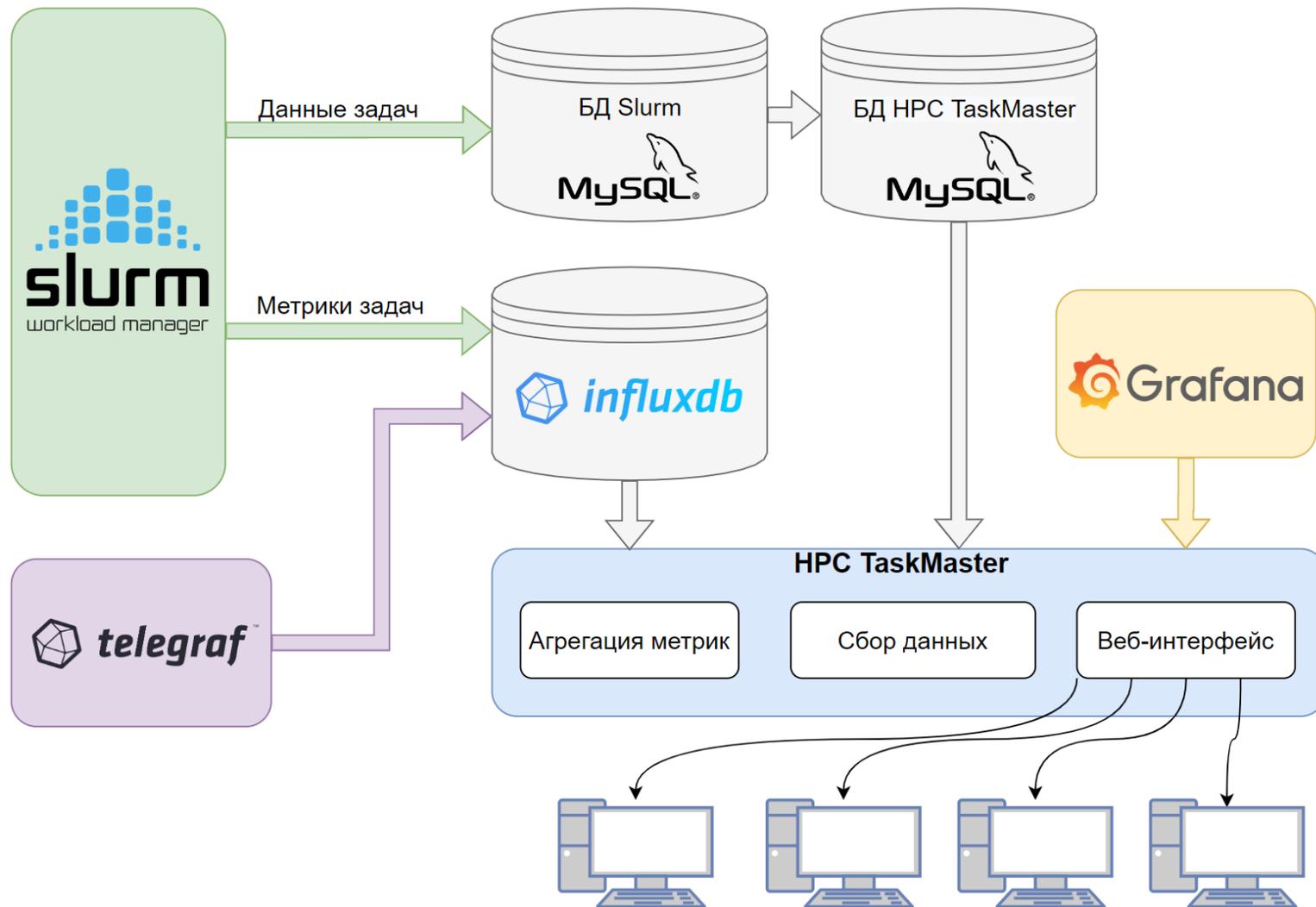
5



- 1570 пользователей выполняют свои расчеты на суперкомпьютере *sCHARISMa*
- Пользователи могут ошибочно запускать неэффективные задачи, которые могут замедлить работу остальных пользователей
- Под неэффективными подразумеваются задачи которые используют не все выделенные ресурсы
- Необходимо автоматически определять неэффективных задач, а также собирать статистику пользователей для обнаружения у них проблем с запусками и проведения консультаций
- Разработанная система позволяет сэкономить до 20% дорогостоящего машинного времени суперкомпьютера



Описание системы HPC TaskMaster



Открытый
исходный код

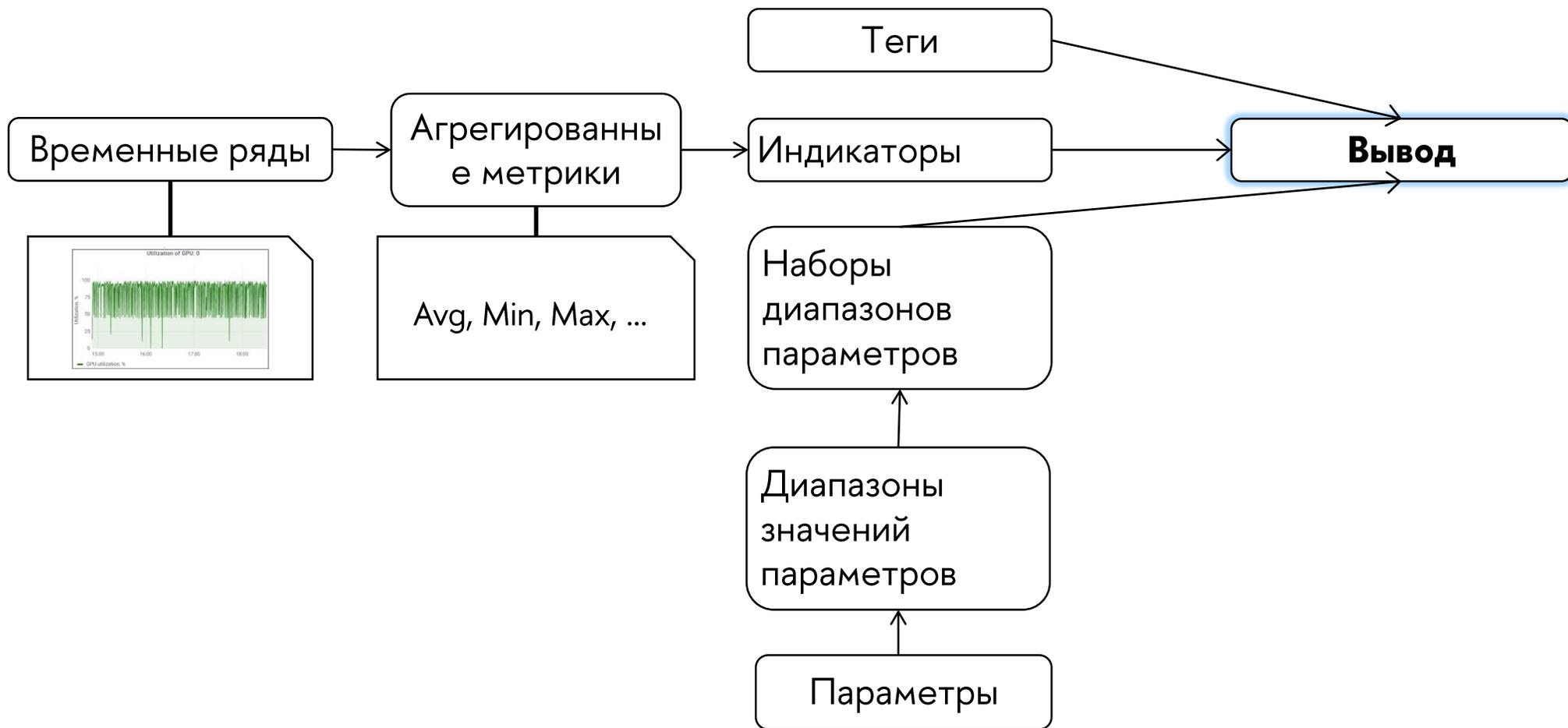
Информация о
задачах, а не об
узлах

Автоматический
анализ метрик

Интерактивные
графики



Этапы анализа эффективности задачи





№	Выводы
1	Рекомендация для задач Quantum Espresso
2	Рекомендация для задач Gromacs
3	Данная программа изнашивает SSD диски суперкомпьютера слишком частой записью данных
4	Приложение Quantum Espresso может работать быстрее
5	Можно выделить GPU другого типа
6	Малое использование оперативной памяти
7	Использование GPU в задачах Quantum Espresso
8	Задача не использует CPU
9	Задача не использует GPU
10	Ускорение задач GROMACS
11	Задачи Gromacs могут вычисляться быстрее
12	Рекомендация для задач LAMMPS
13	Недостаточно данных для анализа
14	Простаивание GPU
15	Простаивание CPU
16	Задача использует GPU недостаточно эффективно
17	Задача использует CPU недостаточно эффективно
18	СХД: Слишком высокая нагрузка
19	Неэффективное использование параллельных вычислений на GPU
20	Неэффективное использование параллельных вычислений на CPU

№	Индикаторы
1	Lustre: высокая нагрузка на запись (предупреждение)
2	Lustre: высокая нагрузка на чтение (рекомендация)
3	CPU: Хорошая средняя загрузка
4	Присутствие пропусков в метриках

№	Теги
1	Тип задачи - Stata
2	Тип задачи - IQ Tree
3	Тип задачи - GNU Octave
4	Тип задачи - Amber
5	Тип задачи - Abinit



Иерархия выводов

9

ID	Приоритет	Вывод	CPU: низкая средняя загрузка	CPU: низкая загрузка отдельных ядер	GPU: низкая средняя загрузка	GPU: используется мало графической памяти (рекомендация)	RAM: используется мало оперативной памяти (рекомендация)	SSD: высокая нагрузка на чтение (рекомендация)
9	1	Неэффективное использование srun/salloc - GPU			+			
1	2	Неэффективное использование srun/salloc - CPU	+	+				
10	3	Неэффективное использование srun/salloc	+	+	+			
5	4	Неэффективное использование Jupyter Notebook	+	+	+			
34	5	Рекомендация для задач Quantum Espresso			+			
35	6	Рекомендация для задач Gromacs			+			
20	7	Данная программа изнашивает SSD диски суперкомпьютера слишком частой записью данных						+
32	8	Приложение Quantum Espresso может работать быстрее	+					
37	9	Можно выделить GPU другого типа			+	+		
30	10	Малое использование оперативной памяти					+	
7	11	Запущена не HPC задача	+					
36	12	Использование GPU в задачах Quantum Espresso						
14	13	Задача не использует одно или несколько ядер CPU		+				
8	14	Задача не использует CPU	+	+				
4	15	Задача не использует GPU			+	+		



Пример графика непараллельной задачи

Был проведен анализ работы двух методов для обнаружения непараллельных задач: корреляция Спирмена и алгоритм САП-трансформ.

$$a_i = \frac{6 \sum_{j=0}^{k-1} (2j - k + 1) y_{i+j}}{hk(k^2 - 1)}, i \in (1, 2, \dots, n - k + 1).$$

Формула локального тренда

$$coss_k(y, x) = \frac{\sum_{i=1}^m a_{yi} \cdot a_{xi}}{\sqrt{\sum_{i=1}^m a_{yi}^2 \cdot \sum_{j=1}^m a_{xj}^2}}$$

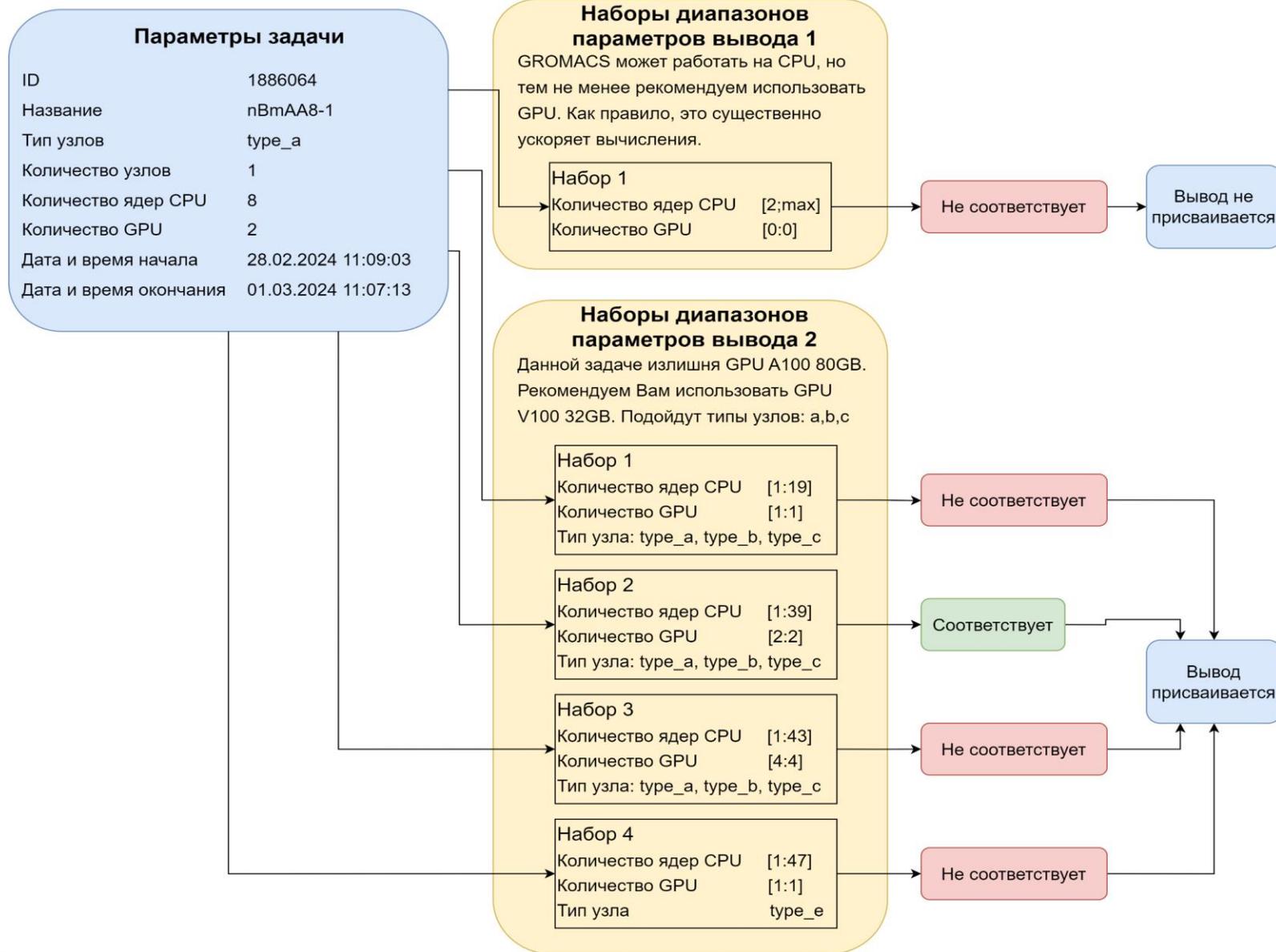
Мера сравнения локальных трендов

Для некоторых длинных задач с большим числом выделенных ресурсов мера САП-трансформ считалась долго, поэтому алгоритм был изменен: для подсчета меры берутся только r последних минут.

Metrica	Spearman	MAP	OptiMAP
Accuracy	0.934088	0.986684	0.986019
ROC AUC	0.630993	0.939587	0.939223
F1	0.407186	0.918699	0.91498
MIN time	0.00119872	0	
MAX time	0.0196644	369.336	
AVG time	0.00234576	1.42191	

Точность определения непараллельных задач у алгоритма САП-трансформ и у оптимизированного алгоритма выше на **80%** по сравнению с алгоритмом, основанным на нахождении корреляции Спирмена.

Примеры рекомендаций пользователям



Удобное формирование вывода на основе параметров

Один вывод может покрывать задачи с различными значениями параметров

Для назначения вывода задаче нужно, чтобы выполнялись все условия хотя бы для одного набора диапазонов параметров



Примеры формирования выводов

Вывод 1

GROMACS может работать на CPU, но тем не менее рекомендуем использовать GPU. Как правило, это существенно ускоряет вычисления.

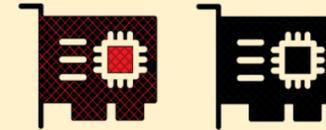
Набор 1	
Количество ядер CPU	[2;max]
Количество GPU	[0:0]



Вывод 2

Данной задаче излишня GPU A100 80GB. Рекомендуем Вам использовать GPU V100 32GB. Подойдут типы узлов: a,b,c

Набор 1	
Количество GPU	[1:1]
Тип узла	type_e



Задача

ID	1886064
Название	nVmAA8-1
Тип узлов	type_d
Количество узлов	1
Количество ядер CPU	8
Количество GPU	0
Дата и время начала	28.02.2024 11:09:03
Дата и время окончания	01.03.2024 11:07:13



Анализ на соответствие вывода 1 задаче

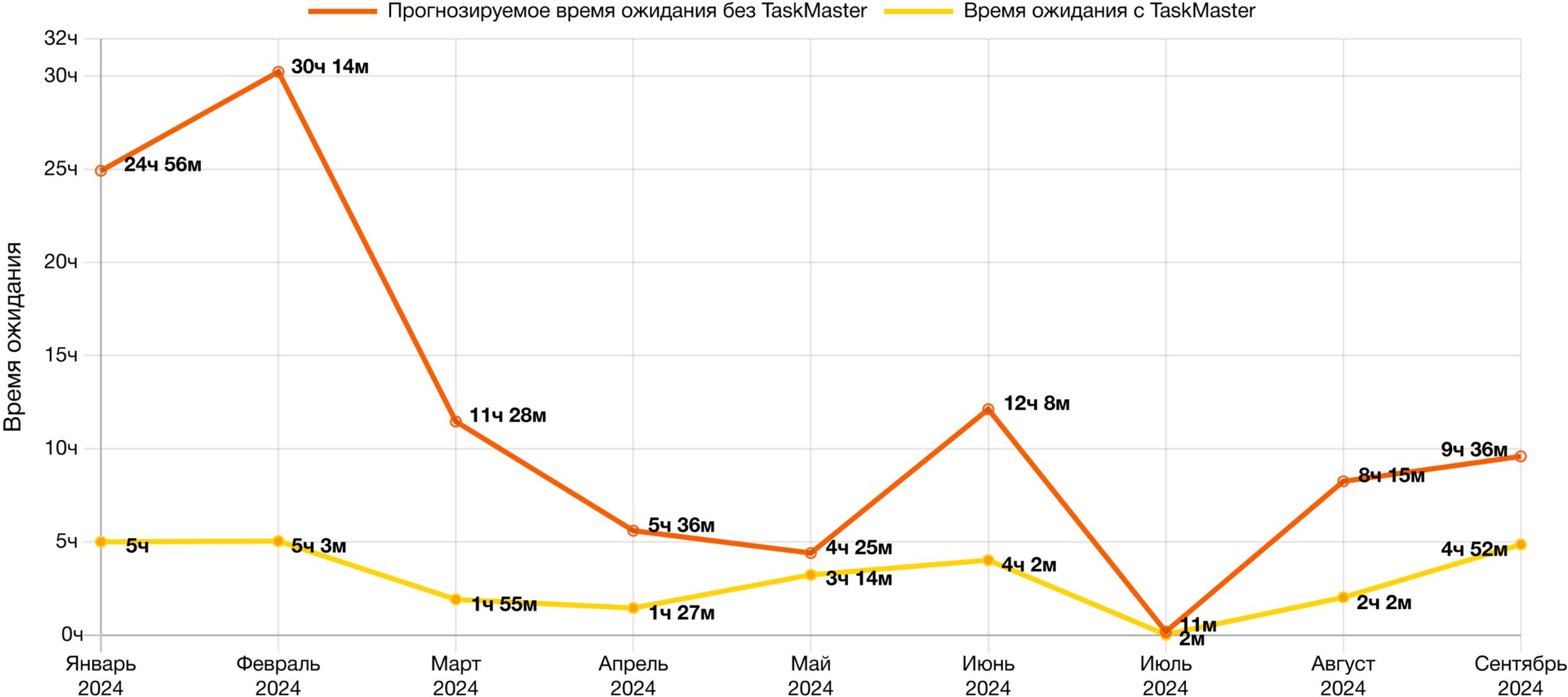
Соответствует

Анализ на соответствие вывода 2 задаче

Не соответствует



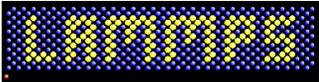
Эффект от HPC TaskMaster





Наиболее используемое ПО для моделирования материалов и веществ

14

ПО	% ресурсов GPU	% ресурсов CPU
	10	8
	6	7
	2	5
	1	3



Пример

15

Вариант выполнения	Ускоренный	Сбалансированный	Экономичный
Используемые ресурсы			
Количество задач, решаемых одновременно	1	2	4
Время на задачу (час)	$1/2$	$2/3$	1
Количество задач, решенных за 1 час	2	3	4

Р Экономия вычислительных ресурсов

Постановка задачи

Имеется параллельный вычислительный комплекс с высокой загруженностью.

Главная цель – эффективное использование ресурсов.

Дополнительно – уменьшение времени выполнения задачи

Решение

Стоимость использования вычислительного оборудования i -й задачей в час: $p_i = (n_i^{cpu} \cdot p_{cpu} + n_i^{gpu} \cdot p_{gpu}) \cdot t_i$

Стоимость электроэнергии за единицу времени: $p_i^e = (n_i^{cpu} \cdot c_{cpu} \cdot u_i^{cpu} + n_i^{gpu} \cdot c_{gpu} \cdot u_i^{gpu}) \cdot e_p \cdot t_i$

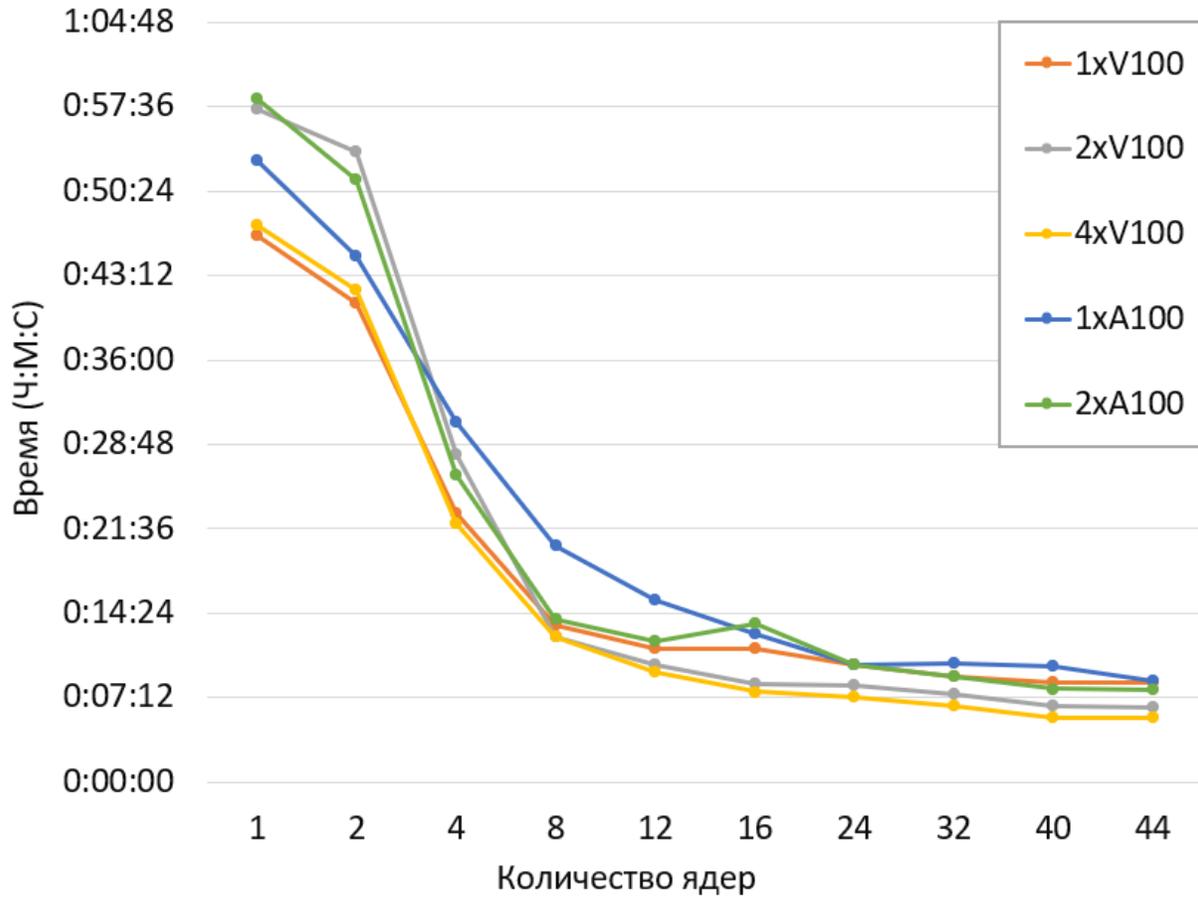
Варианты экономии

- вычислительных ресурсов: $p_i \rightarrow \min$
- ресурсов и электроэнергии: $p_i + p_i^e \rightarrow \min$
- ресурсов, финансов но с учетом срочности: $w_1(p_i + p_i^e) + w_2 t_i \rightarrow \min$

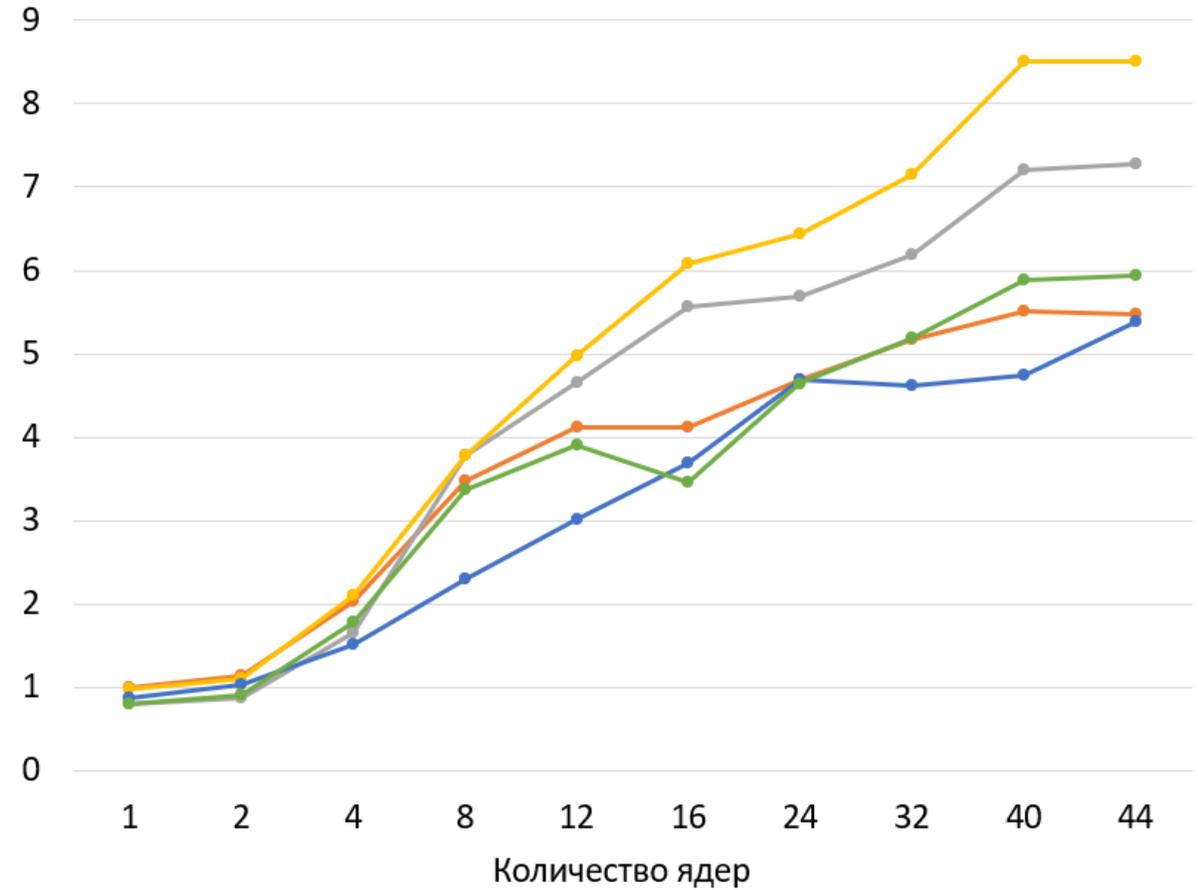
Здесь: n_{cpu}, n_{gpu} – количество CPU и GPU,
 p_{cpu} и p_{gpu} – стоимости использования 1 часа CPU-ядра и GPU,
 e_p – стоимость электроэнергии в час,
 u_{cpu}, u_{gpu} – средняя загрузка CPU и GPU соответственно $[0, 1]$,
 c_{cpu}, c_{gpu} – максимальное электропотребление для CPU и GPU (Ватт),
 t_i – время выполнения задачи (час),
 w_1, w_2 – веса, регулирующие важность критериев.



Время выполнения



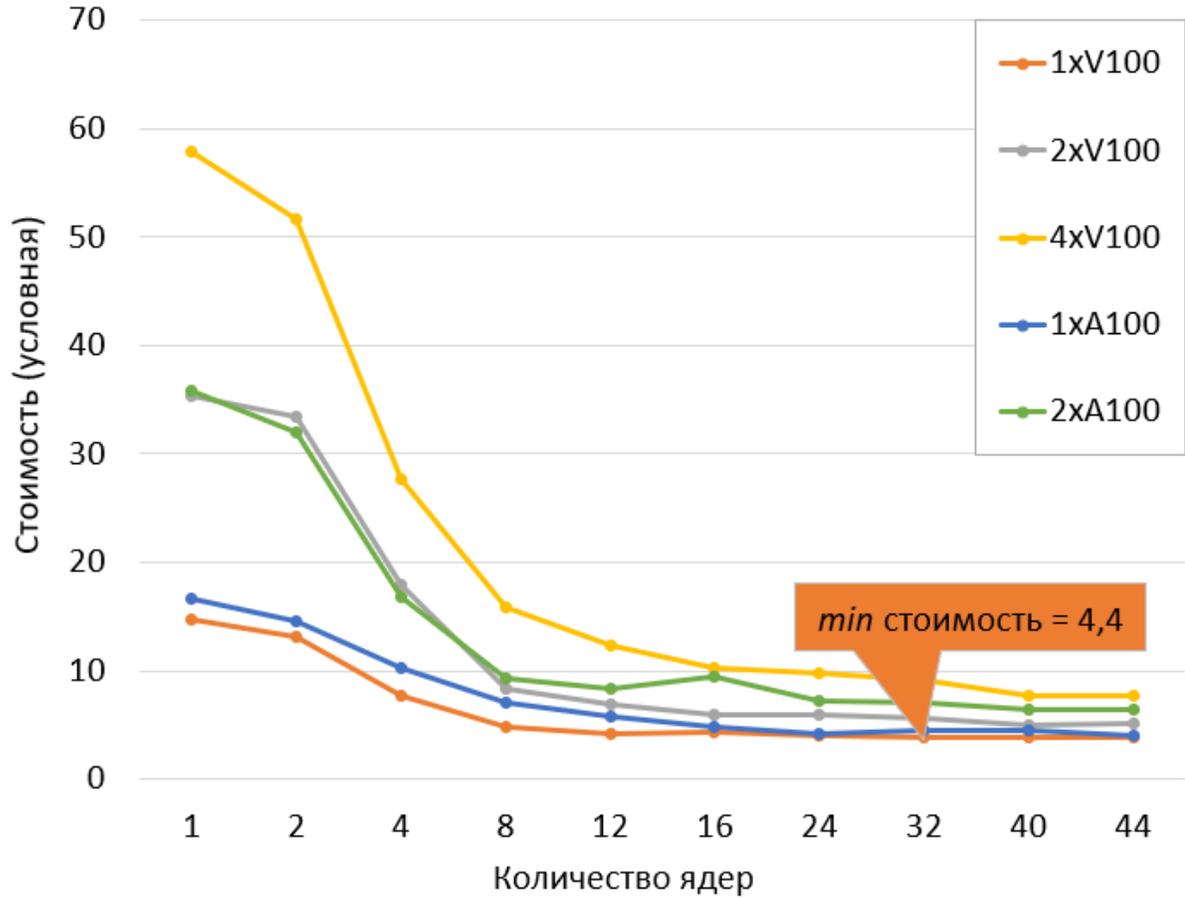
Ускорение





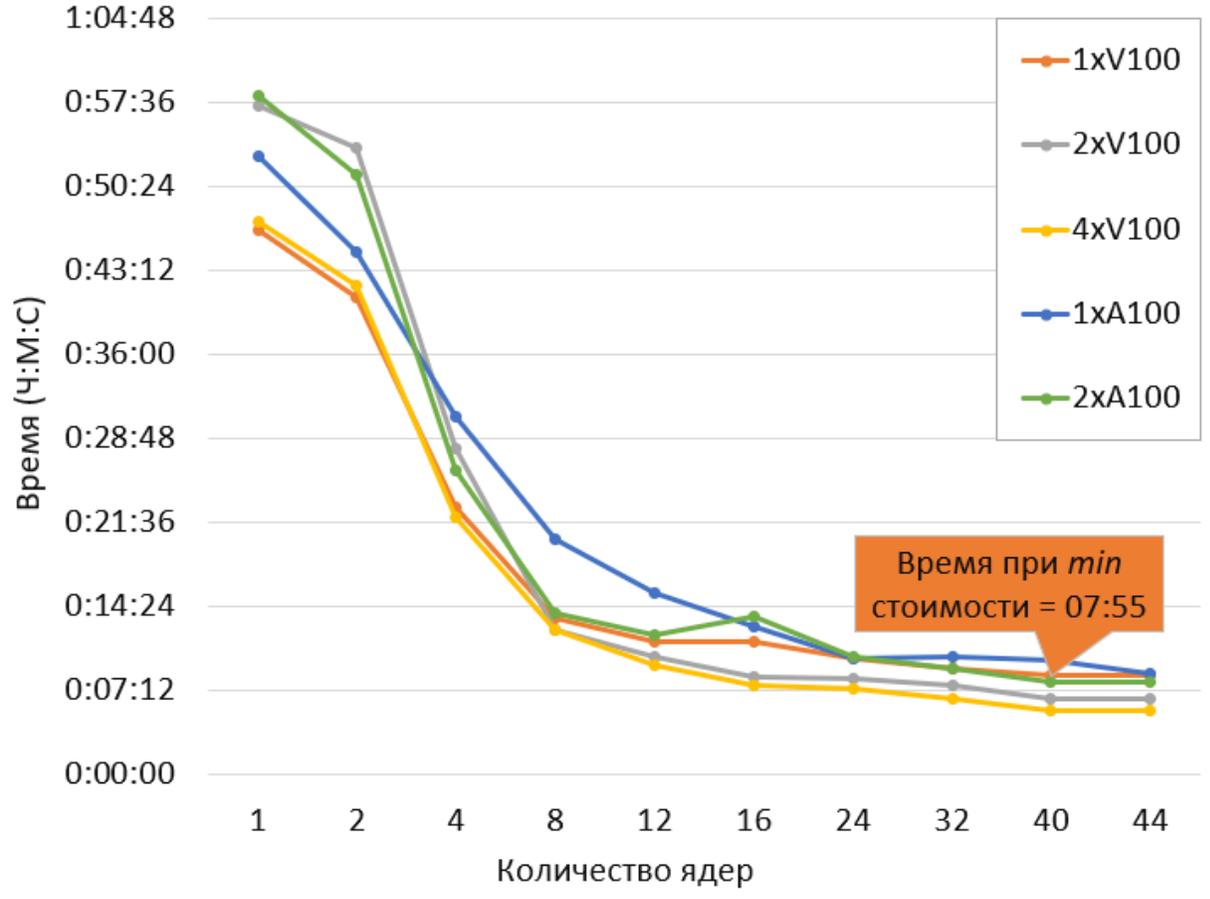
ПО GROMACS: стоимость выполнения задачи

Стоимость выполнения задачи



min стоимость = 4,4

Время выполнения



Время при min стоимости = 07:55



Рекомендации для начинающих пользователей

☰

ОТДЕЛ
СУПЕРКОМПЬЮТЕРНОГО
МОДЕЛИРОВАНИЯ

- Принципы использования суперкомпьютера >
- Вычислительные ресурсы >
- Программное обеспечение >
- Пользователям >
- Научная работа >
- Научно-технический совет >
- О нас >

Руководитель –
Костенецкий Павел
Сергеевич

GROMACS
fast, flexible & free



Gromacs - пакет программ для моделирования физико-химических процессов в молекулярной динамике. Разработан командой Германа Берендсена, работающей в отделении биофизической химии университета Гронингена. В настоящее время развивается и поддерживается усилиями энтузиастов, в число которых входят представители университета Уппсалы и королевского технологического института. Пакет предназначен для моделирования биомолекул (например, молекул белков и

липидов), имеющих много связанных взаимодействий между атомами. Обеспечивает высокую скорость расчётов для несвязанных взаимодействий.

Официальный сайт: <http://www.gromacs.org/>

Использование GPU-ускорителей (рекомендуется)

В данном примере, для расчёта будет выделено 20 ядер и 1 GPU на одном вычислительном узле.

Содержимое файла `gromacs-gpu.sbatch`:

```
#!/bin/bash
#SBATCH --job-name=gromacs           # Название задачи
#SBATCH --time=01:00:00              # Максимальное время выполнения
#SBATCH --ntasks 20                  # Количество MPI процессов
#SBATCH --nodes 1                    # Все CPU и GPU будут на одном вычислительном узле
#SBATCH --gpus 1                     # Требуемое кол-во GPU
#SBATCH --constraint="type_a|type_b" # Предпочтительные типы вычислительных узлов
#SBATCH --mail-user=ваша_почта       # Укажите ваш email для отправки уведомлений
#SBATCH --mail-type=END,FAIL          # События, требующие уведомления

module load GROMACS/2022.2            # Загрузка модуля GROMACS

srun --mpi=pmix_v3 gmx_mpi mdrun -ntomp 1 -v -s benchmark.tpr

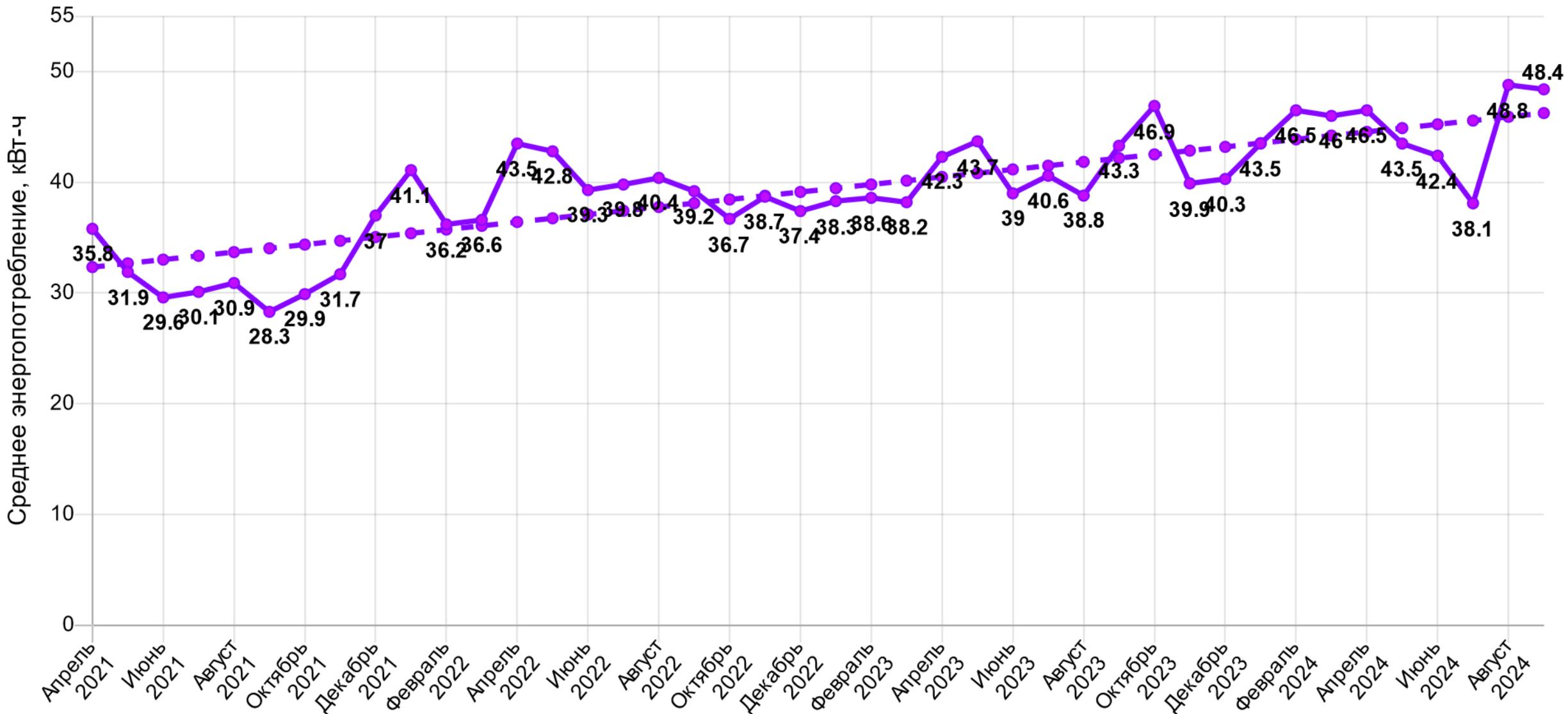
# Рекомендуемые пропорции ресурсов на CHARISMa:
# -на узлах типов A, B, C: 1 GPU и 20 ядер, 2 GPU и 40 ядер; 4 GPU и 44/48 ядер.
# -на узлах типа E: 1 GPU и 48 ядер.
# Как можно больше заполняйте память GPU. Отслеживайте загрузку памяти через: https://lk.hpc.hse.ru
```

Для постановки задачи в очередь выполните команду `sbatch gromacs-gpu.sbatch`



Энергопотребление вычислительных узлов

20





Цифровой блок НИУ ВШЭ
Отдел суперкомпьютерного моделирования

Москва,
23 сентября 2024 г.

Примеры научных проектов, выполненных на суперкомпьютере НИУ ВШЭ

Заказчик: Программа фундаментальных исследований

Исполнитель: Международная лаборатория суперкомпьютерного атомистического моделирования и многомасштабного анализа

Руководитель проекта: Стегайлов В.В.

Цель исследования:

Оптимизация и разработка параллельных алгоритмов классических и квантовых молекулярно-динамических расчетов для многомасштабного моделирования материалов. Решение научных задач в области вычислительного материаловедения.

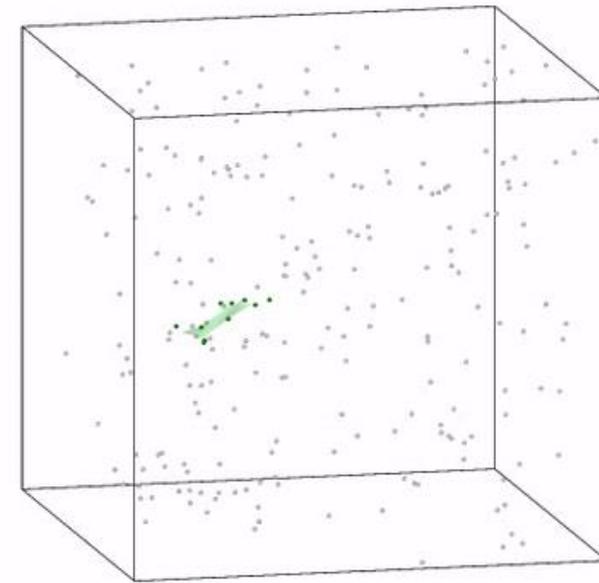
Результат:

- Разработана аналитическая модель для оптимизации алгоритма матричного умножения на нескольких ГПУ. Проведен анализ производительности алгоритма на серверах с ГПУ Nvidia и AMD, объединенных Nvlink и PCIe4.0.
- Проведены исследования фазовых превращений аморфных льдов, электронной структуры магнетита, плазменного фазового перехода в разогретом плотном водороде и процесса сборки/разворачивания ДНК.

Публикации:

1. Choi Y. R., Nikolskiy V., Stegailov V. in: *Parallel Computational Technologies: 16th International Conference, PCT 2022, Dubna, Russia, March 29–31, 2022, Revised Selected Papers*. Springer, 2022. Ch. 12. P. 158-171.
2. Belonoshko A., Smirnov G. *A Comparison of Experimental and Ab Initio Structural Data on Fe under Extreme Conditions // Metals, 2023 (Q1)*
3. Choi Y. R., Stegailov V. in: *22nd International Conference, MMST 2022, Nizhny Novgorod, Russia, November 14–17, 2022, Revised Selected Papers*. Springer, 2022. Ch. 23. P. 281-292.
4. Garkul A., Stegailov V. // *Scientific Reports*. 2022. Vol. 12. No. 1. Article 13325. **(Q1)**
5. Shutikova M. I., V. V. Stegailov. // *Journal of Physics: Cond. Matter*. 2022. Vol. 34. No. 47. Article 475701.
6. Aliev T. A., Timralieva A. A., Kurakina T. A., Katsuba K. E., Egorycheva Y. A., Dubovichenko M. V., Kuttyrev M.A., Shilovskikh V.V., Orekhov N., Kondratyuk N., Semenov S.N., Kolpashchikov D.M., Skorb E. V. // *Nano Select*. 2022. Vol. 3. No. 11. P. 1526-1536.
7. Fedorov I.D., V. V. Stegailov. // *ChemPhysChem*. 2023. V. 24. No. 6. Article e202200730.

Transformation HDA -> LDA during decompression -0.2 -> -0.5 GPa



Density 1.14 -> 0.98 g/cm³ T = 77 K Time 5840 ps.

Заказчик: Программа фундаментальных исследований НИУ ВШЭ
Исполнитель: Департамент электронной инженерии МИЭМ НИУ ВШЭ
Менеджер проекта: А.С. Васенко

Цель исследования:

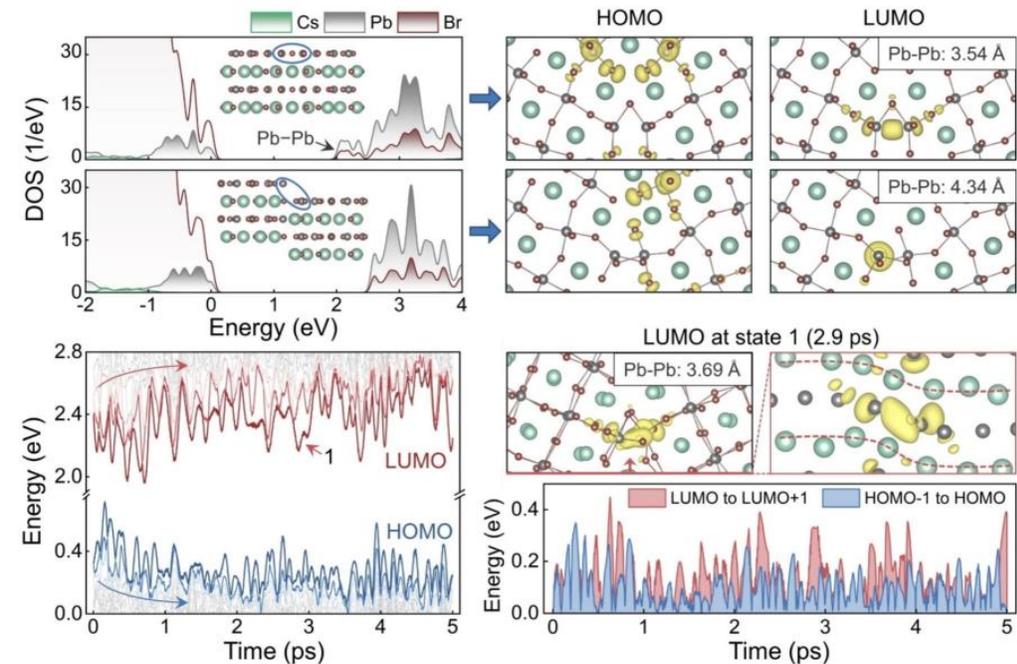
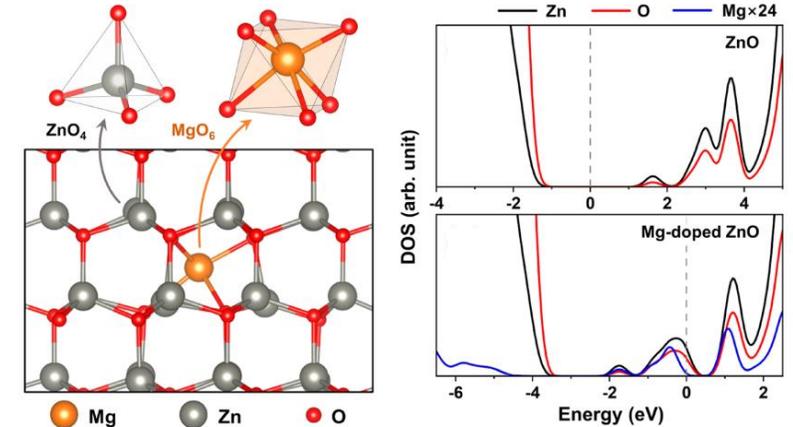
Изучите электронную структуру сверхпроводников и других функциональных материалов с помощью квантово-механических расчетов и проанализируйте их новые свойства.

Результаты:

- Создана модель ZnO, легированного Mg, и с помощью метода функционала плотности (DFT) рассчитана плотность состояний (DOS) как чистого, так и легированного ZnO. Примеси Mg в ZnO служат донорами электронов для заполнения электронов в зоне проводимости, изменяя краевые состояния зоны и уменьшая ширину запрещенной зоны.
- Граница зерна (120) в CsPbBr₃ приводит к локализованным электронным состояниям ниже зоны проводимости. Спонтанное скольжение двух зерен устраняет эти локализованные состояния, но их тепловые флуктуации могут воспроизвести такие состояния и создать уровень энергии глубоко в запрещенной зоне.

Опубликованные статьи:

- Abinash Das, Dongyu Liu, Riu Riu Wary, Andrey S. Vasenko, Oleg V. Prezhdo, and Ranjith G. Nair. Enhancement of Photocatalytic and Photoelectrochemical Performance of ZnO by Mg Doping: Experimental and Density Functional Theory Insights// J. Phys. Chem. Lett., 2023 (Q1)
- Dongyu Liu, Yifan Wu, Andrey S. Vasenko, Oleg V. Prezhdo. Grain boundary sliding and distortion on a nanosecond timescale induce trap states in CsPbBr₃: ab initio investigation with machine learning force field // Nanoscale, 2023 (Q1)





Заказчик: HSE

Исполнитель: LAMBDA HSE

Руководитель проекта: Лазарев М.В.

Цель исследования:

Мы предлагаем подход машинного обучения для быстрой оценки свойств 2D-материала с учетом структуры решетки и конфигурации дефектов. Метод предлагает способ представления конфигурации 2D-материалов с дефектами, позволяющий нейронной сети быстро и точно обучаться. Мы сравниваем нашу методологию с современными подходами и демонстрируем снижение ошибки прогнозирования энергии как минимум в 3,7 раза. Кроме того, наш подход на порядок более эффективен с точки зрения ресурсов, чем его конкуренты, как в части обучения, так и в части вывода.

Результат:

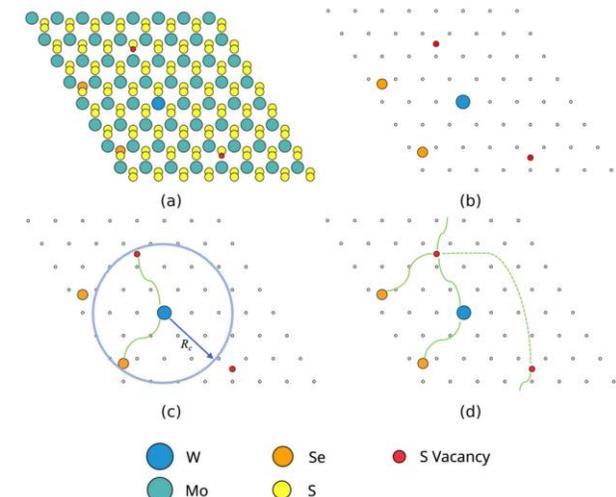
- Свойства материалов предсказываются GNN

Публикации:

1. Kazeev, Nikita, Abdalaziz Rashid Al-Maeeni, Ignat Romanov, Maxim Faleev, Ruslan Lukin, Alexander Tormasov, A. H. Castro Neto, Kostya S. Novoselov, Pengru Huang, and Andrey Ustyuzhanin. "Sparse representation for machine learning the properties of defects in 2D materials." *npj Computational Materials* 9, no. 1 (2023): 113. **(Q1)**

Formation energy per site MAE, meV; lower is better						
Material	Density	SchNet	GemNet	MEGNet (Full)	CatBoost	Sparse (MEGNet)
Combined	Both	631 ± 31	483 ± 91	158 ± 47	164 ± 5	43 ± 5
BP	High	2088 ± 72	1490 ± 429	198 ± 211	382 ± 30	80 ± 10
GaSe	High	245 ± 12	230 ± 41	107 ± 25	103 ± 4	48 ± 7
InSe	High	268 ± 19	247 ± 26	95 ± 27	137 ± 5	35 ± 2
MoS ₂	High	321 ± 100	535 ± 206	136 ± 22	136 ± 5	23 ± 5
WSe ₂	High	536 ± 123	575 ± 181	112 ± 33	162 ± 6	23 ± 4
h-BN	High	1442 ± 68	697 ± 315	496 ± 229	363 ± 17	127 ± 16
MoS ₂	Low	65 ± 5	44 ± 14	58 ± 11	12.6 ± 0.4	4 ± 1
WSe ₂	Low	85 ± 22	42 ± 9	65 ± 16	16.3 ± 0.8	6 ± 1

HOMO – LUMO gap MAE, meV; lower is better						
Material	Density	SchNet	GemNet	MEGNet (Full)	CatBoost	Sparse (MEGNet)
Combined	Both	224 ± 111	166 ± 42	112 ± 3	117 ± 1	112 ± 3
BP	High	208 ± 20	176 ± 10	170 ± 4	174 ± 2	187 ± 9
GaSe	High	309 ± 83	196 ± 11	178 ± 8	173 ± 4	194 ± 11
InSe	High	214 ± 69	178 ± 22	156 ± 7	155 ± 1	167 ± 15
MoS ₂	High	204 ± 121	174 ± 111	54 ± 4	71 ± 4	39 ± 4
WSe ₂	High	186 ± 177	268 ± 182	47 ± 3	106 ± 6	38 ± 4
h-BN	High	244 ± 24	227 ± 6	233 ± 4	208 ± 3	260 ± 14
MoS ₂	Low	187 ± 180	46 ± 42	30 ± 2	26.7 ± 0.8	5.7 ± 0.2
WSe ₂	Low	236 ± 224	64 ± 46	32 ± 5	18.3 ± 0.6	8.1 ± 0.6



Руководитель проекта: Баховадинов М.С.

Цель исследования:

Исследование модели одномерных фермионов с парными прыжками в беспорядке методом DMRG в HPC CHARISMA

Результат:

Изучаются эффекты беспорядка в одномерной системе фермионов с одночастичными и парными перескоками. С помощью численных расчетов получена фазовая диаграмма модели в чистом случае. Показано, что в режиме сильных парных прыжков член беспорядка irrelevantна в смысле ренормгруппы и электронная жидкость с алгебраически затухающими корреляциями при конечном, но малом беспорядке, сохраняется. При больших же амплитудах беспорядка система переходит в фазу локализации по механизму Березинского-Костерлица-Таулесса. Фазовая диаграмма неупорядоченной модели получена на основе спада пространственных корреляторов и усредненного по беспорядку параметра жидкости Латтингера К. Найденный переход происходит по сценарию Джиамарчи-Шульца.

Публикации:

1. Bahovadinov M.S., Kurlov D.V., Matveenko S.I., Altshuler B.L., Shlyapnikov G.V. Many-body localization transition in a frustrated XY chain // *Phys. Rev. B*, 2022 (Q1)
2. Bahovadinov M.S., Buijsman W., Fedorov A.K., Gritsev V., Kurlov D.V. Many-body localization of Z3 Fock parafermions // *Phys. Rev. B*, 2022 (Q1)

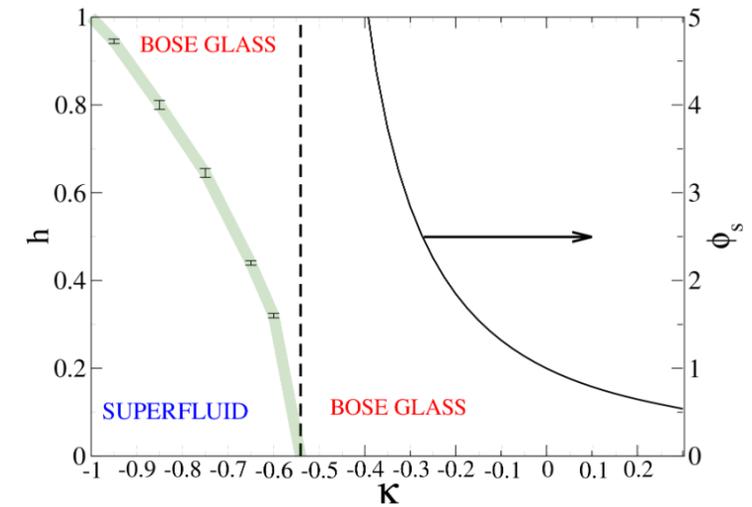
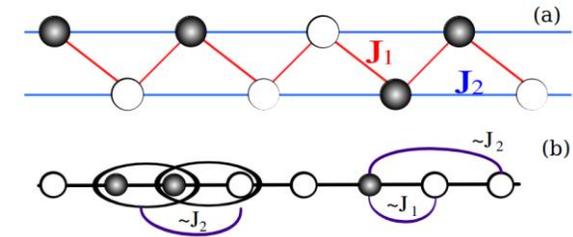


FIG. 7. Phase diagram of the model in the $h-\kappa$ plane and the exponent ϕ as function of κ (solid curve). The dashed vertical line at $\kappa \approx -0.54$ separates the exhibited two regimes: for $-1 \leq \kappa \lesssim -0.54$ and $h < h_c$ the SF phase persists. At larger disorder strength $h > h_c$ one enters the BG phase. In the second regime for $\kappa > -0.54$ arbitrarily weak disorder drives the system into the BG phase. The solid strip serves as eye-guide.

Заказчик: Программа фундаментальных исследований

Исполнитель: Международная лаборатория суперкомпьютерного атомистического моделирования и многомасштабного анализа

Руководитель проекта: Писарев В.В.

Цель исследования:

Проект направлен на исследование динамики атомов в конденсированных фазах методом атомистического моделирования и выявления связи особенностей динамики с макроскопическими свойствами

Результат:

- Обнаружено, что парная энтропия как структурный параметр может характеризовать стеклование в разнородных классах веществ
- При использовании квантовой молекулярной динамики удалось объяснить механизм фазового перехода первого рода в соли Фриделя.
- Раскрыты пути адсорбции нитрит-иона на поверхности цементного минерала

Публикации:

1. Tararushkin E.V., Pisarev V.V., Kalinichev A.G. Interaction of Nitrite Ions with Hydrated Portlandite Surfaces: Atomistic Computer Simulation Study // *Materials*, 2023, 16(14) P. 5026. **(Q2)**
2. Glushak A.A., Smirnov G.S. Partial proton ordering during phase transition in Friedel's salt// *Applied Clay Science*. 2023. V. 243. P. 107072. **(Q1)**
3. Tararushkin E., Pisarev V., Kalinichev A.G. Equation of State, Compressibility, and Vibrational Properties of Brucite over Wide Pressure and Temperature Ranges: Atomistic Computer Simulations with the Modified ClayFF Classical Force Field // *Minerals*. 2023. V. 13. P. 408. **(Q2)**

