

УДК 004.22, 004.272.4, 004.62

АНАЛИЗ ДАННЫХ О ЗАГРУЗКЕ ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ ПЛАТФОРМ ПОЛЬЗОВАТЕЛЬСКИМИ ЗАДАЧАМИ НА ПРИМЕРЕ ГЕТЕРОГЕННОЙ ВЫЧИСЛИТЕЛЬНОЙ ПЛАТФОРМЫ HYBRILIT

Полегаева Екатерина Игоревна¹, Пряхина Дарья Игоревна²,
Стрельцова Оксана Ивановна³

¹Студент;

ГБОУ ВО МО «Университет «Дубна»,

Россия, 141980, Московская обл., г. Дубна, ул. Университетская, д. 19;

e-mail: robin_goul@mail.ru.

²Научный сотрудник;

Объединенный институт ядерных исследований,

Россия, 141980, Московская обл., г. Дубна, ул. Жолио-Кюри, д. 6;

Старший преподаватель;

ГБОУ ВО МО «Университет «Дубна»,

Россия, 141980, Московская обл., г. Дубна, ул. Университетская, д. 19;

e-mail: pryahinad@jinr.ru.

³Старший научный сотрудник;

Объединенный институт ядерных исследований,

Россия, 141980, Московская обл., г. Дубна, ул. Жолио-Кюри, д. 6;

Кандидат физико-математических наук, доцент;

ГБОУ ВО МО «Университет «Дубна»,

Россия, 141980, Московская обл., г. Дубна, ул. Университетская, д. 19;

e-mail: strel@jinr.ru

Актуальность исследования, приведенного в данной статье, заключается в возможности агрегировать статистическую информацию об использовании ресурсов платформы различными группами пользователей. В статье проводится анализ данных по использованию ресурсов учебно-тестового полигона платформы HybriLIT, на основе которого можно построить модели для прогнозирования дальнейшей загруженности платформы с целью рационального распределения имеющихся вычислительных ресурсов и системы хранения данных, а также повышения эффективности их использования. Цель работы – представить исследования, направленные на интеллектуальный анализ данных используемых ресурсов при запуске задач различными группами пользователей и времени их выполнения на учебно-тестовом полигоне платформы HybriLIT. Гетерогенная вычислительная платформа HybriLIT, состоящая из учебно-тестового полигона и суперкомпьютера «Говорун», является частью многофункционального информационно-вычислительного комплекса Лаборатории информационных технологий им. М.Г. Меццержкова Объединенного института ядерных исследований. Платформа имеет гетерогенную структуру вычислительных узлов и позволяет запускать параллельные приложения для проведения расчетов на различных вычислительных архитектурах. Сводная информация о запущенных на HybriLIT задачах, автоматически записывающаяся в базу данных менеджером и планировщиком ресурсов SLURM, представляет интерес для анализа. В результате был проведен глубокий анализ данных по нескольким критериям для каждого ресурса и для каждой группы пользователей учебно-тестового полигона гетерогенной платформы HybriLIT.

Ключевые слова: интеллектуальный анализ данных, высокопроизводительные вычислительные платформы, эффективное использование ресурсов.

Для цитирования:

Полегаева Е. И., Пряхина Д. И., Стрельцова О. И. Анализ данных о загрузке высокопроизводительных платформ пользовательскими задачами на примере гетерогенной вычислительной платформы HybriLIT // Системный анализ в науке и образовании: сетевое научное издание. 2021. № 2. С. 67–76. URL : <http://sanse.ru/download/437>.

ANALYSIS OF DATA ON THE LOAD OF HIGH-PERFORMANCE PLATFORMS BY USER TASKS ON THE EXAMPLE OF A HETEROGENEOUS COMPUTING PLATFORM HYBRILIT

Polegaeva Ekaterina I.¹, Priakhina Daria I.², Streltsova Oksana I.³

¹Student;

Dubna State University,

19 Universitetskaya Str., Dubna, Moscow region, 141980, Russia;

e-mail: robin_goul@mail.ru.

²Researcher;

Joint Institute for Nuclear Research,

6 Joliot-Curie Str., Dubna, Moscow region, 141980, Russia;

Senior teacher;

Dubna State University,

19 Universitetskaya Str., Dubna, Moscow region, 141980, Russia;

e-mail: pryahinad@jinr.ru.

³Senior researcher;

Joint Institute for Nuclear Research,

6 Joliot-Curie Str., Dubna, Moscow region, 141980, Russia;

Candidate of physical and mathematical sciences, associate professor;

Dubna State University,

19 Universitetskaya Str., Dubna, Moscow region, 141980, Russia;

e-mail: strel@jinr.ru.

The relevance of the research presented in this article lies in the ability to aggregate statistical information about the use of platform resources by various user groups. The article analyzes the data on the use of the resources of the training and testing ground of the HybriLIT platform, on the basis of which it is possible to build models for predicting the further workload of the platform in order to rationally allocate the available computing resources and the data storage system, as well as increase the efficiency of their use. The purpose of the work is to present the research aimed at data mining of the resources used when launching tasks by various groups of users and the time of their execution on the training and testing ground of the HybriLIT platform. The heterogeneous HybriLIT computing platform, consisting of a training and testing ground and the Govorun supercomputer, is part of the multifunctional information and computing complex of the Meshcheryakov Information Technology Laboratory of the Joint Institute for Nuclear Research. The platform has a heterogeneous structure of computing nodes and allows you to run parallel applications for performing calculations on various computing architectures. The summary information about the tasks running on HybriLIT, which is automatically recorded in the database by the SLURM resource manager and scheduler, is of interest for analysis. As a result, a deep analysis of data was carried out according to several criteria for each resource and for each group of users of the training and testing ground of the heterogeneous HybriLIT platform.

Keywords: data mining, high performance computing platforms, efficient use of resources.

For citation:

Polegaeva E., Priakhina D., Streltsova O. Analysis of data on the loading of high-performance platforms by user tasks on the example of the heterogeneous computing platform HybriLIT, 2021;(2):67–76(In Russ). Available from: <http://sanse.ru/download/437>.

Введение

Современные тенденции развития высокопроизводительных платформ (суперкомпьютеров) связаны не только с построением гетерогенных вычислительных платформ, в вычислительное поле которых включены ускорители вычислений, но и с развитием специализированных систем хранения данных [1]. Суперкомпьютеры, как правило, это системы общего назначения и предназначены для решения задач разных типов, связанных и с массивно-параллельными расчетами, и обработкой дан-

ных. К таким платформам общего назначения относится и Гетерогенная вычислительная платформа *HybriLIT*, которая является частью Многофункционального информационно-вычислительного комплекса Лаборатории информационных технологий им. М.Г. Мещерякова Объединенного института ядерных исследований (ЛИТ ОИЯИ) [2, 3] – базовой установки ОИЯИ. Основным назначением платформы *HybriLIT* является обеспечения проведения массивно-параллельных расчетов для решения задач теоретической физики, обработки и анализа данных для экспериментов в области ядерной физики Проблемно-тематического плана ОИЯИ. Платформа *HybriLIT* состоит из учебно-тестового полигона и суперкомпьютера «Говорун», объединенных единой программно-информационной средой [3]. Результаты, представленные в статье, получены на данных о выполненных задачах на учебно-тестовом полигоне.

Высокопроизводительные платформы в основном не являются системами с интерактивной организацией расчетов. Вычисления на них проходят в режиме очередей, и пользовательские задачи выполняются через систему управления заданиями *SLURM* [4].

На *HybriLIT* в планировщике заданий *SLURM* настраиваются очереди для распределения ресурсов между различными группами пользователей и различными ограничениями на время выполнения задач, например, в очереди «*scri*» доступны вычислительные узлы с центральными процессорами, в очереди «*tut*» установлено ограничение на время выполнения расчетов два часа и доступны, как центральные процессы и графические ускорители для проведения расчетов студентами и участниками учебных курсов. На платформе в качестве единицы для возможности резервирования ресурсов являются логические ядра *CPU*, отдельно указываются количество необходимых ускорителей вычислений и объем оперативной памяти. Формирование очередей и другие настройки параметров проводятся исходя из анализа потребностей пользователей. Однако, представляет большой интерес в систематизации данных об использованных ресурсах за различные периоды времени и анализ об использовании как различных типов вычислительных архитектур, так и анализ использования ресурсов различными группами пользователей с целью повышения эффективности их использования.

В программно-информационную среду платформы *HybriLIT* внедрены несколько сервисов мониторинга, позволяющих в интерактивном режиме отслеживать состояния загруженности узлов как по вычислительным компонентам, так и по использованию памяти узла. Агрегирование информации о проведенных расчетах и анализ таких данных за различные промежутки времени представляет интерес, поскольку позволит анализировать риски перегруженности оборудования и прогнозировать их дальнейшее использование. В работе представлены исследования, направленные на интеллектуальный анализ данных используемых ресурсов при запуске задач различными группами пользователей и времени их выполнения на учебно-тестовом полигоне платформы *HybriLIT*.

Работа выполнена на ресурсах *HybriLIT*, на развернутой экосистеме для решения задач машинного и глубокого обучения. В состав экосистемы входит *Jupyter Notebook* – веб-приложение с открытым исходным кодом (режим доступа: <https://jinr.jhub.com/>). Анализ данных проводился на языке *Python* с помощью таких библиотек, как *pandas* [5], *numpy* [6], *datetime* [7], *matplotlib* [8], *sklearn* [9].

1. Платформа *HybriLIT*

Гетерогенная платформа *HybriLIT* находится под управлением операционной системы *Scientific Linux 7.5*, с планировщиком задач *SLURM* и установленным программным обеспечением. *SLURM* – высоко масштабируемый отказоустойчивый менеджер кластеров и планировщик заданий с открытым исходным кодом, который обеспечивает три основные функции:

- выделяет эксклюзивный и/или неэксклюзивный доступ к ресурсам для пользователей в течение некоторого периода времени;
- обеспечивает основу для запуска, работы и мониторинга работы на множестве выделенных узлов;
- поддерживает очередь ожидающих заданий и управляет общей загрузкой ресурсов в процессе выполнения работы.

Запуск задачи осуществляется с помощью постановки ее в очередь на выполнение. На текущий момент *HybriLIT* содержит 6 очередей, далее приведены некоторые из них [10]:

- *interactive* – включает один вычислительный узел с двумя процессорами *Intel Xeon E5-2695 v2 12-cores*, одной видеокартой *NVIDIA Tesla K20X*, один сопроцессор *Intel Xeon Coprocessor 5110P*. Очередь используется по умолчанию, в частности для запуска тестовых программ. Время расчетов для этой очереди ограничено и составляет 1 день;
- *cpu* – включает 7 вычислительных узлов с двумя процессорами *Intel Xeon E5-2695 v2 12-cores* на каждом. Очередь применяется для запуска приложений, использующих для расчетов центральные процессоры;
- *gpu* – очередь включает 6 вычислительных узлов: три видеокарты *NVIDIA Tesla K40 (Atlas)* на каждом, три вычислительных узла с двумя видеокартами *NVIDIA Tesla K80* на каждом. Данная очередь предназначена для запуска приложений, использующих графические ускорители для расчетов;
- *long* – очередь включает два вычислительных узла с видеокартой *NVIDIA Tesla*. Очередь используется для запуска приложений, требующих длительных (до 14 дней) вычислений.

2. Подготовка данных

Исходными данными для проведения анализа являлись данные из базы данных (БД) планировщика *SLURM*. Для анализа использовалась копия БД, которая содержит информацию о запусках задач за период с апреля 2018 года по август 2020 года. БД состоит из 47 таблиц, которые можно распределить на следующие категории: данные по использованию учебно-тестового полигона, данные по использованию суперкомпьютера «Говорун» и общие данные.

Был проведен полный анализ имеющихся в БД таблиц. Для дальнейшей работы из таблиц БД были выбраны данные только по использованию учебно-тестового полигона *HybriLIT* (см. рис. 1). Для удобства данные были собраны в *DataFrame*, таблицу, построенную с помощью библиотеки *pandas* [5], предоставляющей широкий функционал для работы с данными на языке *Python*. Обработка данных в таком виде не требует большого количества вычислительных и временных ресурсов.

group_name	user_name	uid	user_name_count	nodelist	nodecnt	partition	proc_cnt	starttime	duration	account	gres_req	
0	hybrilit	akotov	11126	1	dgx01	1	dgx	2	1523087061	383605	None	gpu:1
1	hybrilit	akotov	11126	1	dgx01	1	dgx	2	1523087061	383605	bltp	gpu:1
2	hybrilit	akotov	11126	1	dgx01	1	dgx	2	1523087061	383605	bltp	
3	hybrilit	akotov	11126	1	dgx01	1	dgx	2	1523087061	383605	dgx	gpu:1
4	hybrilit	akotov	11126	1	dgx01	1	dgx	2	1523087061	383605	dgx	
...
366583	hybrilit	mkompan	11138	1	blade01	1	interactive	20	1597900546	4221	None	
366584	hybrilit	mkompan	11138	1	blade06	1	tut	20	1597900605	5504	None	
366585	hybrilit	mkompan	11138	1	blade06	1	tut	20	1597900626	5483	None	
366586	hybrilit	mkompan	11138	1	blade01	1	interactive	20	1597904690	1718	None	
366587	hybrilit	mkompan	11138	1	blade01	1	interactive	20	1597904767	1641	None	

Рис. 1. Пример подготовленных данных

3. Анализ использования ресурсов гетерогенной вычислительной платформы HybriLIT

3.1. Использование ресурсов различными группами пользователей

На платформе используются следующие узлы: *blade[01-10]*; *dg [01-05]*; *volta[01]*;

Далее представлены графики, отражающие результаты проведенного анализа по различным критериям использования платформы *HybriLIT* по каждой группе пользователей и по каждому ресурсу. Пользователями платформы могут быть либо сотрудники ОИЯИ, которые разделены на группы по критерию принадлежности к определенной лаборатории, либо студенты, в частности университета

«Дубна», либо сотрудники других организаций России или государств-членов ОИЯИ, которые тоже делятся на соответствующие группы.

3.1.1. Общее количество используемых на платформе ресурсов

На рисунке 2 представлен график, отражающий общее количество используемых ресурсов платформы *HybriLIT*, по которому можно сделать вывод о типе самых востребованных ресурсов. В дальнейшем можно будет предположить, какие ресурсы с большей вероятностью потребуются добавить в вычислительный кластер.

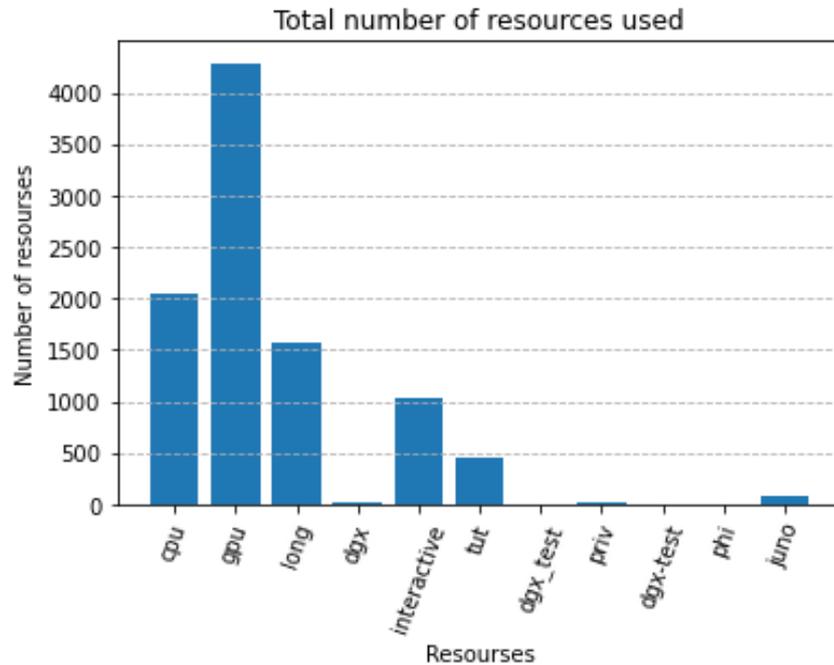


Рис. 2. Общее число всех используемых ресурсов

3.1.2. Количество запущенных задач на различных узлах

На рисунке 3 представлена гистограмма, которая отражает количество задач, запущенных на различных узлах всеми пользователями. Можно сделать вывод, что реже задачи запускаются на узлах типа *dgx*. Самым загруженным узлом является *blade[06]*.

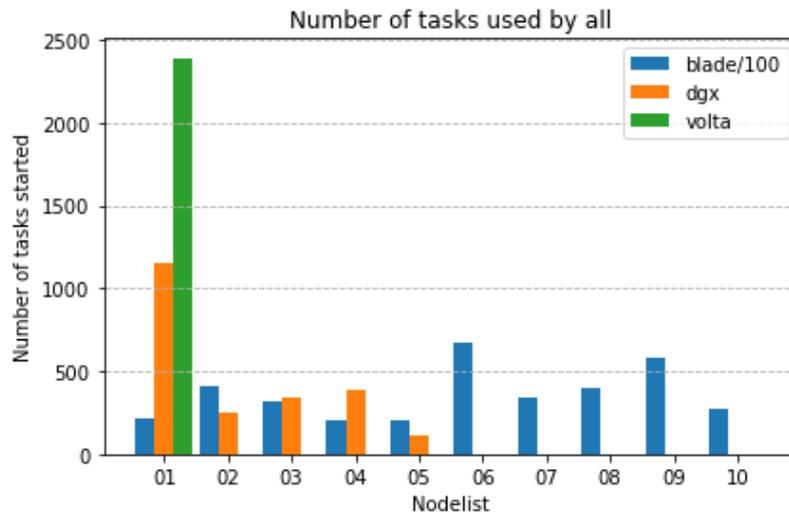


Рис. 3. Число задач, запущенных всеми пользователями на каждом узле

Проведенный по этому критерию анализ позволил получить представление о загруженности узлов. На основе полученных результатов может быть принято решение о перераспределении нагрузки на вычислительные ресурсы.

3.1.3. Количество используемых логических ядер на различных узлах

На рисунке 4 (а) отражено, сколько логических ядер используют студенты, запуская задачи на *cpu* ресурсах. На нем видно, что на *blade*[08] используется несопоставимо малое количество логических ядер, а на *blade*[10] логические ядра не используются вовсе.

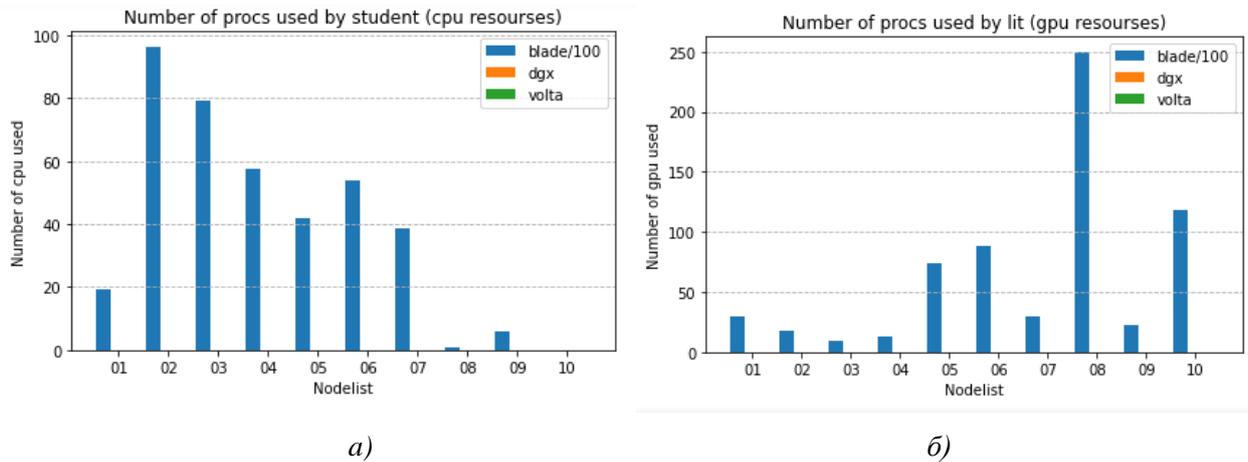


Рис. 4. Число логических ядер, используемых на каждом узле (а) студентами на *cpu* ресурсе (б) пользователями из группы Лаборатории информационных технологий на *gpu* ресурсе

Подобный анализ был проведен для всех групп пользователей, в том числе на рисунке 4 (б) отражено количество логических ядер, используемых пользователями Лаборатории информационных технологий на графических ресурсах. По результатам этого анализа получено общее и детальное представление, сколько логических ядер было использовано за весь период выгрузки данных.

3.1.4. Длительность выполнения задач, запущенных на узлах

На рисунке 5 представлен график длительности задач, запущенных всеми пользователями на каждом узле. Если на узлах *blade* запускалось большее количество задач, а на *dgx* — меньшее (см. п. 3.1.1), то благодаря анализу по этому критерию можно понять, что длительность выполнения имеет обратную картину. На узлах *dgx* запускаются более длительные задачи, а самые продолжительные по времени выполнения задачи запускаются на узле *dgx*[02]. Среди непродолжительных задач, запущенных на узлах *blade*, самые быстро выполнимые задачи запускаются на узле *blade*[08].

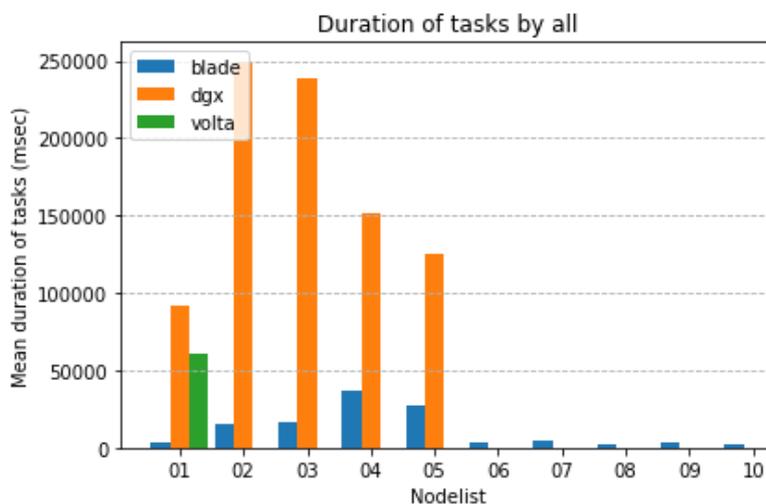


Рис. 5. Длительность выполнения задач, запущенных всеми пользователями на каждом узле

3.1.5. Длительность выполнения задач, запущенных каждой группой пользователей

Анализ по этому критерию позволяет оценить, на сколько длительные задачи запускаются каждой группой пользователей. На рисунке 6 можно увидеть, что самые длительные задачи запускаются пользователями из группы Лаборатории нейтронной физики им. И.М. Франка. Благодаря результатам такого анализа можно более грамотно распределить задачи в очередях, оценив, требуемое время на выполнение задач, запущенных каждой группой пользователей.

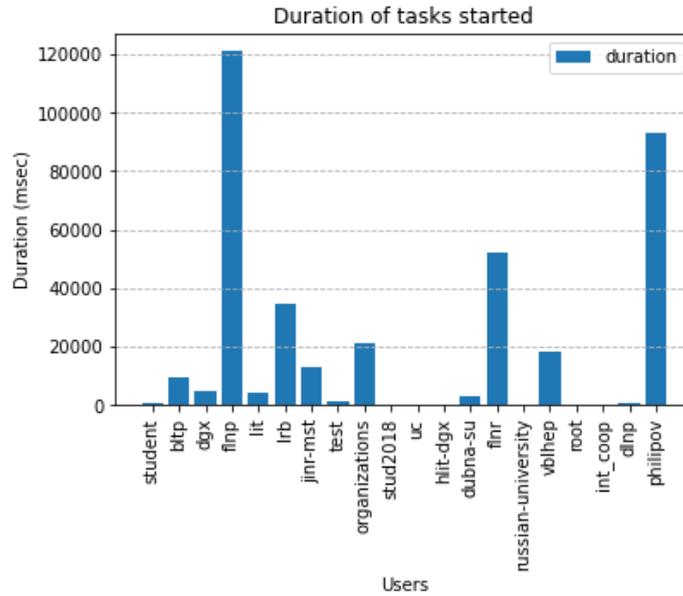


Рис. 6. Длительность выполнения задач, запущенных каждой группой пользователей

3.2.2. Использование ресурсов в течение заданного периода времени

3.2.1. Число вычислительных узлов и логических ядер, используемых за весь период времени

Анализ по этому критерию нужен, чтобы проследить загруженность всех ресурсов на протяжении всего периода имеющейся выгрузки данных. На рисунке 7 изображено количество вычислительных узлов и логических ядер, используемых всеми пользователями в каждый месяц. На графике видно, что увеличение использования логических процессоров (*proc_cnt*) совпадает с увеличением использования вычислительных узлов (*nodecnt*).

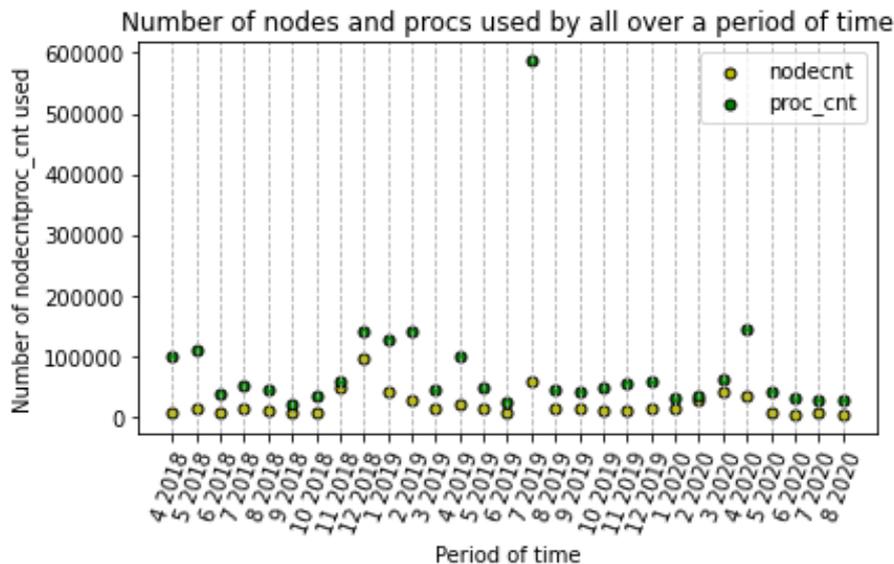


Рис. 7. Число узлов и логических ядер, используемых всеми пользователями за весь период времени

3.2.2. Число логических ядер, используемых при запуске задач на *cpu* и *gpu* ресурсах за весь период времени

Этот анализ был проведен с целью выявления тенденции к использованию *gpu* ресурсов. На рисунке 8 можно наблюдать, что со временем разница в значениях между *cpu* и *gpu* ресурсами сокращается, что можно считать проявлением тенденции к использованию *gpu* ресурсов. Результат анализа дает возможность построить линии тренда и предсказать дальнейшее использование *cpu* и *gpu* ресурсов. Это позволит принимать решение о закупке новых ресурсов при необходимости.

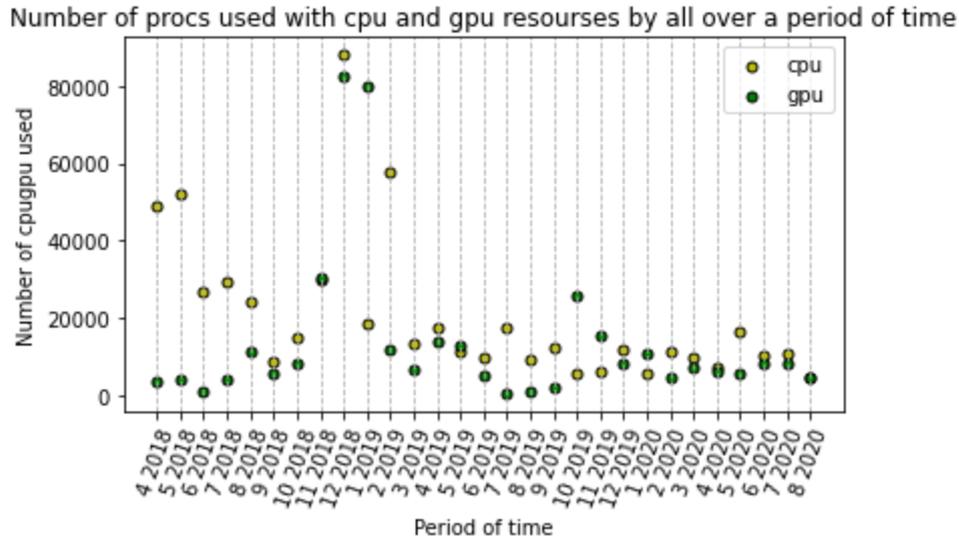


Рис. 8. Число *cpu* и *gpu* ресурсов, используемых всеми пользователями за весь период времени

3.2.3. Число запущенных задач за весь период времени

На рисунке 9 видно, что самое большое число задач было запущено в декабре 2018 года. Результаты данного анализа можно использовать для построения прогнозов по активности пользователей в тот или иной промежуток времени.

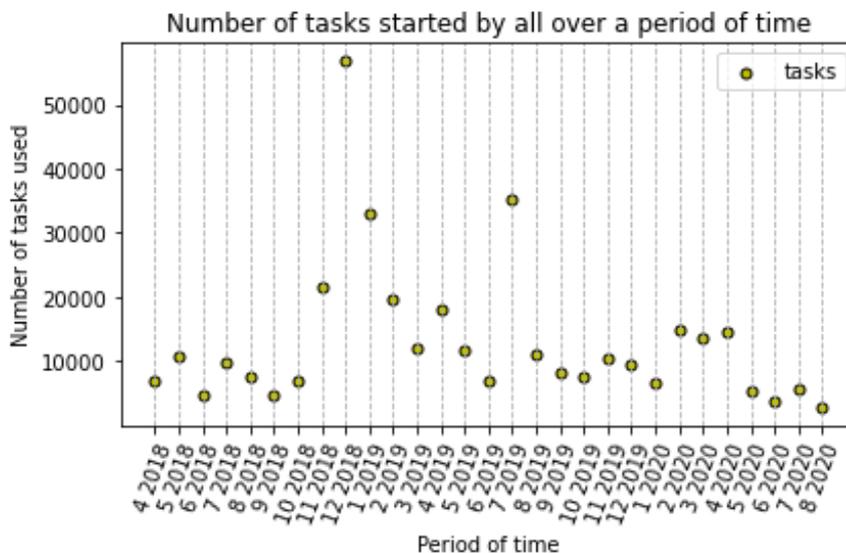


Рис. 9. Число задач, запущенных всеми пользователями за весь период времени

3.2.4. Длительность задач, запущенных за весь период времени

Анализ данных по этому критерию позволяет увидеть, какая средняя длительность запускаемых задач была в каждом месяце. На рисунке 10 видно, что самые быстро исполнимые задачи были запущены в декабре 2018 года. При этом в этот месяц запускалось больше всего задач (см. рис. 9). Из этого можно сделать вывод, что в декабре 2018 года было запущено большое число коротких задач.

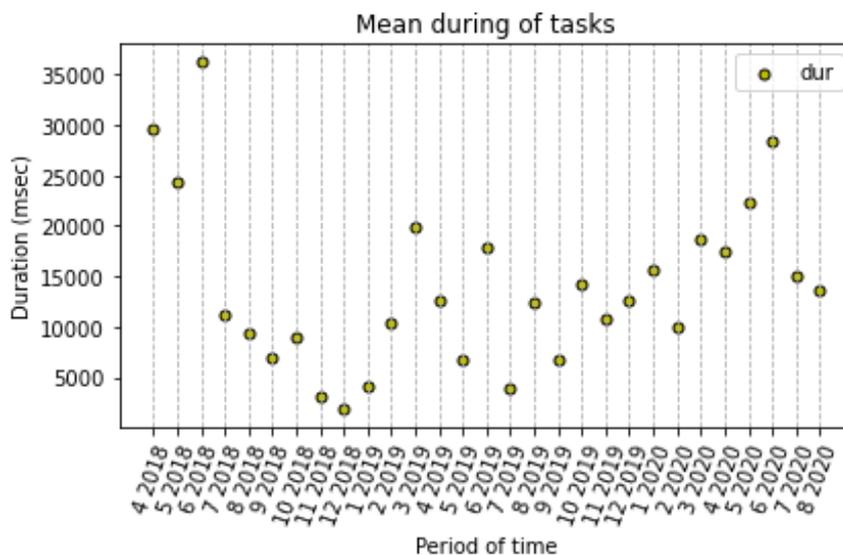


Рис. 10. Длительность задач, запущенных всеми пользователями за весь период времени

Заключение

В результате был проведен глубокий анализ данных по нескольким критериям для каждого ресурса и для каждой группы пользователей учебно-тестового полигона гетерогенной платформы *HybriLIT*. Актуальность данного исследования заключается в возможности агрегировать статистическую информацию об использовании ресурсов платформы различными группами пользователей с целью прогнозирования дальнейшей загруженности платформы на основе моделей, построенных на алгоритмах машинного обучения, что, в свою очередь, позволит рациональнее распределять имеющиеся вычислительные ресурсы и системы хранения данных, а также повысить эффективность их использования.

Список литературы

1. Список мощнейших суперкомпьютеров мира TOP500. URL: <https://www.top500.org/>.
2. МИВК. Многофункциональный информационно-вычислительный комплекс ОИЯИ. URL: <https://micc.jinr.ru/>.
3. Гетерогенная платформа «HybriLIT». URL: <http://hlit.jinr.ru/>.
4. Batch Docs. Планировщик заданий SLURM. URL: <https://batchdocs.web.cern.ch/linuxhpc/index.html>.
5. Абдрахманов М. И. Pandas. Работа с данными. — 2-е изд. — Devpractice Team, 2020. — 171 с.
6. Нуньес-Иглесиас Х. Элегантный SciPy / Х. Нуньес-Иглесиас, Х. Дэншоу, дер. Уолт Ван. — ДМК Пресс, 2018. — 336 с.
7. Бизли Д. Python. Книга рецептов / Д. Бизли, Б. К. Джонс. — ДМК Пресс, 2019. — 646 с.
8. Абдрахманов М. И. Библиотека Matplotlib. — Devpractice Team, 2019. — 125 с.

9. Орельен Ж. Прикладное машинное обучение с помощью Scikit-Learn, Keras и TensorFlow. Концепции, инструменты — 2-е изд. — Диалектика, 2020. — 1040 с.
10. Инструкция по работе. Буклет «Гетерогенная платформа HybriLIT». URL: http://hlit.jinr.ru/for_users/user_guide/#_3.