

ФУНКЦИОНАЛЬНЫЕ ВОЗМОЖНОСТИ СИСТЕМЫ АККАУНТИНГА И ОБРАБОТКИ СТАТИСТИКИ ИСПОЛЬЗОВАНИЯ ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСОВ СУПЕРКОМПЬЮТЕРА «ГОВОРУН»

Д. В. Беляков¹, М. Х. Киракосян¹, М. А. Любимова^{1,},
Д. В. Подгайный¹, О. И. Стрельцова¹*

¹ Объединенный институт ядерных исследований, Дубна, Россия

В статье рассматриваются функциональные возможности системы аккаунтинга и обработки статистики использования вычислительных ресурсов суперкомпьютера «Говорун», представлены примеры практического применения. Основное внимание уделено обработке и визуализации статистических данных, позволяющих оценивать эффективность использования вычислительных ресурсов.

The article discusses the functional capabilities of the accounting system and processing statistics on the use of computing resources of the Govorun supercomputer, and provides examples of practical application. The main attention is paid to the processing and visualization of statistical data that allow evaluating the efficiency of using computing resources.

PACS: 89.20.Ff

ВВЕДЕНИЕ

Суперкомпьютер «Говорун» [1] входит в состав гетерогенной платформы HybriLIT [2], используется для проведения ресурсоемких и массивно-параллельных расчетов и позволяет решать широкий спектр задач, стоящих перед ОИЯИ. Суперкомпьютер «Говорун» представляет собой двухкомпонентную систему, ориентированную на вычисления с использованием ресурсов CPU и GPU.

Для оценки эффективности использования вычислительных ресурсов суперкомпьютера необходимо осуществлять мониторинг производительности, накопление и анализ статистики использования вычислительных ресурсов.

В работе представлена система аккаунтинга, визуализации и анализа статистики использования вычислительных ресурсов суперкомпьютера «Говорун» [3] на основе Business Intelligence системы Yandex DataLens [4], которая является свободно распространяемым программным продуктом коммерческого уровня.

* E-mail: lma@jinr.ru

Разработанная система предназначена для внутреннего использования системными администраторами и руководством суперкомпьютера «Говорун». Она предоставляет информацию о том, как используются вычислительные ресурсы различными группами пользователей, а также позволяет выявлять тенденции и различные аномалии в использовании ресурсов. Источником данных для разработанной системы является база данных планировщика задач SLURM [5].

ПОЛЬЗОВАТЕЛЬСКИЙ ИНТЕРФЕЙС СИСТЕМЫ

Разработанная система представляет статистику по использованию вычислительных ресурсов в виде интерактивного дашборда, элементами которого являются гистограммы, таблицы и индикаторы. Для управления сортировкой в элементах дашборда доступны селекторы, которые позволяют сортировать данные по годам и по счетным очередям или одновременно по обоим параметрам. В элементе «Таблица» доступна сортировка по любому столбцу по возрастанию или убыванию.

На главной странице дашборда (рис. 1) представлена общая обезличенная статистика использования вычислительных ресурсов суперкомпьютера «Говорун», которая включает данные за все время накопления статистики и охватывает все группы пользователей. Группы сформированы по месту работы пользователей (лаборатории ОИЯИ), а сотрудники сторонних организаций объединены в группу «Institutes». Для каждой группы предусмотрена отдельная вкладка дашборда с подробной информацией: количеством пользователей, количеством запущенных задач, затраченными ядро-часами, использованными ресурсами и персональными данными каждого участника. Переключаться между группами можно, выбрав соответствующую вкладку в верхней части страницы.

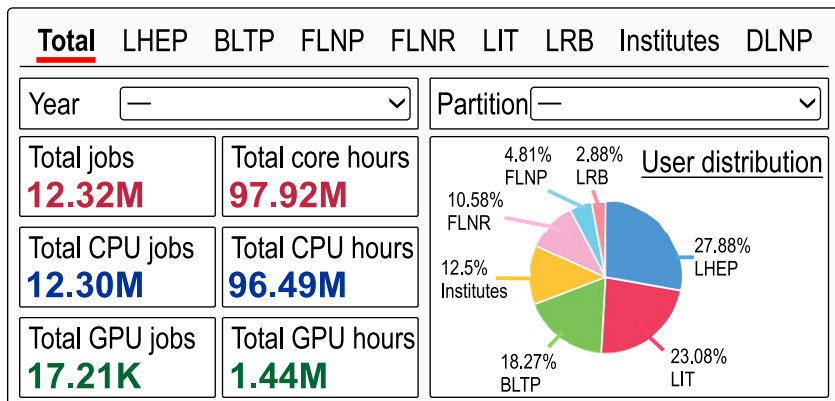


Рис. 1. Система аккаунтинга. Элемент главной страницы

СТАТИСТИКА ИСПОЛЬЗОВАНИЯ ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСОВ

Главная страница дашборда содержит три информационных блока со статистикой: общая, по группам пользователей и по ресурсам.

Общая статистика представлена в виде индикаторов, показывающих количество счетных задач, общее количество затраченных ядро-часов, отдельно количество счетных задач CPU/GPU и количество затраченных CPU-ядро-часов и GPU-часов. Это дает первичное представление о загруженности вычислительных ресурсов суперкомпьютера «Говорун».

Отображение статистики по группам пользователей включает в себя несколько представлений. Распределение пользователей по лабораториям в процентном соотношении представлено в виде круговой диаграммы, что помогает оценить, насколько востребованы вычислительные ресурсы каждой лабораторией. Общее количество счетных задач и затраченных CPU-ядро-часов и GPU-часов для каждой группы, а также количество счетных задач, использующих ресурсы CPU и GPU, и, соответственно, количество затраченного времени по ресурсам показано в сводной таблице. Данные о распределении ресурсов CPU и GPU между группами визуализированы в виде круговых диаграмм для удобства восприятия. В конце блока приведена статистика по наиболее ресурсоемким проектам ОИЯИ, в том числе MPD, VM@N и проектам Лаборатории теоретической физики.

Блок статистики по ресурсам показывает данные об использовании счетных очередей и типов ресурсов, распределении задач по количеству использованных ядер CPU. Для визуализации распределения нагрузки по счетным очередям используются две гистограммы. Первая показывает количество запущенных на счет задач в каждой очереди, а вторая — затраченные CPU-ядро-часы. Эти гистограммы позволяют оценивать нагрузку на каждую очередь с позиции востребованности.

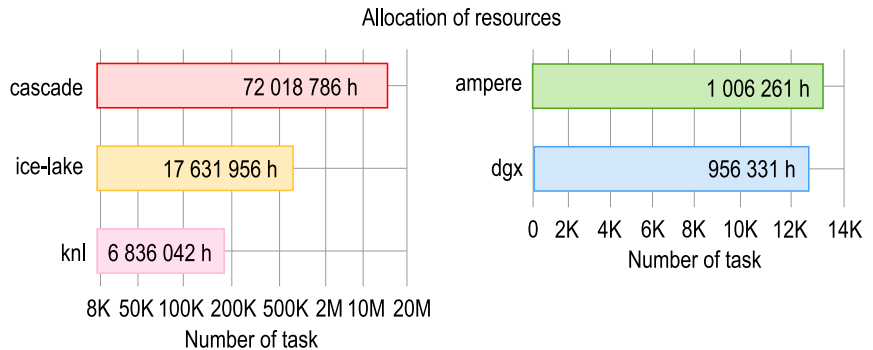


Рис. 2. Распределение нагрузки по типам ресурсов

Таблица 1. Распределение задач по использованным CPU-ядрам. Сортировка по убыванию в столбце «Количество задач»

CPU-ядра	Количество задач	Процент от всех задач	Количество ядро-часов	Процент от всех ядро-часов
2	2 783 812	38,54	5 849 780,01	7,55
1	2 731 684	38,09	413 161,00	0,53
8	623 764	8,70	1 436 081,78	1,85

Особый интерес представляет распределение нагрузки по типам ресурсов (рис. 2), поскольку данное распределение используется при принятии решений по вопросам модернизации вычислительных ресурсов суперкомпьютера «Говорун».

Распределение задач по количеству использованных ядер CPU представлено в табл. 1. Такое представление позволяет оценивать пользовательские задачи по типам расчетов (одноядерные, параллельные или массивно-параллельные вычисления).

АНАЛИЗ СТАТИСТИКИ. ПРИМЕРЫ ПРАКТИЧЕСКОГО ПРИМЕНЕНИЯ

Руководство суперкомпьютера «Говорун» с помощью разработанной системы может оценивать востребованность вычислительных ресурсов, что помогает поддерживать принятие решений по развитию суперкомпьютера. Например, простой анализ по годам показывает, что потребность в вычислительных ресурсах с каждым годом растет. В 2023 г. было запущено на счет 984 195 задач, в 2024 г. — 6 199 806 задач, а в первом полугодии 2025 г. — уже 5 133 336 задач.

Системные администраторы могут выявлять тенденции и аномалии в использовании ресурсов. Например, анализ распределения задач по использованным ядрам CPU (см. табл. 1) показывает, что наиболее часто пользователи запускают одно- и двухъядерные задачи CPU. Однако большое количество таких задач (80 % от количества всех задач) заняло малое количество вычислительных ресурсов (менее 9 % от всех затраченных CPU-ядро-часов). При этом небольшое количество задач, занявших для выполнения 1288 и 1932 ядер CPU (табл. 2), потребило значительное

Таблица 2. Распределение задач по использованным CPU-ядрам. Сортировка по убыванию в столбце «Процент от всех ядро-часов»

CPU-ядра	Количество задач	Процент от всех задач	Количество ядро-часов	Процент от всех ядро-часов
1 288	148	00,00	17 158 900,58	22,15
1 932	32	00,00	6 820 006,69	8,80
2	2 783 812	38,54	5 849 780,01	7,55

количество вычислительных ресурсов (30% от всех затраченных CPU-ядро-часов). Такие высокие показатели потребления вычислительных ресурсов требуют дальнейшего, более глубокого изучения с целью оценки эффективности использования ресурсов.

Анализ статистики по группам пользователей решает задачу аккаунтинга. Статистика использования ресурсов отдельным потребителем позволяет оценивать индивидуальный вклад каждого пользователя в использование вычислительных ресурсов, выявлять наиболее активных пользователей, а также определять потребности различных групп.

ЗАКЛЮЧЕНИЕ

В статье представлена система аккаунтинга, визуализации и анализа статистики использования вычислительных ресурсов суперкомпьютера «Говорун». Система позволяет системным администраторам и руководству суперкомпьютера «Говорун» отслеживать использование ресурсов, выявлять тенденции, а также оценивать вклад различных групп пользователей.

Благодаря интерактивным гистограммам, таблицам и индикаторам можно получить полное представление об использовании вычислительных ресурсов, вкладе отдельных лабораторий и проектов, а также выполнить аккаунтинг работы пользователей суперкомпьютера «Говорун». Интерактивный дашборд системы позволяет сортировать данные по различным параметрам, что упрощает анализ и выявление закономерностей.

Таким образом, разработанная система является эффективным инструментом для анализа использования вычислительных ресурсов суперкомпьютера «Говорун».

СПИСОК ЛИТЕРАТУРЫ

1. Суперкомпьютер «Говорун». http://hlit.jinr.ru/supercomputer_govorun/ (дата обращения 17.09.2025).
2. *Anikina A. I., Belyakov D. V., Bezhanian T. J., Kirakosyan M. H., Kokorev A. A., Lyubimova M. A., Matveev M. A., Podgainy D. V., Rahmonova A. R., Shadmehri S., Streltsova O. I., Torosyan S. G., Vala M., Zuev M. I.* Structure and Features of the Software and Information Environment of the HybriLIT Heterogeneous Platform // Distributed Comput. and Commun. Networks (DCCN 2024). Lecture Notes Comput. Sci. 2024. V. 15460. P. 444–457.
3. Разработка системы аккаунтинга и обработки статистики использования вычислительных ресурсов суперкомпьютера «Говорун». http://hlit.jinr.ru/accounting_and_statistics_processing_systems/#_user/ (дата обращения 17.09.2025).
4. DataLens. <https://datalens.tech/> (accessed 17.09.2025).
5. Slurm Workload Manager. <https://slurm.schedmd.com/> (accessed 17.09.2025).