



Разработка подхода к анализу доступности вычислительных узлов суперкомпьютерных комплексов

А. В. Папченко

Московский государственный университет имени М. В. Ломоносова,
факультет вычислительной математики и кибернетики,
Москва, Российская Федерация
ORCID: 0009-0002-0763-2484, e-mail: papchenko30@gmail.com

Д. А. Никитенко

Московский государственный университет имени М. В. Ломоносова,
Научно-исследовательский вычислительный центр,
Москва, Российская Федерация
ORCID: 0000-0002-2864-7995, e-mail: dan@parallel.ru

Аннотация: В работе рассматривается задача анализа доступности вычислительных ресурсов суперкомпьютерного центра коллективного пользования. Актуальность исследования обусловлена тем, что внедрение сторонних решений требует значительных затрат, а штатные средства не позволяют в полной мере оценивать доступность вычислительных систем. Предлагаемый подход основан на разработке специализированного модуля в составе системы Octoshell, объединяющего автоматически собираемые данные планировщика, экспертные комментарии администраторов и механизм накопления информации. Практическая значимость работы состоит в создании гибкого инструмента, ориентированного на задачи сопровождения вычислительных систем. Научная ценность заключается в формировании основы для дальнейшего исследования эксплуатационных, организационных и технических факторов, влияющих на доступность вычислительных ресурсов. Подход реализован на уровне прототипа.

Ключевые слова: суперкомпьютерный центр, вычислительные системы, доступность вычислительных ресурсов, анализ доступности, мониторинг состояния вычислительных узлов, Slurm, Octoshell, административные комментарии, классификация нештатных ситуаций, визуализация эксплуатационных данных.

Благодарности: Исследование выполнено в рамках государственного задания МГУ имени М. В. Ломоносова. При получении результатов использовалось оборудование Центра коллективного пользования сверхвысокопроизводительными вычислительными ресурсами МГУ имени М. В. Ломоносова.

Для цитирования: Папченко А.В., Никитенко Д.А. Разработка подхода к анализу доступности вычислительных узлов суперкомпьютерных комплексов // Вычислительные методы и программирование. 2026. 27, № 3. 359–373. doi 10.26089/NumMet.v27r323.



Development of an approach to analyzing the availability of computing nodes of supercomputer complexes

Anastasia V. Papchenko

Lomonosov Moscow State University,
Faculty of Computational Mathematics and Cybernetics,
Moscow, Russia

ORCID: 0009-0002-0763-2484, e-mail: papchenko30@gmail.com

Dmitry A. Nikitenko

Lomonosov Moscow State University, Research Computing Center,
Moscow, Russia

ORCID: 0000-0002-2864-7995, e-mail: dan@parallel.ru

Abstract: The paper considers the problem of analyzing the availability of computing resources of a supercomputer center for collective use. The relevance of the study is due to the fact that the implementation of third-party solutions requires significant costs, and the standard facilities do not fully assess the availability of computing systems. The proposed approach is based on the development of a specialized module within the Octoshell system that combines automatically collected scheduler data, expert comments from administrators, and a mechanism for accumulating aggregated information. The practical significance of the work is to create a flexible tool focused on the tasks of computer system maintenance. The scientific value lies in forming the basis for further research of operational, organizational and technical factors affecting the availability of computing resources. The approach is implemented at the prototype level.

Keywords: supercomputer center, computing systems, availability of computing resources, availability analysis, monitoring of the state of computing nodes, Slurm, Octoshell, administrative comments, classification of emergency situations, visualization of operational data.

Acknowledgements: The research was carried out within the framework of the state assignment of the Lomonosov Moscow State University. When obtaining the results, the equipment of the Center for the Collective Use of Ultra-high-performance Computing Resources of Lomonosov Moscow State University was used.

For citation: A. V. Papchenko, D. A. Nikitenko, “Development of an approach to analyzing the availability of computing nodes of supercomputer complexes,” *Numerical Methods and Programming*. 27 (3), 359–373 (2026). doi 10.26089/NumMet.v27r323.

1. Введение. Надежное функционирование суперкомпьютерного центра коллективного пользования требует не только контроля формального состояния вычислительных узлов, но и возможности содержательного анализа причин, влияющих на доступность ресурсов. Для администраторов и специалистов по сопровождению важно отслеживать факт изменения состояния узла в контексте происходящих событий [1]: наличие профилактических работ, изменение режима работы системы и иные факторы, способные влиять на возможность использования вычислительных ресурсов. При этом традиционные машинные статусы, формируемые средствами планировщика, не всегда позволяют получить полную и однозначную картину [2].

Одним из возможных путей решения данной задачи является внедрение внешних систем мониторинга и визуализации [3]. Однако подобный подход на практике связан с заметными затратами на интеграцию, настройку и последующее сопровождение. Кроме того, использование сторонних решений не гарантирует, что полученный функционал будет в полной мере соответствовать требованиям конкретного суперкомпьютерного центра, его внутренним регламентам и особенностям эксплуатации. По этой причине более целесообразным представляется развитие специализированного модуля внутри уже используемой системы контроля, способного органично встраиваться в существующую инфраструктуру и учитывать реальные потребности администраторов.



В работе рассматривается подход к разработке такого модуля в составе системы Octoshell [4, 5]. Целью исследования является создание гибкого средства представления и анализа доступности вычислительных ресурсов, объединяющего автоматически собираемые сведения о состоянии узлов, экспертные комментарии администраторов и механизмы накопления агрегированной информации. Предлагаемое решение ориентировано не только на отображение текущего состояния вычислительных ресурсов, но и на формирование основы для более глубокого анализа факторов, влияющих на их доступность [6], что представляет интерес как с практической, так и с исследовательской точки зрения.

2. Обзор способов представления информации о доступности в ведущих суперкомпьютерных центрах коллективного пользования. Предоставление пользователям и администраторам актуальной информации о состоянии вычислительных узлов суперкомпьютерного центра коллективного пользования является очень важной составляющей для надежной работы системы. Под такой информацией понимаются сведения о доступности вычислительных узлов, состоянии очередей заданий, плановых и внеплановых остановках, загрузке систем, а также о технических ограничениях, влияющих на возможность использования ресурсов. Подходы к публикации подобных данных в различных суперкомпьютерных центрах существенно отличаются: одни организации ограничиваются статическими описаниями и годовыми отчетами, тогда как другие публикуют интерактивные панели мониторинга или специализированные страницы с машинными статусами.

Для комплексного изучения поставленной задачи был проведен обзор, который показал, что даже среди ведущих мировых центров высокопроизводительных вычислений отсутствует единый стандарт представления информации о доступности вычислительных ресурсов. Это связано как с организационными особенностями, так и с различиями в требованиях к безопасности, режимам доступа и составу пользовательской аудитории. Вместе с тем можно выделить несколько устойчивых моделей публикации информации: представление только общих технических характеристик систем, публикация агрегированных показателей в годовых отчетах, публикация сообщений о текущих инцидентах и работах, а также использование интерактивных веб-интерфейсов, отображающих состояние ресурсов в режиме, близком к реальному времени.

Показательным примером закрытого, но хорошо структурированного подхода является Lawrence Livermore National Laboratory (LLNL) [7]. На официальных ресурсах лаборатории публикуются сведения о вычислительных платформах и их технических характеристиках, включая информацию о суперкомпьютере El Capitan и других высокопроизводительных системах. Однако оперативные данные о доступности кластеров, мониторинге состояния узлов и средствах управления заданиями не выносятся в полностью открытый доступ. Для получения более подробной информации требуется работа в специализированных средах доступа, разделенных по уровням допуска. Такой подход подчеркивает, что для центров, ориентированных на задачи национального уровня и работу с чувствительными данными, публичное отображение доступности вычислительной инфраструктуры может быть существенно ограничено. С точки зрения внешнего пользователя подобная модель дает представление о масштабах и архитектуре комплекса, но практически не позволяет оценивать текущее состояние ресурсов.

Отдельного внимания заслуживают китайские суперкомпьютерные системы. Китай занимает сильные позиции в области высокопроизводительных вычислений, однако сведения о крупнейших современных установках и их эксплуатационном состоянии публикуются ограниченно. В качестве характерного примера можно рассмотреть Национальный суперкомпьютерный центр в Уси (National Supercomputing Center in Wuxi), на официальном сайте которого представлена информация о системе “Sunway TaihuLight” (“Шэньвэй — свет Тайху”) [8]. Согласно опубликованным данным, центр располагает суперкомпьютером, построенным на отечественных многоядерных процессорах “SW26010”; система имеет пиковую производительность 125.436 PFlops, измеренную производительность 93.015 PFlops, 40960 процессоров, 10649600 вычислительных ядер и общий объем памяти 1310720 GB [9]. На той же странице приведены сведения о системе “CPU-система нового поколения Цинхайского филиала”: ее суммарная производительность составляет 5.3 PFlops, объем памяти — 560 TB, а конфигурация включает узлы на базе Intel Max 9462 и AMD EPYC 9654 [9]. Эти данные позволяют оценить масштаб и технический уровень вычислительной инфраструктуры, однако сайт не содержит открытых оперативных сведений о текущей доступности узлов, состоянии очередей, плановых и внеплановых работах или эксплуатационных событиях.

Иной подход демонстрирует Oak Ridge Leadership Computing Facility (OLCF) [10]. На официальном сайте центра приводится подробная информация о составе вычислительных систем, архитектурах и назначении ресурсов, однако сведения о текущем состоянии вычислительных узлов также не публикуются.

Вместо этого важную роль играют годовые отчеты [11], в которых доступны агрегированные показатели: плановые и фактические значения доступности вычислительных ресурсов, среднее время простоя, среднее время восстановления работоспособности, доля времени, в течение которого узлы были заняты пользовательскими задачами, а также распределение ресурсов между крупными проектами. Такой формат особенно полезен для аналитической оценки эффективности эксплуатации суперкомпьютерного центра, но хуже подходит для ежедневного информирования пользователей о текущем состоянии инфраструктуры. Тем не менее именно OLCF демонстрирует сильную сторону отчетной модели: из агрегированных метрик можно получить представление не только о доступности ресурсов, но и о значимости эксплуатационных процессов, качестве сопровождения систем и характерных рисках при работе крупных вычислительных комплексов.

Одним из наиболее выразительных примеров открытого визуального представления состояния ресурсов является Argonne Leadership Computing Facility (ALCF) [12]. На сайте центра, помимо описания систем и научных программ, представлен раздел *Machine Status*, в котором в наглядной форме отображается состояние суперкомпьютеров Aurora, Polaris, Sophia и Crux (рис. 1). Пользователь получает доступ к ключевым оперативным показателям: числу запущенных заданий, длине очереди, числу зарезервированных и используемых узлов, а также к интегральной оценке загрузки систем. Важным достоинством этого подхода является не только наличие числовых метрик, но и продуманная визуальная форма их представления: данные сгруппированы по системам, легко сравниваются между собой и позволяют быстро оценить уровень занятости ресурсов. Дополнительно ALCF публикует годовые отчеты и сведения о распределении ресурсов по программам и проектам, что делает информационную среду многослойной: пользователь может получить как мгновенную картину состояния, так и более широкий аналитический контекст. Именно такая комбинация оперативной визуализации и агрегированной аналитики выглядит наиболее зрелой с точки зрения организации публичного доступа к информации о доступности суперкомпьютерных ресурсов.

Европейский подход к публикации статуса демонстрирует Jülich Supercomputing Centre (JSC) [13]. На официальном сайте вычислительного центра размещается ссылка на сервис *JSC Service Status* [14], который предоставляет открытый доступ к сообщениям о состоянии кластерных систем, файловых систем и вспомогательных сервисов. Данный ресурс, в отличие от полноценных панелей мониторинга, ориентирован прежде всего на публикацию эксплуатационных уведомлений: плановых работ, технических ограничений и сообщений о возникших инцидентах. В этом случае пользователь не видит развернутых метрик загрузки или доступности узлов в реальном времени, однако получает важную информацию о работоспособности инфраструктуры через событийную модель. Такой подход можно считать компромиссным: он

Machine Status

Mouse over diagrams for details

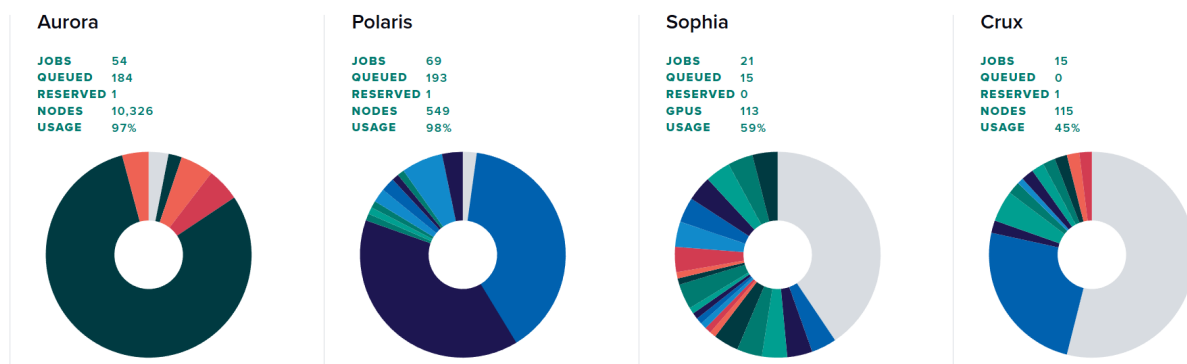


Рис. 1. Информационные панели, отражающие состояние доступности вычислительных ресурсов на официальной странице ALCF

Fig. 1. Information panels reflecting the availability status of computing resources on the official ALCF page



проще в реализации и безопаснее с точки зрения публикации внутренних данных, но при этом остается полезным для информирования пользователей о доступности ресурсов и ожидаемых ограничениях.

Наиболее развитую форму публичного представления информации о доступности среди рассмотренных центров демонстрирует RIKEN Center for Computational Science (R-CCS) [15], сопровождающий суперкомпьютер Fugaku. На сайте центра используется публичный дашборд (рис. 2) *Operating Status* [16], построенный на платформе Grafana. Он отображает в режиме, близком к реальному времени, ключевые показатели функционирования вычислительных ресурсов: количество выполняющихся и ожидающих заданий, статистику занятых, свободных и зарезервированных узлов, уровень загрузки CPU, параметры энергопотребления, состояние узлов, а также отдельные характеристики сетевого взаимодействия. Существенным преимуществом данного решения является то, что пользователь получает не изолированную текстовую сводку, а целостную визуальную картину состояния суперкомпьютера за выбранный временной интервал. Благодаря этому можно анализировать не только текущий статус, но и динамику изменения нагрузки. В контексте задач разработки средств отображения доступности вычислительных ресурсов именно пример R-CCS представляется одним из наиболее ценных, поскольку сочетает оперативность и высокую наглядность.

Интересный баланс между оперативностью и простотой представления реализован в итальянском центре CINECA. На официальном портале в разделе *System status in real time* [17] публикуются компактные панели для ключевых систем, включая Leonardo Booster, Leonardo DCGP и Galileo100. Эти панели содержат процентные показатели статуса вычислительных узлов, сведения о количестве активных, свободных и недоступных узлов, информацию о числе выполняемых и ожидающих заданий, а также о числе подключенных пользователей. Обновление информации осуществляется с небольшим интервалом в 10 мин, что позволяет использовать данный раздел как инструмент текущего мониторинга для широкой аудитории. По сравнению с более насыщенными аналитическими дашбордами, например в Grafana, решение CINECA выглядит проще, однако именно эта простота оказывается сильной стороной: пользователь сразу видит наиболее важные характеристики доступности без перегрузки интерфейса второстепенными деталями. Такой подход особенно ценен для публичных сервисов, где важно обеспечить быстрое восприятие информации.

Сопоставление рассмотренных примеров позволяет выделить несколько характерных тенденций. Во-первых, наиболее распространенной практикой остается публикация агрегированных сведений в годовых отчетах и обзорных материалах, тогда как полноценный открытый доступ к оперативным метрикам реализован лишь у части центров. Во-вторых, формы представления информации различаются по степени

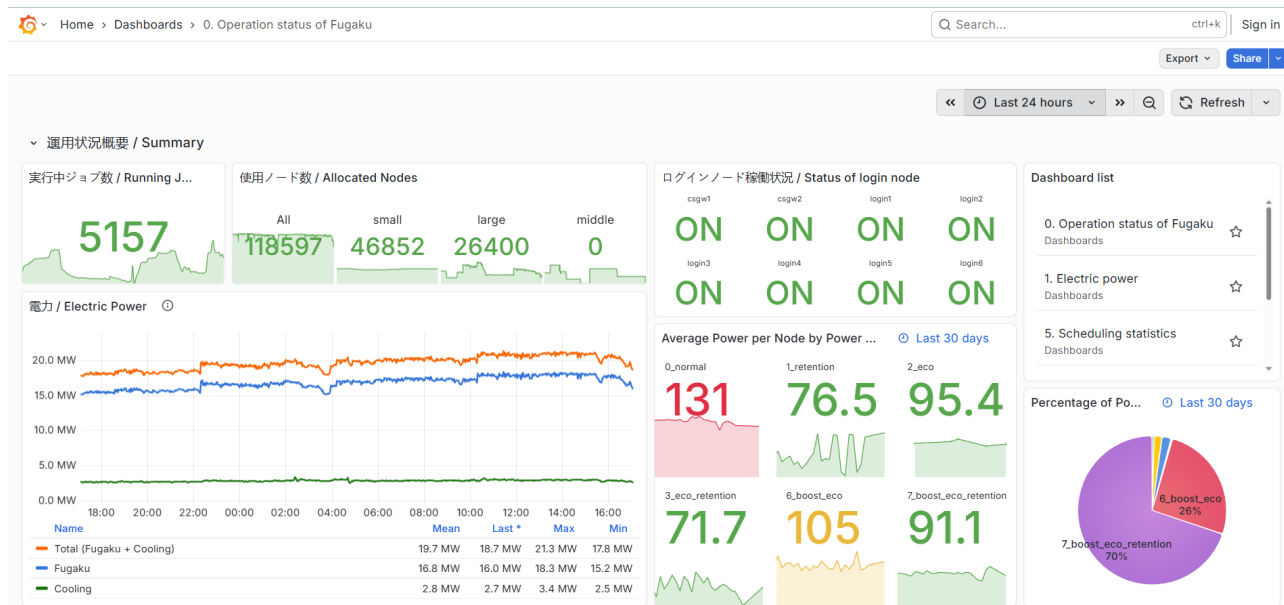


Рис. 2. Панель мониторинга состояния суперкомпьютерного комплекса R-CCS на платформе Grafana

Fig. 2. The status monitoring panel of the R-CCS supercomputer complex on the Grafana platform

детализации: от текстовых уведомлений о состоянии сервисов до интерактивных панелей с отображением загрузки, статуса узлов и характеристик очередей. В-третьих, при проектировании публичных интерфейсов центры вынуждены искать баланс между наглядностью, полезностью для пользователей и ограничениями, связанными с безопасностью и внутренними регламентами.

С практической точки зрения наиболее информативными можно считать три подхода. Первый основан на использовании полноценных дашбордов реального времени, как в R-CCS, где пользователь получает многомерное представление о состоянии системы. Второй связан с компактными и интуитивно понятными панелями статуса, как в CINECA и ALCF, где отображаются только ключевые показатели доступности. Третий представляет собой событийную модель, применяемую в JSC, где основное внимание уделяется публикации сообщений об инцидентах, регламентных работах и технических ограничениях. Напротив, модель LLNL характеризуется низкой степенью публичной информации: полезна для общего знакомства с центром и оценки масштабов его деятельности, но ограничено применима при анализе текущей доступности вычислительных ресурсов.

3. Подход к анализу доступности вычислительных узлов. В предыдущем разделе был продемонстрирован подробный обзор среди ведущих суперкомпьютерных центров наиболее ярких примеров реализации разного вида представлений информации о доступности вычислительных ресурсов в суперкомпьютерных центрах коллективного пользования. Было выделено три основных способа представления информации:

- публикация агрегированных сведений в годовых отчетах и обзорных материалах;
- публикация сообщений об инцидентах, регламентных работах и технических ограничениях;
- публикация интерактивных дашбордов (обновление информации от 10 до 60 мин) с демонстрацией основных параметров доступности вычислительных ресурсов.

Данная информация легла в основу разработки модуля демонстрации доступности вычислительных ресурсов суперкомпьютерного центра МГУ, основой которого является суперкомпьютер “Ломоносов-2” [18]. Модуль реализуется как расширение существующей системы контроля суперкомпьютерного центра Octoshell и будет демонстрировать данные для пользователей с правами администратора. Для структурированного представления информации организацию данных планируется разделить на три основные вкладки:

1. Аналитика доступности узлов в реальном времени (данные, приходящие из логов).
2. Комментарии/пометки администратора (ручной ввод/форма для ввода).
3. Агрегированные данные из логов планировщика и ручного ввода.

3.1. Архитектура решения. Архитектура разработанного решения включает в себя несколько взаимосвязанных компонентов, обеспечивающих сбор, хранение и представление информации о состоянии вычислительных узлов суперкомпьютера. Общая схема взаимодействия компонентов представлена на рис. 3. В основе решения лежит интеграция автоматического сбора данных из системы управления вычислительными ресурсами Slurm, ручного ввода комментариев администраторами, централизованного хранения информации в базе данных Octoshell и последующего визуального представления данных в виде графиков для пользователей с группой доступа администраторов.

Суперкомпьютерный кластер является источником информации о состоянии вычислительных узлов. Управление ресурсами кластера осуществляется планировщиком Slurm, который взаимодействует с узлами и предоставляет сведения об их состоянии. Эти данные используются для формирования снимков состояния системы.

Octoshell, как система контроля суперкомпьютерного комплекса, выступает в роли связующего компонента, на базе которого происходит хранение и обработка поступающей информации. В базе данных Octoshell сохраняются два основных типа данных: *snapshots* и *comments*. Снимки состояния (*snapshots*) формируются на основе информации, поступающей из Slurm, и отражают текущее состояние узлов суперкомпьютеров. Комментарии (*comments*) содержат текстовые пояснения и замечания, которые вводятся администраторами вручную. Более подробно о структуре базы данных разработанного модуля будет говориться в следующих разделах.

Администраторы участвуют в работе системы как источник экспертной информации. Они дополняют автоматически собираемые данные комментариями, необходимыми для уточнения причин сбоев,

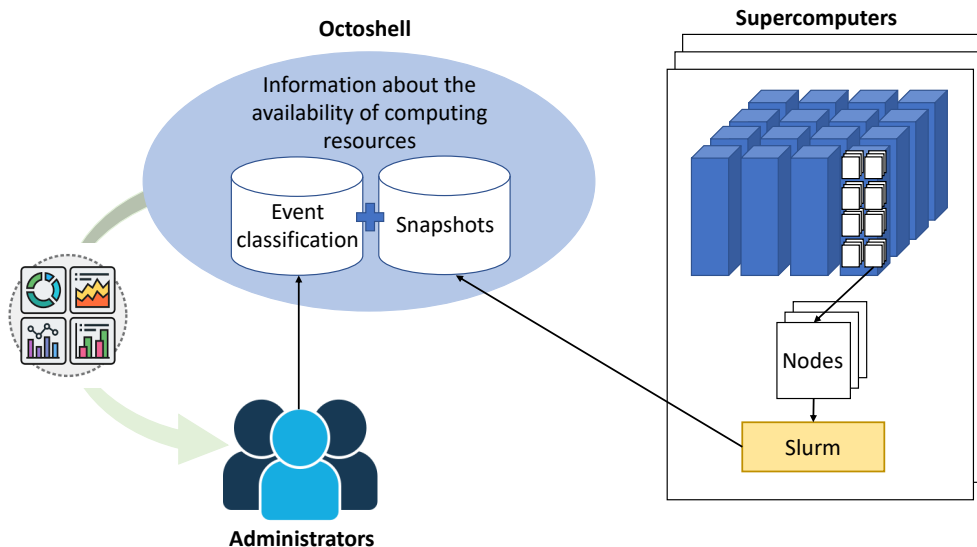


Рис. 3. Общая схема архитектуры разработанного решения
 Fig. 3. The general scheme of the architecture of the developed solution

описания особенностей работы узлов и фиксации важных событий, которые не могут быть однозначно определены автоматически. Таким образом, в системе сочетаются машинные данные и экспертные сведения, вводимые человеком.

Графики представляют собой уровень визуализации данных. На основе сохраненной информации формируются графические представления, предназначенные для администраторов. Это позволяет наглядно анализировать состояние вычислительных узлов, отслеживать изменения во времени и использовать накопленные данные для принятия решений по сопровождению и эксплуатации вычислительного кластера.

Таким образом, архитектура решения строится вокруг централизованного хранения снимков состояния и комментариев в Octoshell, объединяя автоматический сбор данных из Slurm, ручной ввод информации администраторами и последующее представление результатов в виде графиков.

3.2. Организация сбора данных и классификация нештатных ситуаций. Одной из ключевых задач в рамках данной работы является формирование целостного представления о состоянии вычислительной системы на основе как автоматически собираемых технических данных, так и экспертной информации, поступающей от администраторов. Такой подход позволяет учитывать не только формально фиксируемые изменения состояния узлов, но и внешние либо организационные факторы, способные существенно влиять на доступность и производительность вычислительных ресурсов.

В разработанном решении подсистема сбора данных строится на сочетании двух взаимодополняющих источников информации: машинного и ручного. Машинный источник обеспечивает получение объективных сведений о фактическом состоянии узлов со стороны планировщика заданий, тогда как ручной источник предназначен для фиксации событий, причин и контекста, которые не всегда могут быть однозначно определены автоматически.

§ 3.2.1. Сбор и обработка данных, поступающих от планировщика заданий. Машинный источник данных отражает фактическое состояние вычислительных узлов со стороны планировщика и используется в качестве базового слоя для построения аналитики [1]. В качестве планировщика в рассматриваемой системе используется Slurm. Сбор машинных данных осуществляется путем периодического получения вывода команды `sinfo -a`. Полученный результат сохраняется в виде снимка состояния системы, после чего выполняется его последующая обработка и нормализация до уровня отдельных узлов. Для каждого узла в системе сохраняются агрегированное состояние, принадлежность к определенной очереди, а также признак наличия указанной причины изменения состояния (`has_reason`). История изменений представ-

ляется в виде временных интервалов действия записи с использованием полей `valid_from` и `valid_to`. Это дает возможность получать как актуальный срез состояния системы, так и выполнять анализ за произвольный временной период.

На этапе прототипирования в качестве основного механизма сбора данных был выбран периодический разбор вывода команды `sinfo -a`, поскольку существовала необходимость быстро интегрировать модуль с действующей инфраструктурой Octoshell без существенной переработки уже используемых административных сценариев. Вместе с тем более предпочтительным направлением дальнейшего развития является использование программного интерфейса Slurm REST API, реализуемого через службу `slurmrestd` и предназначенного для взаимодействия внешних клиентов со Slurm через REST-запросы [19]. Переход к такому механизму позволит сделать получение данных более формализованным и уменьшить зависимость точности анализа от интервала между периодическими снимками.

§ 3.2.2. *Классификация нештатных ситуаций и экспертные данные администраторов.* Несмотря на высокую информативность машинных данных, они не всегда позволяют определить реальные причины возникновения проблем или изменений в работе системы. По этой причине в архитектуру решения был включен ручной источник данных, предназначенный для фиксации информации, поступающей непосредственно от администраторов. К таким сведениям относятся комментарии, пояснения, служебные пометки и иные записи, позволяющие интерпретировать происходящие события с учетом эксплуатационного контекста. В данной работе экспертные комментарии рассматриваются шире, чем формальная причина изменения состояния узла, фиксируемая средствами планировщика. Они предназначены для описания эксплуатационного контекста: какие работы проводились в системе, какие внешние или организационные события наблюдались в рассматриваемый момент времени и каким образом они могли повлиять на доступность вычислительных ресурсов. Такая информация важна не только для текущего администрирования, но и для последующего восстановления картины событий, анализа причин ограничений и подготовки эксплуатационных отчетов.

Администратор задает содержание пометки, указывает интервал ее действия, определяет уровень серьезности события и, при необходимости, привязывает запись к конкретным вычислительным узлам. Наличие такого механизма существенно расширяет аналитические возможности системы, поскольку позволяет учитывать факторы, которые не отражаются напрямую в логах планировщика, но при этом оказывают влияние на доступность, производительность и режим эксплуатации вычислительных ресурсов.

На текущем этапе была разработана классификация нештатных ситуаций, связанных с функционированием вычислительной системы. Данная классификация является одной из важнейших составляющих предложенного решения, поскольку позволяет перевести разрозненные наблюдения и комментарии администраторов в структурированную форму, пригодную для хранения, анализа и дальнейшего использования в административном интерфейсе. Предложенная классификация была сформирована в процессе общения с администраторами и отражает их практический опыт эксплуатации вычислительного комплекса. Следует отметить, что данный перечень не является окончательным: по мере развития системы и накопления новых наблюдений он может быть расширен и уточнен. Однако для первой реализации был сформирован базовый набор категорий, позволяющий описывать наиболее характерные и значимые события.

Разработанная система классификации включает три основные группы и набор подкатегорий в рамках каждой из них:

1. Нештатные ситуации

- аномально высокая уличная температура;
- отключение или недоступность части системы охлаждения;
- ограничение энергоснабжения;
- аварийная ситуация общего характера;
- снижение производительности сети межсоединений;
- массовый отказ части вычислительных узлов;
- недоступность или проблемы с узлами логина;
- недоступность или нестабильность системы очередей.

2. Изменение режима работы суперкомпьютерного центра

- выделение ресурсов для больших задач;



- выделение ресурсов для приоритетных проектов;
- введение в эксплуатацию нового оборудования;
- временное снижение пользовательских лимитов;
- снижение производительности сети межсоединений;
- включение режима бенчмарков или тестирования производительности;
- выделение ресурсов под образовательные нужды (курсы, мастер-классы);
- тестирование новых настроек системного программного обеспечения.

3. Профилактические и ремонтные работы

- обновление системного программного обеспечения;
- изменения в настройках файловой системы;
- обновление системы поддержки функционирования суперкомпьютерного центра;
- профилактика или ремонт дисковой подсистемы;
- работы с сетевой инфраструктурой;
- работы с энергоснабжением и системой бесперебойного питания;
- работы с системой охлаждения;
- замена или расширение вычислительных узлов;
- плановые работы в машинном зале.

3.3. Структура базы данных модуля. Для реализации модуля демонстрации доступности вычислительных ресурсов суперкомпьютерного центра была разработана дополнительная структура базы данных (рис. 4), расширяющая существующую схему системы Octoshell. Новые таблицы не дублируют уже имеющиеся сущности, а дополняют их, связываясь с таблицами `USERS`, `CORE_CLUSTERS` и `CORE_PARTITIONS`.

Основу расширения составляет подсистема хранения автоматически собираемых данных о состоянии вычислительных ресурсов. Таблица `CORE_ANALYTICS_SNAPSHOTS` хранит снимки состояния кластера, т.е. результаты отдельных запусков парсинга данных, полученных с помощью команды `sinfo`. Для каждого снимка фиксируются кластер, момент получения данных, команда-источник, версия парсера и полный исходный текст результата. Таблица `CORE_ANALYTICS_NODES` представляет собой справочник вычислительных узлов внутри кластера и содержит сведения об их именах и префиксах. Центральной таблицей аналитической части является `CORE_ANALYTICS_NODE_STATES`, в которой для каждого снимка сохраняется состояние конкретного узла в конкретной очереди. В этой таблице фиксируются основное состояние узла, дополнительное подсостояние, признак наличия причины, а также временной интервал актуальности записи, задаваемый полями `valid_from` и `valid_to`.

Для хранения информации, вводимой администраторами вручную, предусмотрена таблица под названием `CORE_COMMENTS`. Она содержит комментарии, связанные с кластером и автором записи из таблицы `USERS`. Для каждого комментария сохраняются заголовок, текстовое описание, период актуальности, уровень серьезности, а также ссылки на классификацию причины и конкретную причину. Подробней о классификациях причин рассказывается в разделе 3.2. Эти данные образуют ручной слой описания событий и инцидентов, который дополняет автоматически собранную информацию и позволяет фиксировать сведения, не извлекаемые напрямую из системных команд.

Поскольку один комментарий может относиться к нескольким узлам, а один узел может упоминаться в нескольких комментариях, связь между комментариями и узлами реализована через промежуточную таблицу `CORE_COMMENTS_NODES`. Это обеспечивает гибкую привязку экспертных замечаний к конкретным вычислительным узлам и позволяет использовать комментарии при анализе причин недоступности или ограничений в работе ресурсов.

4. Реализация модуля мониторинга доступности вычислительных ресурсов. Практическая реализация предложенного решения выполнена в качестве дополнения к существующей системе Octoshell в виде отдельного административного модуля, предназначенного для анализа доступности вычислительных ресурсов суперкомпьютерного центра. В отличие от предыдущих разделов, где были рассмотрены архитектура решения, организация сбора данных и структура базы данных, в данном разделе основное внимание уделяется пользовательскому представлению информации и общим принципам организации интерфейса.

Модуль ориентирован на пользователей с административными правами и предназначен для совместного использования автоматически собираемых сведений о состоянии вычислительных ресурсов и



экспертной информации, вводимой вручную. В основу пользовательского интерфейса положено разделение данных на несколько взаимодополняющих уровней представления, соответствующих различным задачам анализа. Это позволяет не смешивать оперативные сведения о текущем состоянии узлов, комментарии эксплуатационного характера и обобщенные результаты анализа, формируемые по накопленной истории наблюдений. Структурно интерфейс модуля организован в виде трех основных вкладок.

- Первая вкладка предназначена для отображения аналитики доступности узлов в реальном времени (рис. 5). В ней представляются сведения, формируемые на основе автоматически получаемых данных планировщика, что позволяет получать актуальные данные о состоянии вычислительных узлов в каждый момент времени и оперативно фиксировать признаки нештатных ситуаций. Прежде всего такой сбор данных необходим для общей картины восприятия состояния системы и для последующего перехода к более детальному анализу.
- Вторая вкладка обеспечивает поддержку формы, которая собирает экспертные данные от администраторов суперкомпьютерного комплекса (рис. 6). Ее назначение состоит в дополнении машинных данных сведениями, которые невозможно однозначно восстановить только по журналам и состояниям планировщика. В данном разделе интерфейса фиксируются пояснения к наблюдаемым событиям, сведения о профилактических работах, организационных изменениях режима эксплуатации, а также иные замечания, позволяющие интерпретировать причины ограничений доступности. Наличие такого слоя особенно важно в тех случаях, когда одинаковые машинные состояния могут быть вызваны различными по смыслу эксплуатационными причинами. Именно для этого раздела была разработана классификация нештатных ситуаций (см. § 3.2.2), которая является одним из важнейших результатов данного исследования.
- Третья вкладка предназначена для представления агрегированных данных, формируемых на основе накопленных записей автоматического и ручного происхождения (рис. 7). В отличие от оперативной аналитики, ориентированной на текущее состояние системы, данный раздел служит для выявления устойчивых закономерностей в работе вычислительного комплекса. С его помощью можно анали-

Очередь / раздел	Всего узлов (по состояниям)	Заняты (alloc)	Свободны (idle)	Недоступны (down)	Прочие
compute	357	195	6	56	100
compute_prio	1118	610	29	184	295
test	17	0	14	0	3
phi	4	0	0	4	0
service	2	0	0	0	2
pascal_prio	160	140	3	12	5
volta1_prio	16	0	14	1	1
volta2_prio	18	2	4	3	9
pec	8	0	0	5	3

Снимки систем

Время снимка	Система	Всего	Заняты (alloc)	Свободны (idle)	Недоступны (down)	Прочие
12 апр., 00:18	test	342	227	47	1	67
25 марта, 23:19	test	1700	800	30	313	557

Текущее состояние узлов



Рис. 5. Прототип интерфейса вкладки с информацией от планировщика заданий, который содержит тестовые данные

Fig. 5. A prototype of the tab interface with information from the task scheduler, which contains test data

Новый комментарий

Система * Категория *

Комментарий привязывается к системе и (опционально) к конкретным узлам.

Причина *

Заголовок *

Описание (опционально)

Действует с * Действует до (опционально) Серьёзность *

Затронутые узлы (опционально)
 Можно искать по имени. Выбирай узлы только в рамках выбранной системы.

 Выберите систему выше — тогда появится список узлов.

Рис. 6. Прототип интерфейса вкладки с формой обратной связи от администраторов (интерфейс для создания комментариев)

Fig. 6. Prototype of the tab interface with a feedback form from administrators (interface for creating comments)

Параметры графиков

С По

Полупрозрачные зоны — интервалы комментариев администраторов.

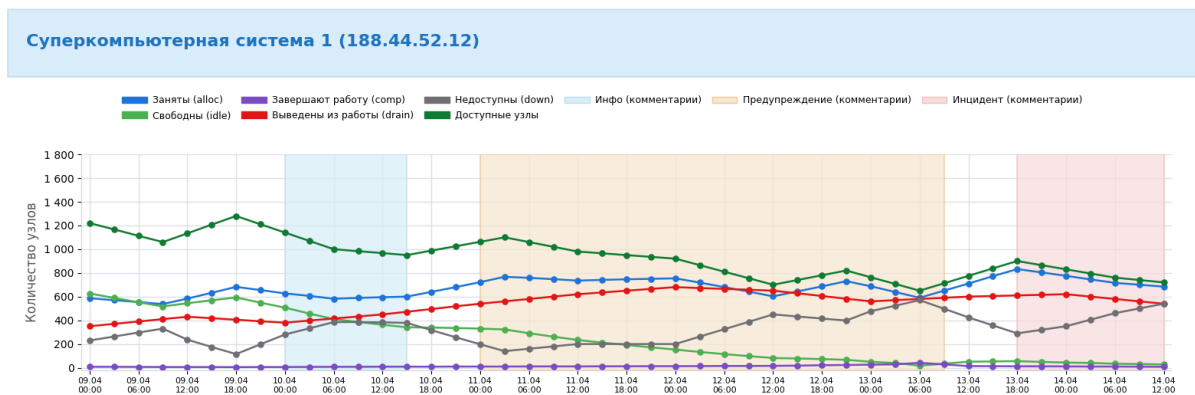


Рис. 7. Прототип интерфейса вкладки с агрегированной информацией, сформированной на основе данных планировщика заданий и комментариев администратора, содержащий тестовые данные

Fig. 7. A prototype of the tab interface with aggregated information based on task scheduler data and administrator comments, containing test data



зировать повторяемость событий, сопоставлять различные типы нештатных ситуаций, оценивать длительность ограничений и рассматривать общую динамику доступности ресурсов за продолжительные интервалы времени. Тем самым интерфейс поддерживает не только мониторинг, но и последующий аналитический разбор эксплуатационной истории.

При проектировании пользовательского интерфейса особое внимание уделялось наглядности, компактности и логическому разделению различных типов данных. Автоматически собираемая информация отображает фактическое состояние вычислительных узлов, тогда как комментарии администратора обеспечивают необходимый контекст для интерпретации наблюдаемых изменений. Агрегированный уровень, в свою очередь, позволяет перейти от отдельных событий к их обобщению и анализу. Такое построение интерфейса обеспечивает последовательный переход от текущего наблюдения к более глубокому пониманию причин и характера ограничений в работе вычислительных ресурсов.

5. Заключение. Предложен подход к анализу доступности вычислительных ресурсов суперкомпьютерного центра коллективного пользования, ориентированный на использование в составе существующей системы Octoshell. Проведенный обзор показал, что применяемые в различных суперкомпьютерных центрах способы представления информации о состоянии ресурсов существенно различаются по степени детализации, форме визуализации и уровню открытости, а универсального решения, одинаково пригодного для всех эксплуатационных сценариев, фактически не существует. Это подтверждает целесообразность разработки специализированного модуля, учитывающего особенности конкретной вычислительной инфраструктуры и административных задач. В рамках работы были определены основные принципы организации такого модуля:

- объединение машинных данных, получаемых от планировщика, с ручными комментариями администраторов;
- введение классификации нештатных ситуаций и эксплуатационных событий;
- разработка структуры хранения информации и пользовательского представления данных.

Предложенное решение позволяет перейти от фиксации отдельных состояний узлов к более содержательному описанию факторов, сопровождающих изменения доступности вычислительных ресурсов, и тем самым расширяет аналитические возможности системы.

Практическая значимость работы состоит в том, что разрабатываемый модуль представляет собой гибкое средство, изначально ориентированное на потребности сопровождения вычислительных кластеров и не требующее внедрения внешних платформ с неопределенной полезностью и дополнительными издержками на адаптацию. На текущем этапе разработанный модуль является внутренним административным расширением системы Octoshell, предназначенным для использования в инфраструктуре суперкомпьютерного центра МГУ, и не размещен в открытом доступе в сети Интернет. Возможность публикации исходных кодов модуля и условий лицензирования будут рассмотрены после завершения апробации и согласования с правилами сопровождения системы Octoshell.

Научная ценность заключается в создании основы для дальнейшего исследования влияния различных эксплуатационных, организационных и технических факторов на доступность вычислительных узлов. В перспективе развитие модуля может быть связано с расширением классификации событий, накоплением статистики и построением более глубоких методов анализа, позволяющих количественно оценивать степень воздействия отдельных факторов на состояние вычислительных ресурсов. Отдельным направлением дальнейшего развития может стать совместное использование разработанного модуля с программными средствами прогнозирования времени запуска пользовательских задач, основанными на статистическом анализе истории очереди [20].

Список литературы

1. Nikitenko D.A., Voevodin V.V., Zhumatiy S.A. Deep analysis of job state statistics on Lomonosov-2 supercomputer // *Supercomputing Frontiers and Innovations*. 2018. 5, N 2. 4–10. doi 10.14529/jsfi180201.
2. Voevodin V.V., Chulkevich R.A., Kostenetskiy P.S., et al. Administration, monitoring and analysis of supercomputers in Russia: a survey of 10 HPC centers // *Supercomputing Frontiers and Innovations*. 2021. 8, N 3. 82–103. doi 10.14529/jsfi210305.
3. Nikitenko D.A., Zhumatiy S.A., Shvets P.A. Making large-scale systems observable — another inescapable step towards exascale // *Supercomputing Frontiers and Innovations*. 2016. 3, N 2. 72–79. doi 10.14529/jsfi160205.

4. Nikitenko D., Zhumatiy S., Paokin A., et al. Evolution of the Octoshell HPC center management system // 13th International Conference Parallel Computational Technologies (PCT 2019), Kaliningrad, Russia, April 2–4, 2019. Communications in Computer and Information Science. Vol. 1063. Cham: Springer Nature, 2019. pp. 19–33. doi 10.1007/978-3-030-28163-2_2.
5. Паокин А.В., Никитенко Д.А., Жумати́й С.А. Исследование поддержки вариативности рабочих процессов в системе Octoshell // Параллельные вычислительные технологии — XV международная конференция (ПаВТ’2021), Волгоград, март 30 — апрель 1, 2021. Челябинск: ЮУрГУ, 2021. 233–243.
6. Voevodin V.V., Shaikhislamov D.I., Nikitenko D.A. How to assess the quality of supercomputer resource usage // Supercomputing Frontiers and Innovations. 2022. 9, N 3. 4–18. doi 10.14529/jsfi220301.
7. Lawrence Livermore National Laboratory. <https://www.llnl.gov/>. Cited June 20, 2026.
8. National Supercomputing Center in Wuxi. <https://www.nscwx.cn/about/71.html>. Cited June 20, 2026.
9. National Supercomputing Center in Wuxi. Hardware resources. <https://www.nscwx.cn/Resource/81.html>. Cited June 20, 2026.
10. Oak Ridge Leadership Computing Facility - The OLCF was established at Oak Ridge National Laboratory in 2004 with the mission of standing up a supercomputer 100 times more powerful than the leading systems of the day. <https://www.olcf.ornl.gov/>. Cited June 20, 2026.
11. Center Reports - Oak Ridge Leadership Computing Facility. <https://www.olcf.ornl.gov/about-olcf/center-reports/>. Cited June 20, 2026.
12. Home | Argonne Leadership Computing Facility. <https://www.alcf.anl.gov/>. Cited June 20, 2026.
13. Jülich Supercomputing Centre (JSC). <https://www.fz-juelich.de/en/jsc>. Cited June 20, 2026.
14. JSC Service Status. <https://status.jsc.fz-juelich.de/>. Cited June 20, 2026.
15. RIKEN Center for Computational Science RIKEN Website | RIKEN R-CCS. <https://www.r-ccs.riken.jp/en/>. Cited June 26, 2025.
16. Operating Status | RIKEN Center for Computational Science RIKEN Website. <https://www.r-ccs.riken.jp/en/fugaku/opp-status/>. Cited June 26, 2026.
17. HPC Cineca. <https://www.hpc.cineca.it/>. Cited June 26, 2026.
18. Voevodin V.V., Antonov A.S., Nikitenko D.A., et al. Supercomputer Lomonosov-2: large scale, deep monitoring and fine analytics for the user community // Supercomputing Frontiers and Innovations. 2019. 6, N 2. 4–11. doi 10.14529/jsfi190201.
19. Slurm Workload Manager - REST API Details. https://slurm.schedmd.com/rest.html?utm_source. Cited June 20, 2026.
20. Яценко И.С., Сальников А.Н. Программное обеспечение для прогнозирования времени запуска задачи путем статистического анализа истории очереди задач // Физика элементарных частиц и атомного ядра. 2024. 55, № 3. 580–584.

Получена
17 апреля 2026 г.

Принята
26 мая 2026 г.

Опубликована
8 июля 2026 г.

Информация об авторах

Анастасия Вячеславовна Папченко — бакалавр физических наук, Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики, Ленинские горы, 1, стр. 4, 119234, Москва, Российская Федерация.

Дмитрий Александрович Никитенко — к.ф.-м.н., вед. науч. сотр.; Московский государственный университет имени М. В. Ломоносова, Научно-исследовательский вычислительный центр, Ленинские горы, 1, стр. 4, 119234, Москва, Российская Федерация.

References

1. D. A. Nikitenko, V. V. Voevodin, and S. A. Zhumatiy, “Deep Analysis of Job State Statistics on Lomonosov-2 Supercomputer,” Supercomputing Frontiers and Innovations 5 (2), 4–10 (2018). doi 10.14529/jsfi180201.
2. V. V. Voevodin, R. A. Chulkevich, P. S. Kostenetskiy, et al., “Administration, Monitoring and Analysis of Supercomputers in Russia: a Survey of 10 HPC Centers,” Supercomputing Frontiers and Innovations 8 (3), 82–103 (2021). doi 10.14529/jsfi210305.



3. D. A. Nikitenko, S. A. Zhumatiy, and P. A. Shvets, “Making Large-Scale Systems Observable — Another Inescapable Step Towards Exascale,” *Supercomputing Frontiers and Innovations* **3** (2), 72–79 (2016). doi [10.14529/jsfi160205](https://doi.org/10.14529/jsfi160205).
4. D. Nikitenko, S. Zhumatiy, A. Paokin, et al., “Evolution of the Octoshell HPC Center Management System,” in *13th International Conference Parallel Computational Technologies (PCT 2019), Kaliningrad, Russia, April 2–4, 2019*. Communications in Computer and Information Science. Vol 1063. (Springer Nature, Cham, 2019), pp. 19–33. doi [10.1007/978-3-030-28163-2_2](https://doi.org/10.1007/978-3-030-28163-2_2).
5. A. V. Paokin, D. A. Nikitenko, and S. A. Zhumatiy, “Study of Support for Workflow Variability in the Octoshell System,” in *Parallel Computing Technologies — XV International Conference (PCT’2021), Volgograd, March 30 — April 1, 2021* (YuUrGU, Chelyabinsk, 2021), pp. 233–243.
6. V. V. Voevodin, D. I. Shaikhislamov, and D. A. Nikitenko, “How to Assess the Quality of Supercomputer Resource Usage,” *Supercomputing Frontiers and Innovations* **9** (3), 4–18 (2022). doi [10.14529/jsfi220301](https://doi.org/10.14529/jsfi220301).
7. Lawrence Livermore National Laboratory. <https://www.llnl.gov/>. Cited June 20, 2026.
8. National Supercomputing Center in Wuxi. <https://www.nscwx.cn/about/71.html>. Cited June 20, 2026.
9. National Supercomputing Center in Wuxi. Hardware Resources. <https://www.nscwx.cn/Resource/81.html>. Cited June 20, 2026.
10. Oak Ridge Leadership Computing Facility - The OLCF was established at Oak Ridge National Laboratory in 2004 with the mission of standing up a supercomputer 100 times more powerful than the leading systems of the day. <https://www.olcf.ornl.gov/>. Cited June 20, 2026.
11. Center Reports - Oak Ridge Leadership Computing Facility. <https://www.olcf.ornl.gov/about-olcf/center-reports/>. Cited June 20, 2026.
12. Home | Argonne Leadership Computing Facility. <https://www.alcf.anl.gov/>. Cited June 20, 2026.
13. Jülich Supercomputing Centre (JSC). <https://www.fz-juelich.de/en/jsc>. Cited June 20, 2026.
14. JSC Service Status. <https://status.jsc.fz-juelich.de/>. Cited June 20, 2026.
15. RIKEN Center for Computational Science RIKEN Website | RIKEN R-CCS. <https://www.r-ccs.riken.jp/en/>. Cited June 26, 2026.
16. Operating Status | RIKEN Center for Computational Science RIKEN Website. <https://www.r-ccs.riken.jp/en/fugaku/opp-status/>. Cited June 26, 2026.
17. HPC Cineca. <https://www.hpc.cineca.it/>. Cited June 26, 2026.
18. V. V. Voevodin, A. S. Antonov, D. A. Nikitenko, et al., “Supercomputer Lomonosov-2: Large Scale, Deep Monitoring and Fine Analytics for the User Community,” *Supercomputing Frontiers and Innovations* **6** (2), 4–11 (2019). doi [10.14529/jsfi190201](https://doi.org/10.14529/jsfi190201).
19. Slurm Workload Manager - REST API Details. https://slurm.schedmd.com/rest.html?utm_source. Cited June 20, 2026.
20. I. S. Yashchenko and A. N. Salnikov, “Software for Predicting the Start Time of a Task by Statistical Analysis of the History of the Task Queue,” *Physics of Elementary Particles and the Atomic Nucleus* **55** (3), 580–584 (2024).

Received
 April 17, 2026

Accepted
 May 26, 2026

Published
 July 8, 2026

Information about the authors

Anastasia V. Papchenko — Bachelor of Physical Sciences, Lomonosov Moscow State University, Faculty of Computational Mathematics and Cybernetics, Leninskie Gory, 1, building 4, 119234, Moscow, Russia.

Dmitry A. Nikitenko — Ph. D., Leading Scientist; Lomonosov Moscow State University, Research Computing Center, Leninskie Gory, 1, building 4, 119234, Moscow, Russia.