

*Московский государственный университет имени М.В.Ломоносова  
Научно-исследовательский вычислительный центр  
Центр коллективного пользования сверхвысокопроизводительными вычислительными ресурсами МГУ  
(Суперкомпьютерный комплекс)*

# *О работе суперкомпьютерного комплекса Московского университета*

*12 марта 2019 г.*

# Суперкомпьютерный ЦКП МГУ



# Суперкомпьютерный ЦКП МГУ

(parallel.ru/cluster)

PARALLEL.RU

Суперкомпьютер "ЛОМОНОСОВ-2"

ЦКП | Начинающим | Суперкомпьютеры | Прикладное ПО | Дополнительно | Информация | О Центре

- О Сервере
- Что нового?
- Компьютеры
- Технологии
- Исследования
- Метакомпьютинг
- Информация
- Конференции
- Центры
- История
- Россия
- Новости
- Общение
- Наш сервис
- ЦКП

Общая характеристика

Пиковая производительность	4,946 Пф
Производительность на тесте Linpack	2,478 Пф
Число вычислительных узлов	1 696
Основной тип процессора	Intel Haswell
Тип ускорителя	Nvidia Tesla

## Регламент доступа к вычислительному комплексу МГУ

- Общие положения
- Регламент доступа
- Правила доступа и использования ресурсов суперкомпьютерного комплекса
- Перерегистрация

### Общие положения

- Ресурсы суперкомпьютерного комплекса Московского университета предназначены для поддержки фундаментальных научных исследований и учебного процесса, а также выполнения специальных проектов, требующих привлечения суперкомпьютерных систем с петафлопсным уровнем производительности.
- Распределение приборного времени и очередность работ на оборудовании ЦКП определяется в соответствии с приоритетами:
  - учебная работа в соответствии с учебным планом,
  - научная работа профессорско-преподавательского состава, научных работников МГУ,
  - научная работа подразделений МГУ в соответствии с планом научных работ,
  - выполнение научных проектов мирового уровня по приоритетным направлениям в рамках Стратегии научно-технологического развития Российской Федерации,
  - оказание услуг сторонним организациям на возмездной основе.
- Работа студентов, аспирантов и сотрудников МГУ осуществляется на безвозмездной основе.
- Использование вычислительных ресурсов производится в удаленном режиме (через Интернет) для различных научных проектов, главным образом, подразделениями МГУ.

### Регламент доступа

- Условия допуска к работе на оборудовании СК:
  - допуск предоставляется только для проведения научной работы, для проведения практических занятий в рамках образовательных программ и для выполнения вычислений, требующих большой вычислительной мощности;
  - допущенные лица обязуются выполнять правила использования СК.

# Суперкомпьютерный ЦКП МГУ в России и в мире

## Текущий рейтинг

29-я редакция от 24.09.2018г.



№	Место	Кол-во CPU/ядер		Производительность (Тфлор/с)		Разработчик
				Linpack	Пиковая	
1	Москва Московский государственный университет имени М.В.Ломоносова 2018 г.	1696/64384	узлов: 1536 (Xeon E5-2697v3 [Acc: Tesla K40M] 2.6 GHz 64 GB RAM) узлов: 160 (Xeon Gold 6126 [Acc: 2x Tesla P100] 2.6 GHz 96 GB RAM) сеть: Infiniband FDR/Infiniband FDR/Gigabit Ethernet	2,478.00	4,946.79	T-Платформы
2	Москва <a href="#">Главный вычислительный центр Федеральной службы по гидрометеорологии и мониторингу окружающей среды</a> 2018 г.	1952/35136	узлов: 976 (2xXeon E5-2697v4 2.3 GHz 128 GB RAM) сеть: Aries/Aries + Infiniband/Aries + Gigabit Ethernet	1,200.35	1,293.00	T-Платформы, CRAY
3	Москва <a href="#">Московский государственный университет имени М.В.Ломоносова</a> 2012 г.	12422/82468	узлов: 4160 (2xXeon 5570 2.93 GHz 12 GB RAM) узлов: 777 (2xXeon E5630 [Acc: 2xTesla X2070] 2.53 GHz 12 GB RAM) узлов: 640 (2xXeon 5670 2.93 GHz 24 GB RAM) узлов: 288 (2xXeon E5630 [Acc: 2xTesla X2070] 2.53 GHz 24 GB RAM) узлов: 260 (2xXeon 5570 2.93 GHz 24 GB RAM) узлов: 40 (2xXeon 5670 2.93 GHz 48 GB RAM) узлов: 30 (2xPowerXCell 8i 3.2 GHz 16 GB RAM) узлов: 4 (4xXeon E7650 2.26 GHz 512 GB RAM) сеть: Infiniband QDR/Gigabit Ethernet/Gigabit Ethernet	901.90	1,700.21	T-Платформы
4	Москва НИЦ "Курчатовский Институт" 2018 г.	1070/21146	узлов: 364 (2xXeon E5-2680v3 2.5 GHz 128 GB RAM) узлов: 148 (2xXeon E5-2650v2 [Acc: 2x Tesla K80] 2.6 GHz 128 GB RAM) узлов: 23 (2xXeon E5-2680v3 [Acc: 3x Tesla K80] 2.5 GHz 128 GB RAM) сеть: Infiniband FDR	755.53	1,100.55	НИЦ "Курчатовский Институт"
5	Санкт-Петербург <a href="#">Суперкомпьютерный центр Санкт-Петербургский политехнический университет</a> 2017 г.	1468/20552	узлов: 623 (2xXeon E5-2680v3 2.5 GHz 128 GB RAM) узлов: 56 (2xXeon E5-2680v3 2.5 GHz 128 GB RAM) узлов: 36 (2xXeon E5-2680v3 2.5 GHz 128 GB RAM) узлов: 8 (2xXeon E5-2680v3 2.5 GHz 128 GB RAM) узлов: 8 (2xXeon E5-2680v3 2.5 GHz 128 GB RAM) узлов: 3 (2xXeon E5-2680v3 2.5 GHz 128 GB RAM) сеть: Infiniband FDR			

	Top50 России	Top500 мира
Ломоносов-2	1	79
Ломоносов	3	487



*Суперкомпьютерный комплекс МГУ –  
ключевой элемент  
национальной суперкомпьютерной инфраструктуры*

*Суперкомпьютерный центр МГУ на сегодня:*

*Пользователи: 2955*

*Научные проекты: 880*

*Факультеты / Институты МГУ : 21*

*Институты РАН : 95*

*ВУЗы России: 102*

*Города России: 50*

# Национальный масштаб ЦКП – СКЦ МГУ

Файл Правка Вид Журнал Закладки Инструменты Справка

Суперкомпьютер «Ломоносов-2»

ckp-rf.ru/usu/439942/

АВТОРИЗАЦИЯ

Современная исследовательская инфраструктура Российской Федерации  
открытость, доступность, инновационность

НОВОСТИ ЦКП УНУ ДОКУМЕНТЫ ОБЩЕНИЕ ОТКРЫТЫЕ ДАННЫЕ БИБЛИОТЕКА КОНТАКТЫ

УНИКАЛЬНЫЕ НАУЧНЫЕ УСТАНОВКИ

НА ГЛАВНУЮ ПОСЛЕДНИЕ НОВОСТИ ВСЕ НОВОСТИ

### Суперкомпьютер «Ломоносов-2»

УНУ создана в 2014 году

**АДРЕС**

Центральный  
119192, г. Москва, Ломоносовский пр-т, д. 27, к. 1  
<http://parallel.ru/cluster>

**РУКОВОДИТЕЛЬ РАБОТ**

Тихонравов Александр Владимирович  
(495) 9395424  
[tikhv@srcc.msu.ru](mailto:tikhv@srcc.msu.ru)

- Открытый конкурс по мероприятию 3.1.1 ФЦП ИнР 2014-2020
- Информационный бюллетень Национальной контактной точки "Исследовательские инфраструктуры"
- Серия вебинаров для ЦКП и базовых организаций
- Ликвидация Испытательного центра

СВЕДЕНИЯ О РЕЗУЛЬТАТИВНОСТИ ЗА 2016 ГОД (ДААННЫЕ МОНИТОРИНГА)

УЧАСТИЕ В МОНИТОРИНГЕ	ЧИСЛО ОРГАНИЗАЦИЙ-ПОЛЬЗОВАТЕЛЕЙ, ЕД.	ЧИСЛО ПУБЛИКАЦИЙ, ЕД.	ЗАГРУЗКА ОРГАНИЗ.
			%

# Суперкомпьютерный комплекс МГУ и конкурс РНФ

Перечень объектов инфраструктуры, на базе которых будут реализовываться проекты, поддержанные по итогам конкурса 2019 года на получение грантов Российского научного фонда по мероприятию «Проведение исследований на базе существующей научной инфраструктуры мирового уровня» Президентской программы исследовательских проектов, реализуемых ведущими учеными, в том числе молодыми учеными

7	<p>Суперкомпьютерный комплекс МГУ имени М.В.Ломоносова, Центр коллективного пользования сверхвысокопроизводительными вычислительными ресурсами Московского университета</p> <p>федеральное государственное бюджетное образовательное учреждение высшего образования "Московский государственный университет имени М.В.Ломоносова", г Москва</p> <p><a href="http://parallel.ru/cluster">http://parallel.ru/cluster</a></p>	<p><a href="#">19-73-20104</a></p> <p><a href="#">19-73-20115</a></p> <p><a href="#">19-73-20032</a></p> <p><a href="#">19-72-20086</a></p> <p><a href="#">19-72-20089</a></p> <p><a href="#">19-73-20030</a></p>
---	--	---



# Конкурс Минобрнауки РФ на развитие ЦКП России (150 млн.руб. на 2017-2018 г.г.)

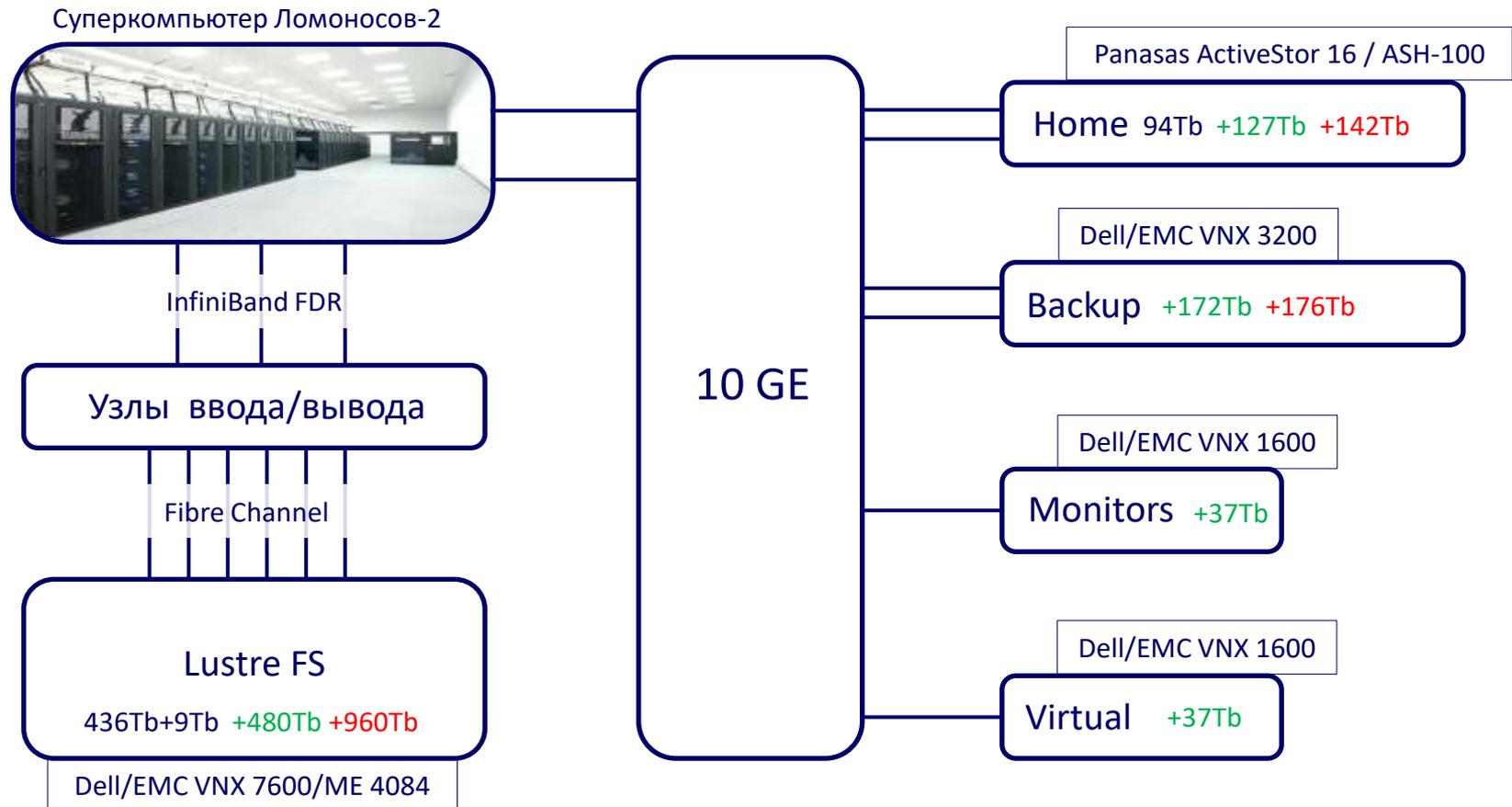
---

Ключевые требования конкурса (Соглашения):

- Число внешних пользователей,
- Доля молодых исследователей,
- Публикации за время действия проекта:
  - дата публикации,
  - публикация в научном журнале (WoS/Scopus),
  - ссылка на Соглашение в тексте публикации по строгому формату,
- Софинансирование проекта,
- ...

Реализация Соглашения: открытие аккредитива 28 декабря...

# Структура системы хранения данных, Ломоносов-2 (лот Минобрнауки РФ, 150 млн.руб.)



Вторая очередь расширения СХД СКЦ МГУ  
будет завершена в мае 2019 года.

# Расширение инфраструктуры ПО, Ломоносов-2

(лот Минобрнауки РФ, 150 млн.руб.)

---

Предварительный опрос пользователей суперкомпьютерного комплекса МГУ.

Заключены контракты на поставку:

Первый этап:

- Schrodinger; MATLAB; MOLPRO; FlowVision.

Второй этап:

- Crystal14 (+upgrade); TeraChem; TotalView; BIOVIA Materials Studio.

PS. Запрет (жесткие специальные условия) на покупку ПО в рамках Соглашения с Минобрнауки РФ:

- CHARMM, Amber.

*Техподдержка и сопровождение СКЦ МГУ*

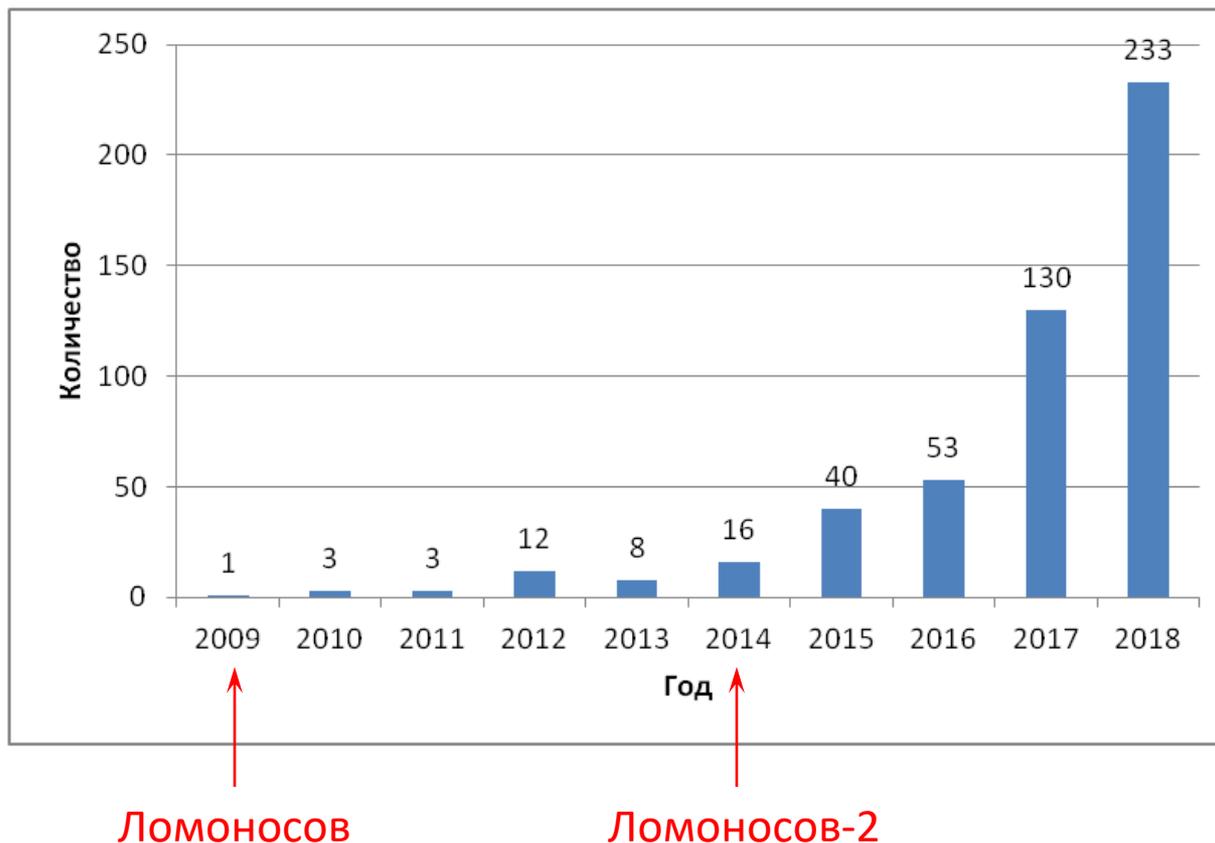
# *Техподдержка и сопровождение СКЦ МГУ*

---

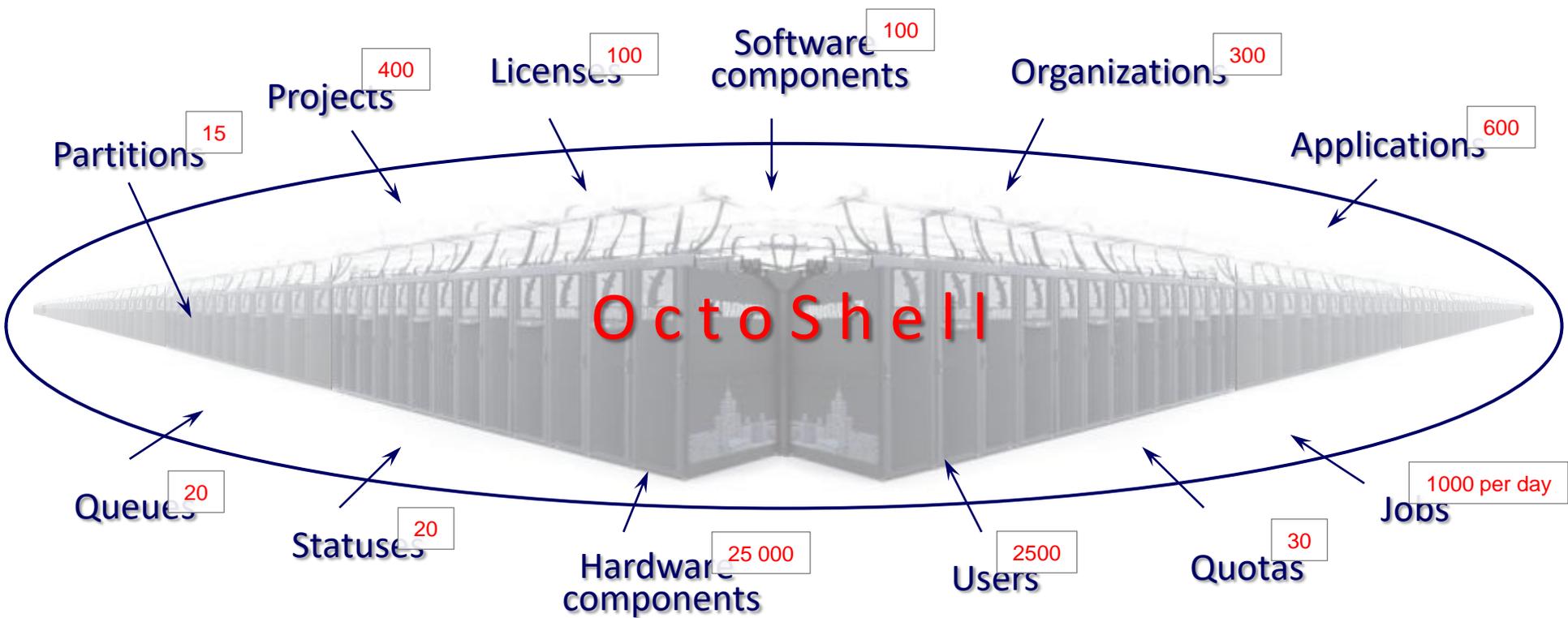
- Нетривиальные масштабы суперкомпьютерного комплекса МГУ.
- Высокая востребованность суперкомпьютерных ресурсов. Значительные масштабы сообщества пользователей СКЦ МГУ. Разнообразие тематик, требований, методов организации вычислительного процесса.
- Специальные технологии обеспечения эффективности СКЦ МГУ.
- Уникальность оборудования.
- Устаревание оборудования.

# Возраст оборудования СКЦ МГУ

Распределение суперкомпьютеров списка Top500 по годам выпуска



# Масштаб суперкомпьютерного комплекса МГУ



# OctoShell: система оперативной поддержки СКЦ МГУ (users.parallel.ru)

The image displays a collage of screenshots from the OctoShell web application, which is used for managing resources and providing support at the Supercomputer Center of Moscow State University. The interface is primarily in Russian and includes several key sections:

- Проекты (Projects):** A page for managing projects, featuring a search bar, filters for status and organization, and a table listing projects with columns for ID, name, status, and user.
- Организации (Organizations):** A page for managing organizations, including a search bar and a table listing organizations with columns for ID, abbreviation, and name.
- Список заявок в поддержку (Support Request List):** A page for managing support requests, featuring a search bar, filters for project, author, and status, and a table listing requests with columns for ID, problem, and status.
- Виды квот на ресурсы (Resource Quota Types):** A table showing different types of resource quotas, with columns for ID, name, and value.
- Логи кластеров (Cluster Logs):** A page for viewing cluster logs, featuring a search bar and a table listing logs with columns for cluster, project, and log content.

# Группа поддержки и сопровождения СКЦ МГУ

(текущая работа: 1 марта 2018 – 1 марта 2019)

---

- Запросов в группу поддержки от пользователей:
  - Всего: 856

из них:

- отработано: 707
- в работе, ожидается реакция от пользователя: 146
- в работе, ожидается реакция от поддержки: 3
- ожидает обработки: 0

# Группа поддержки и сопровождения СКЦ МГУ

## (инфраструктура ПО на Ломоносове-2)

---

Планировщик потока задач:

- SLURM;

MPI:

- OpenMPI - 9 сборок, 4 версии,
- Intel MPI 4 сборок;

Компиляторы:

- Intel 3 версии,
- gcc 3 версии,
- pgc 3 версии,
- clang 2 версии;

Библиотеки:

- CUDA 3 версии,
- MKL 5 сборок, 4 версии;

Анализ эффективности приложений:

- Valgrind;

Контейнеризация:

- Singularity;

Прикладные пакеты:

- Gromacs (2 версии, 3 сборки), VASP (2 версии), FlowVision, Molpro, NAMD, NWChem, Espresso, WIEN2K, Abinit, cp2k, LAMMPS (2 сборки, 2 версии), magma, MATLAB, CHARM, FireFly, Schrodinger, OpenFoam, SPILADY...;

Менеджер пакетов Python:

- anaconda;

Поддержка ML/DL/AI:

anaconda, caffe, Tensorflow, Pytorch, Torchvision, opencv2-python, Keras, numpy, scipy, scikit, singularity, BigDFT, Caffe2, Chainer, Microsoft Cognitive Toolkit, Deep Cognition Studio, DIGITS, H2O Driverless AI, Kenetica, OmniSci, MXNet, NVCaffe, PaddlePaddle, TensorRT Inference Server, TensorRT, Theano, Torch...

# Группа поддержки и сопровождения СКЦ МГУ (инфраструктура ПО на Ломоносове-2 и вопросы лицензирования)

## VASP

С пятницы 15 февраля доступ к пакету VASP будет предоставлен только тем, кто направит запрос. Это нужно сделать через сайт [users.parallel.ru](http://users.parallel.ru) (раздел "пакеты") и далее опять есть две альтернативы:

1. если Вы работаете на основе одной из 10 лицензий МГУ (ситуация 1), то одобрение доступа к пакету будет сделано на основе присланных руководителями групп списков;
2. если Вы работаете на основе другой лицензии (ситуация 2), то для одобрения нужно приложить сканы подтверждающих документов.

Если сейчас Вы не  
доступ по такой

18.02.2019 12:09, Doris.Vogtenhuber<doris.vogtenhuber@univie.ac.at> пишет:

Dear Vladimir Voevodin,  
thank you very much for your efforts to check whether colleagues who run VASP on your supercomputers do have a valid license to do so! We appreciate this very much!

Please let me confirm that Professor Alexandr Saletsky has a valid license to run vasp.5 (including the latest release, vasp.5.4.4) since December 2009 (license number 5-229)

Prof Saletsky and Andriy Klavsyuk are both registered as members of the VASP group.



Dr. Doris Vogtenhuber

# Ежегодные отчеты пользователей СКЦ МГУ

---

- Публикации с указанием использования СКЦ МГУ (приложить сканы).
  - Краткие карточки проекта на русском и английском языках.
  - Проверка и подтверждение актуальности информации о проекте и о его участниках.
  - Краткий опрос по структуре проекта и его особенностях (указание используемого инструментария, технологий программирования, потребности в разного рода образовательных курсах и тренингах и т.п.).
  - Иллюстрации по результатам проекта (в научно-популярном виде).
  - Заявка на следующий год.
- 
- Отчеты по проектам представляют ТОЛЬКО их руководители, с учетом данных от ВСЕХ участников проекта.
- 
- Около 350 содержательных отчетов каждый год...



*Модуль анализа эффективности  
приложений в Октошелле*

# Модуль Октошелла для анализа эффективности. Главная страница

Появилась вкладка «Эффективность» в Октошелл, которая показывает результаты анализа эффективности по всем задачам (только для СК Ломоносов-2)

Octoshell Рабочий кабинет Кабинет администратора Справка vadin@parallel.ru Выход rus eng

Профиль Поддержка Проекты Перерегистрации Пакеты Статистика **Эффективность** Комментарии

### Информация о завершенных задачах

На данной странице вы можете посмотреть информацию о ваших задачах, а также оценить эффективность их выполнения. Для отображения только интересных задач укажите ниже требуемые фильтры.

Показаны задачи: 0..19 из 19

Нет скрытых проблем с производительностью.

Найденные проблемы	ID задачи	Начало счета	Конец счета	Статус	Число узлов	Время счета (часы)	Размер задачи (ЦПУ-часы)	Загрузка ЦПУ	Загрузка ГПУ	Load average	IPC	Получено байт по MPI (МБ/с)	Передано байт по MPI (МБ/с)
☒	697858	2018-07-25 13:12:09	2018-07-25 13:36:56	cancelled	1	0.4	5.8	86.7	0.0	25.4	0.36	0.0	0.0
☒	697859	2018-07-25 13:12:09	2018-07-25 13:36:56	cancelled	1	0.4	5.8	83.9	0.0	24.7	0.35	0.0	0.0
☒	697857	2018-07-25 13:12:06	2018-07-25 13:36:56	cancelled	1	0.4	5.8	55.2	0.0	18.7	0.36	0.0	0.0
☒	697867	2018-07-25 13:40:02	2018-07-25 14:23:28	completed	1	0.7	10.1	34.7	0.0	10.7	0.46	0.0	0.0

Заметный дисбаланс внутри узлов по использованию памяти.

# Базовая информация по отдельной задаче

Информация по задаче № **CENSORED**

## Общая информация

Суперкомпьютер	lomonosov-2
ID задачи	<b>CENSORED</b>
Логин	<b>CENSORED</b>
Статус завершения задачи	COMPLETED
Раздел суперкомпьютера	pascal
Число ядер	24
Число узлов	2
Постановка в очередь	01/11/19 16:17:43
Начало счета	01/11/19 16:17:44
Конец счета	01/11/19 16:36:29
Время счета (часы)	0.3

## Производительность

Метрика	Значение	Общая оценка
Средняя загрузка ЦПУ (%)	1.37	low
Среднее LoadAVG	10.72	good
Среднее IPC		
Средняя загрузка ГПУ (%)	0.00	low
Интенсивность передачи данных по MPI (МБ/с)	0.00	low
Интенсивность чтения из файловой системы (МБ/с)	0.00	low
Интенсивность записи в файловую систему (МБ/с)	0.01	low

## Базовые свойства

Задача имеет необычную локальность обращений в память
В задаче мало MPI коммуникаций
Слишком низкая загрузка процессора при данном уровне loadavg
Слишком низкая интенсивность работы с памятью.

Строка запуска:

**CENSORED SORED**

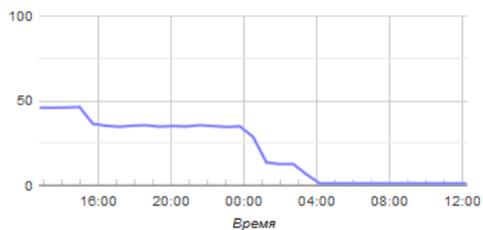
# Данные мониторинга об эффективности выполнения программы

Динамика поведения задачи во время выполнения

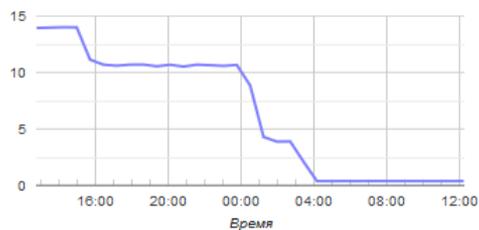
[Более подробное описание](#)

*Еще больше подробностей о динамике выполнения программы!*

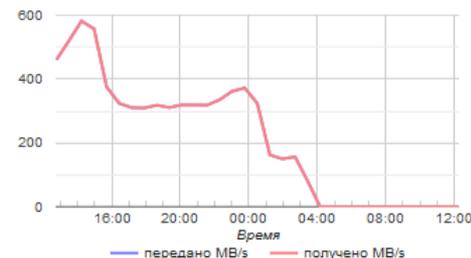
Загрузка ЦПУ (0% .. 100%)



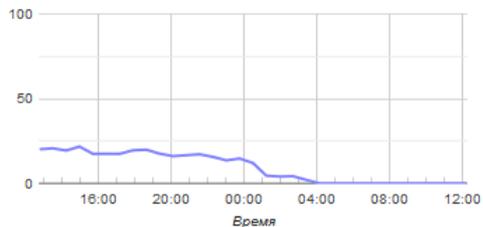
LoadAVG



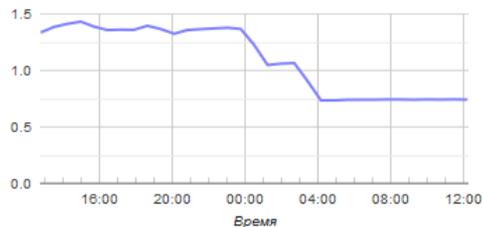
Использование MPI (МБ/с)



Загрузка ГПУ (0% .. 100%)



IPC



Использование файловой системы (МБ/с)



# Информация по найденным потенциальным проблемам с эффективностью

Тип	Описание	Предположение	Рекомендация
   	<p>Задача запущена в разделе для GPU задач, однако практически не использует графические процессоры.</p>	<p>Неправильно выбран раздел для задачи.</p>	<p>Рекомендуется сменить раздел.</p>
   	<p>Задача активно работает с MPI сетью, но сетевая локальность плохая (узлы СК расположены далеко друг от друга).</p>	<p>При запуске задачи менеджером ресурсов был выбран неудачный набор узлов. Рекомендуется при необходимости явно указывать узлы, на которых будет производиться запуск программы, либо пытаться оптимизировать работу с MPI.</p>	<p>Анализ MPI программ -&gt; Профилировка</p>
   	<p>Заметный дисбаланс внутри узлов по использованию памяти.</p>	<p>Такой заметный дисбаланс приводит к снижению эффективности работы программы, поскольку одним узлам приходится ждать другие.</p>	<p>1) Анализ динамики выполнения приложения; 2) Общий анализ последовательных программ; 3) Анализ работы с памятью -&gt; Эффективность.</p>
   	<p>Заметный дисбаланс внутри узлов либо активность сильно отличается на разных узлах.</p>	<p>Разбалансировка нагрузки внутри узла либо между узлами.</p>	<p>1) Анализ динамики выполнения приложения; 2) Общий анализ последовательных программ (если нет MPI); 3) Анализ MPI -&gt; Профилировка (если MPI есть).</p>

# Автоматический анализ задач. Правила

Выполняется анализ всех выполняющихся задач «на лету» с помощью набора правил.

Правило – описание конкретной проблемы с эффективностью. Каждое правило включает:

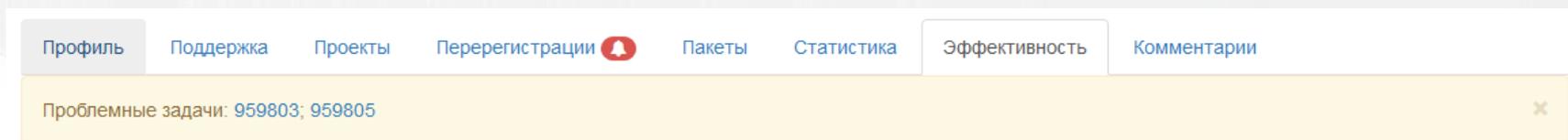
- Описание проблемы (критерий срабатывания).
- Предположение о причинах возникновения.
- Рекомендация по анализу и исправлению.
- Другое (например: оценки критичности проблемы и нашей уверенности в ее наличии).

Реализовано ~30 правил, которые относятся к одной из 9 групп:

1. Зависшая задача.
2. Задача с подозрительно низкой активностью.
3. Дисбаланс загрузки.
4. Некоторые вычислительные ресурсы простаивают (ЦПУ или ГПУ).
5. Некорректное поведение / режим запуска.
6. Неэффективное использование MPI.
7. Малый размер MPI пакетов.
8. Последовательные задачи.
9. Неправильный раздел для запуска.

# Оповещение об аномальных выполняющихся задачах

- Если выполняющаяся задача очень похожа на зависшую или некорректно запущенную – высылаем оповещение
  - Появляется на всех страницах.
  - Появляется до тех пор, пока задач не завершится либо перестанет показывать аномальное поведение.
  - В будущем – оповещение также по электронной почте.



The screenshot shows a navigation bar with the following items: [Профиль](#), [Поддержка](#), [Проекты](#), [Перерегистрации](#) (with a red notification bell icon), [Пакеты](#), [Статистика](#), [Эффективность](#) (highlighted with a white border), and [Комментарии](#). Below the navigation bar is a yellow notification banner with the text: "Проблемные задачи: [959803](#); [959805](#)" and a close button (X) on the right.

## Информация о завершённых задачах

На этой странице вы можете посмотреть информацию о ваших задачах и оценить эффективность их выполнения. Используйте фильтры ниже для отображения только нужных задач.

# Обратная связь

Очень важно получать подтверждение (или опровержение) корректности сделанных выводов!

Найденные проблемы	ID задачи	Начало счета	Конец счета	Статус	Число узлов	Время счета (часы)	Размер задачи (ЦПУ-часы)	Загрузка ЦПУ	Загрузка ГПУ	Load average	IPC	Получено байт по MPI (МБ/с)	Передано байт по MPI (МБ/с)
    <b>CENSORED</b>		2018-07-25 11:03:50	2018-07-25 11:24:13	completed	1	0.3	 4.8	1.0	0.0	0.1	0.23	0.0	0.0

Не показывать больше

-     Задача запущена в разделе для GPU задач, однако практически не использует графические процессоры.
-     Заметный дисбаланс внутри узлов по использованию памяти.
-     Низкая активность использования всех доступных ресурсов (процессоров, памяти, коммуникационной сети, графических ускорителей).
-      Чрезвычайно низкая активность использования всех доступных ресурсов.

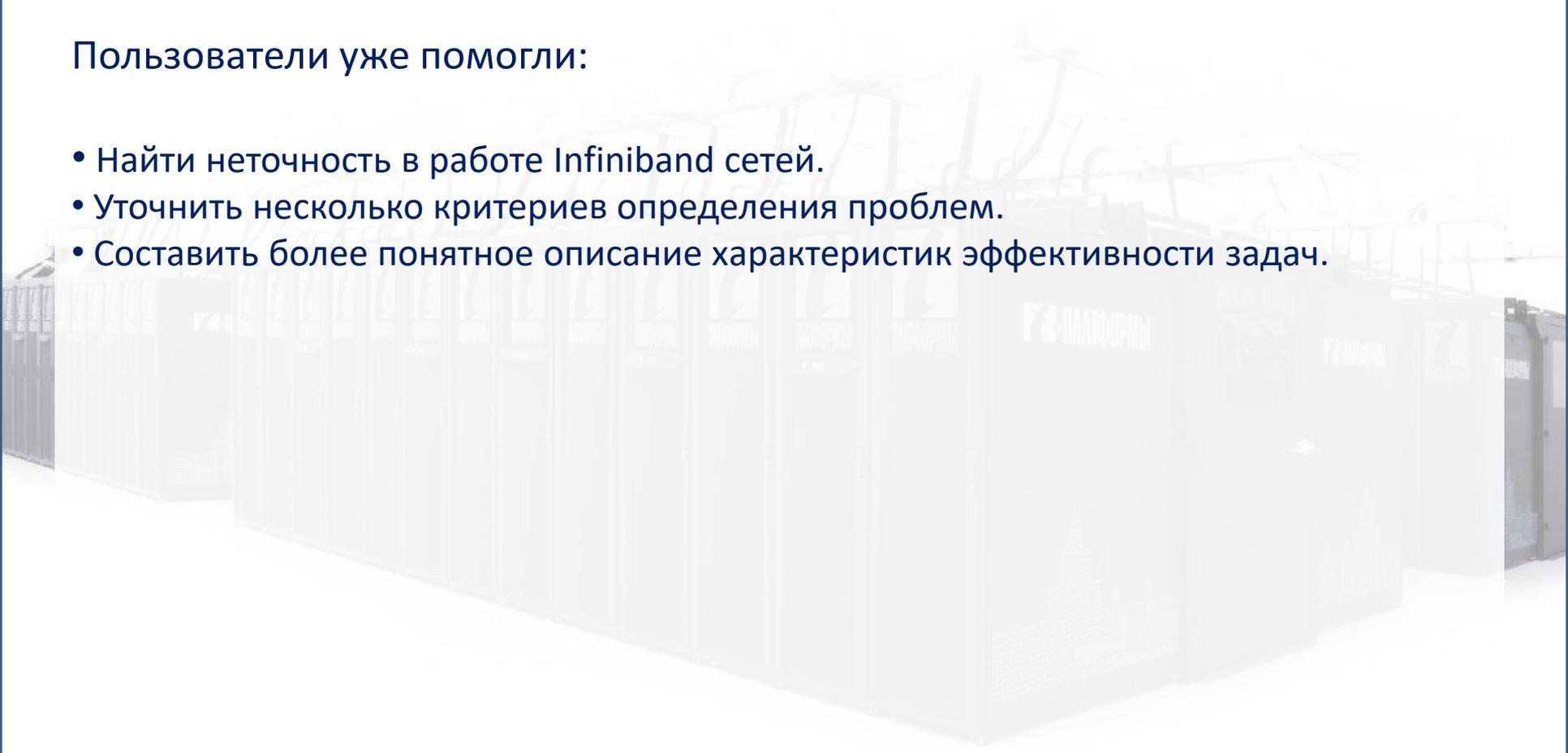
Согласен!

Не согласен! (важно указать: что не так?)

# Примеры обратной связи

Пользователи уже помогли:

- Найти неточность в работе Infiniband сетей.
- Уточнить несколько критериев определения проблем.
- Составить более понятное описание характеристик эффективности задач.



*Примеры: интересные и не очень...*

# Пример: влияние сетевой локальности

Два одинаковых запуска:

- Все параметры запуска одинаковы
- Но время выполнения и интенсивность использования MPI отличается

Найденные проблемы	ID задачи	Начало счета	Конец счета	Статус	Число узлов	Время счета, в часах	Размер задачи (ЦПУ-часы)	Загрузка ЦПУ	Загрузка ГПУ	Load average	IPC	Получено байт по MPI, в МБ
		2018-10-20 07:37:52	2018-10-21 11:46:24	completed	24	28.1	 9455.8	47.1	0.0	14.0	1.68	550.0
		2018-10-20 11:11:53	2018-10-21 17:18:01	completed	24	30.1	 10114.3	46.9	0.0	14.0	1.71	514.2

# Пример: влияние сетевой локальности

Два одинаковых запуска:

- Все параметры запуска одинаковы
- Но время выполнения и интенсивность использования MPI отличается
- **Причина:** плохая сетевая локальность (6 коммутаторов вместо 3)

Найденные проблемы	ID задачи	Начало счета	Конец счета	Статус	Число узлов	Время счета, в часах	Размер задачи (ЦПУ-часы)	Загрузка ЦПУ	Загрузка ГПУ	Load average	IPC	Получено байт по MPI, в МБ
	CENSORED	2018-10-20 07:37:52	2018-10-21 11:46:24	completed	24	28.1	 9455.8	47.1	0.0	14.0	1.68	550.0
	CENSORED	2018-10-20 11:11:53	2018-10-21 17:18:01	completed	24	30.1	 10114.3	46.9	0.0	14.0	1.71	514.2

ⓧ

🔇 🗣️ 👍 🔄 Задача активно работает с MPI сетью, но сетевая локальность плохая (узлы СК расположены далеко друг от друга).

🔇 🗣️ 👍 🛑 Задача запущена в разделе для GPU задач, однако практически не использует графические процессоры.

🔇 🗣️ 👍 🛑 Слишком маленькие средние размеры ФС IB пакетов при достаточно высокой интенсивности использования коммуникационной сети.

# Пример: неэффективная работа с MPI

## Общая информация

Суперкомпьютер	lomonosov-2
ID задачи	<b>CENSORED</b>
Логин	<b>CENSORED CENSORED</b>
Статус завершения задачи	COMPLETED
Раздел суперкомпьютера	compute
Число ядер	224
Число узлов	16
Постановка в очередь	02/16/19 22:03:00
Начало счета	02/17/19 22:25:50
Конец счета	02/19/19 04:47:52
Время счета (часы)	30.4

## Производительность

Метрика	Значение	Общая оценка
Средняя загрузка ЦПУ (%)	39.54	average
Среднее LoadAVG	14.14	good
Среднее IPC	1.97	good
Средняя загрузка ГПУ (%)	0.02	low
Интенсивность передачи данных по MPI (МБ/с)	378.71	good
Интенсивность чтения из файловой системы (МБ/с)	0.00	low
Интенсивность записи в файловую систему (МБ/с)	0.18	low

## Базовые свойства

Задача хорошо использует ресурсы СК
Задача активно работает с памятью
В задаче много MPI коммуникаций
Задача имеет низкую локальность обращений в память

Строка запуска: `/opt/mpl/wrappers/ompi` **CENSORED CENSORED CENSORED**

## Найденные потенциальные проблемы с эффективностью

В данном разделе приведен список проблем с эффективностью, которые были найдены для данной задачи. Для каждой проблемы приведено ее описание (какой признак возникновения проблемы был обнаружен), предположение (в чем, на наш взгляд, может заключаться причина возникновения проблемы) и рекомендация (что мы советуем сделать для ее устранения). Во многих случаях в рекомендации указано, какой тип дальнейшего анализа стоит проводить; реализацию этого функционала планируется выполнить в будущем.

Возможность оценки корректности и/или изменения обнаруженных проблем с эффективностью также планируется добавить в ближайшем будущем. Сейчас эта возможность реализована только на общей странице со списком задач.

Тип	Описание	Предположение	Рекомендация
	Задача запущена в разделе для GPU задач, однако практически не использует графические процессоры.	Неправильно выбран раздел для задачи.	Рекомендуется сменить раздел.
	Слишком маленькие средние размеры MPI IB пакетов при достаточно высокой интенсивности использования коммуникационной сети.	Накладные расходы на передачу MPI сообщений могут быть существенными.	Анализ MPI программ -> Профилировка (если проблема с MPI сетью)
	Задача активно работает с MPI сетью, но сетевая локальность плохая (узлы СК расположены далеко друг от друга).	При запуске задачи менеджером ресурсов был выбран неудачный набор узлов. Рекомендуется при необходимости явно указывать узлы, на которых будет производиться запуск программы, либо попытаться оптимизировать работу с MPI.	Анализ MPI программ -> Профилировка

# Пример: аномальное поведение задачи

## Общая информация

Суперкомпьютер	lomonosov-2
ID задачи	<b>CENSORED</b>
Логин	<b>CENSORED</b>
Статус завершения задачи	COMPLETED
Раздел суперкомпьютера	comprite
Число ядер	336
Число узлов	24
Постановка в очередь	10/11/18 18.02:15
Начало счета	10/11/18 19:35:17
Конец счета	10/12/18 05:20:15
Время счета (часы)	9.7

Строка запуска **CENSORED CENSORED CENSORED CENSORED CENSORED**

## Производительность

Метрика	Значение	Общая оценка
Средняя загрузка ЦПУ (%)	98.57	good
Среднее LoadAVG	1819.40	low
Среднее IPC	1.15	good
Средняя загрузка ГПУ (%)	0.00	low
Интенсивность передачи данных по MPI (МБ/с)	0.00	low
Интенсивность чтения из файловой системы (МБ/с)	0.00	low
Интенсивность записи в файловую систему (МБ/с)	0.00	low

## Базовые свойства

В задаче мало MPI коммуникаций
Слишком низкая загрузка процессора при данном уровне loadavg
Задача активно работает с памятью
Задача имеет низкую локальность обращений в память

**КРАЙНЕ МНОГО!**

## Найденные потенциальные проблемы с эффективностью

В данном разделе приведен список проблем с эффективностью, которые были найдены для данной задачи. Для каждой проблемы приведено ее описание (какой признак возникновения проблемы был обнаружен), предположение (в чем, на наш взгляд, может заключаться причина возникновения проблемы) и рекомендация (что мы советуем сделать для ее устранения). Во многих случаях в рекомендации указано, какой тип дальнейшего анализа стоит проводить, реализацию этого функционала планируется выполнить в будущем.

Возможность оценки корректности и/или изменения обнаруженных проблем с эффективностью такие планируется добавить в ближайшем будущем. Сейчас эта возможность реализована только на общей странице со списком задач.

Тип	Описание	Предположение	Рекомендация
	Задача запущена в разделе для GPU задач, однако практически не использует графические процессоры.	Неправильно выбран раздел для задачи.	Рекомендуется сменить раздел.
	Задача выполняется аномально неэффективно.	Задача работает некорректно или зависла.	Рекомендуется проверить корректность запуска и при необходимости отменить его.

*Немного статистики с  
суперкомпьютера Ломоносов-2...*

# Использование узлов Ломоносова-2



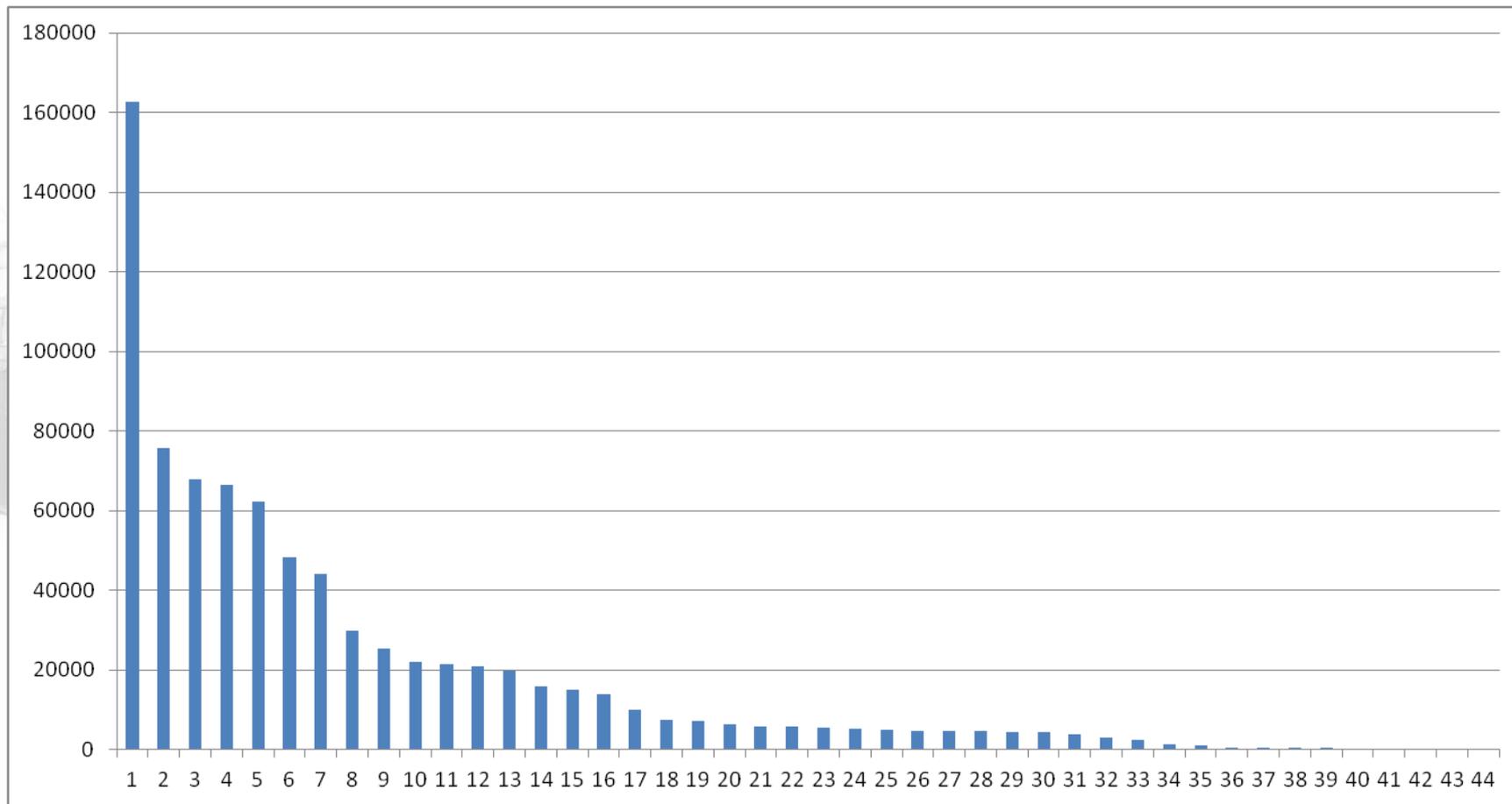
Период: 01.10.18 – 18.12.18.

Доля узлов **Ломоносова-2** со значением CPUload ниже 10%. Несколько значений в одном столбце представляют изменение значений в течение данного дня (данные считались каждый час).

## *Статистика на Ломоносов-2 по найденным потенциальным проблемам*

- Начали собирать с лета 2018 года
- Аномально неэффективные задачи (почти все из них – некорректные запуски или сбои):
  - август 2018 – н.в.: 71 пользователь, 668 задачи, ~258К узлочасов
- Задачи с низкой загрузкой ресурсов (ЦПУ, ГПУ, памяти, I/O):
  - апрель 2018 – н.в.: 127 пользователей, 5151 задач, ~500К узлочасов

# Распределение аномальных приложений по пользователям



Период: 09.10.18 – 20.12.18.

Показан объем процессорно-часов, которые были потрачены аномальными задачами отдельных пользователей Ломоносова-2

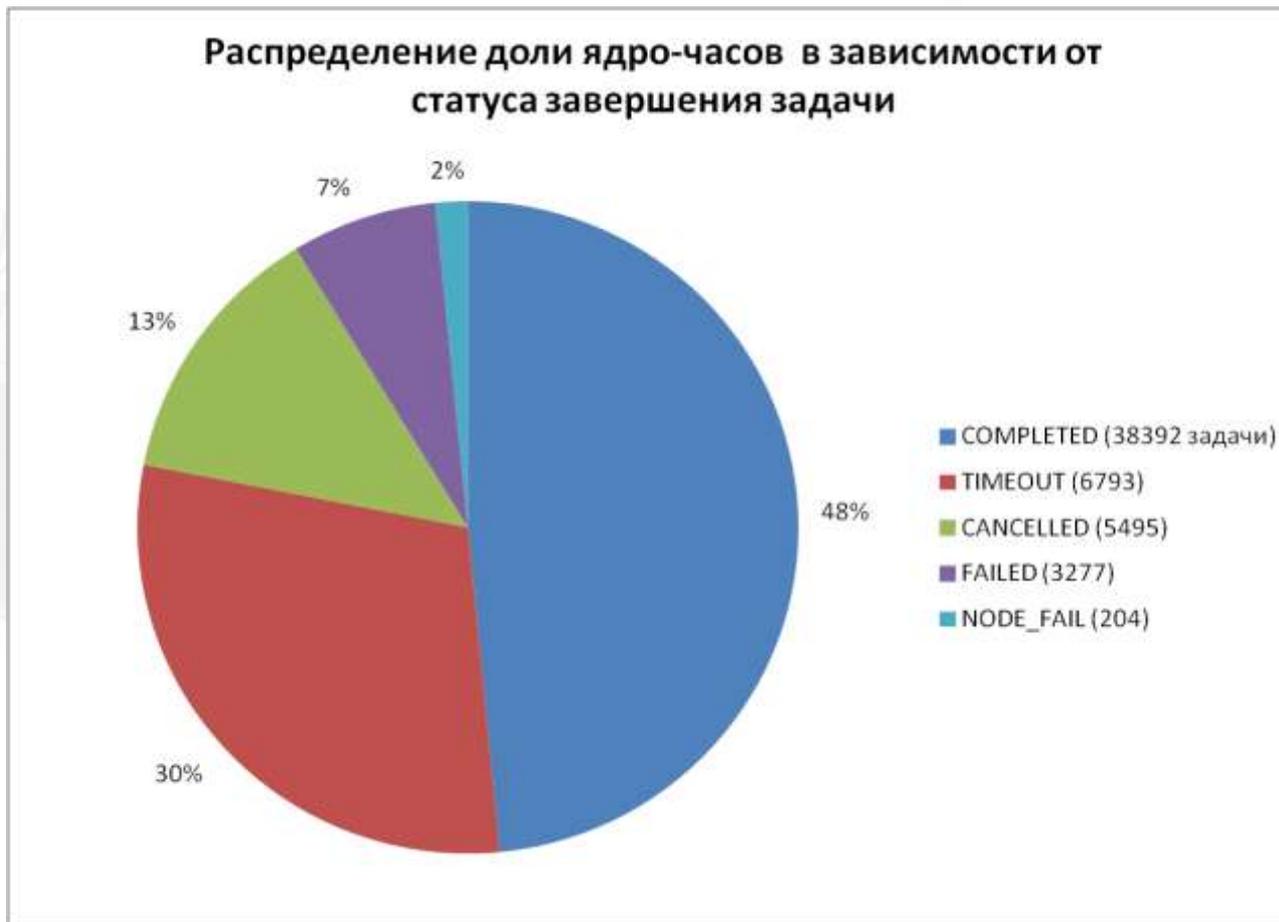
# Статистика на Ломоносов-2 по найденным потенциальным проблемам

- Наблюдается немало и других потенциальных проблем, например:

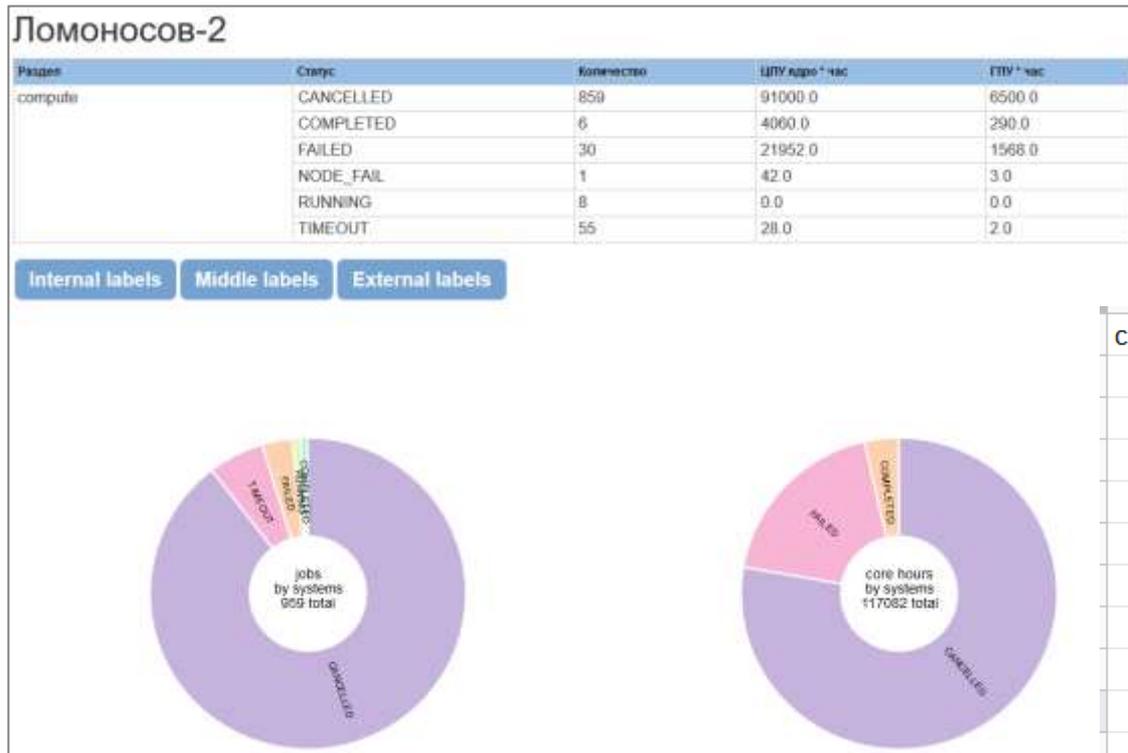
Тип проблемы	Число задач
<b>МРІ</b>	
Плохая сетевая локальность	25 642
МРІ пакеты слишком маленькие	18 317
ФС пакеты слишком маленькие	4 506
<b>Память</b>	
Дисбаланс в использовании памяти	4 655
Низкая локальность	1 826
<b>Общее</b>	
Слишком мало активных процессов на узел	3 517
...	

# Статистика по статусам завершения задач

Ломоносов-2 с начала 2018 года:



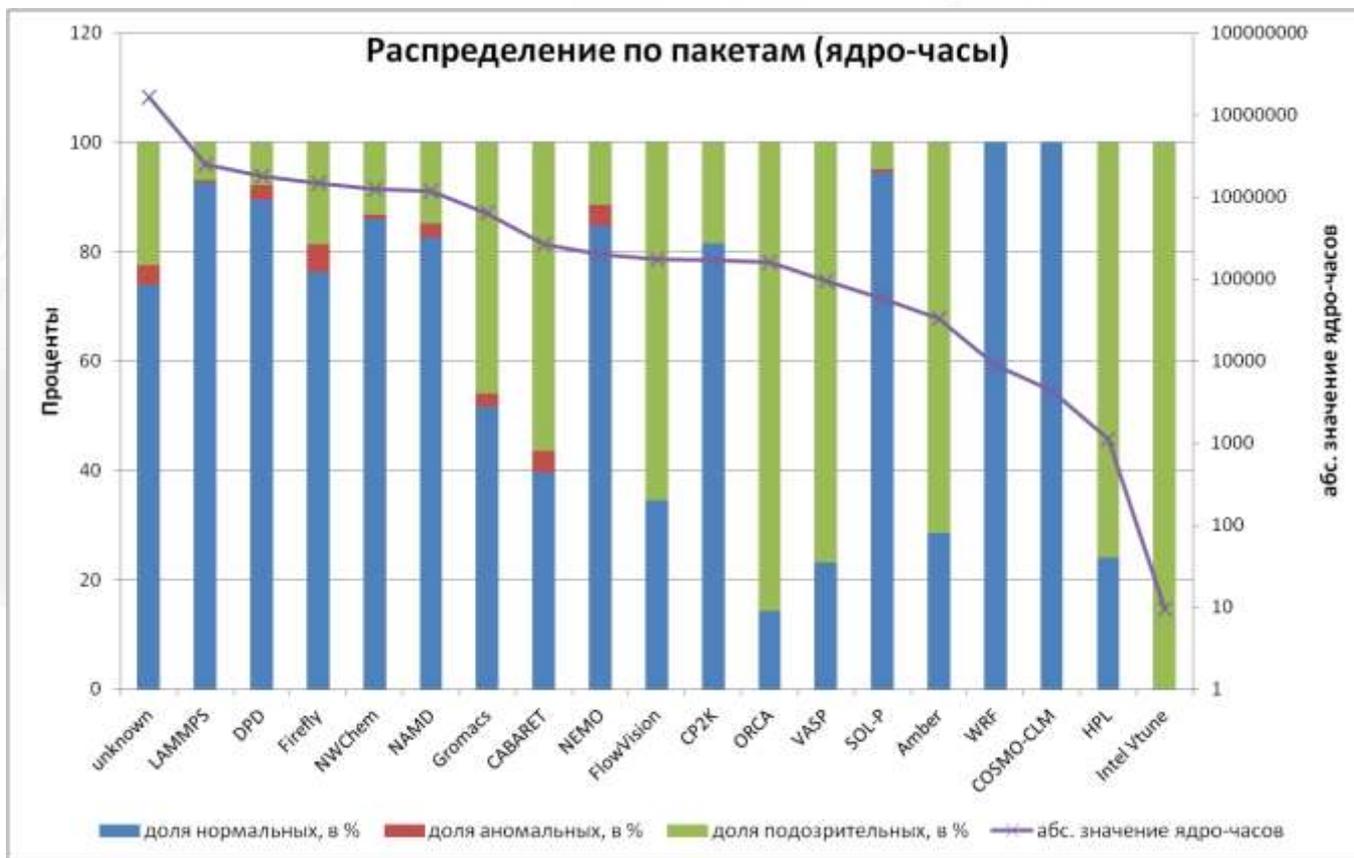
# Анализ статуса завершения приложений конкретного пользователя



count	date	nodes*hours
3	2019-03-11	26.37
4	2019-03-10	102.48
5	2019-03-09	45.92
8	2019-03-07	37.56
7	2019-03-06	326.36
7	2019-03-05	463.16
6	2019-03-04	510.26
3	2019-02-28	68.39
7	2019-02-26	47.52
6	2019-02-20	72.90
3	2019-02-19	152.16
3	2019-02-18	196.10
3	2019-02-15	396.81
3	2019-02-13	255.43

# Статистика по используемым пакетам (XALT)

Ломоносов-2, с 9 октября по 20 декабря 2018:



# Группа поддержки и сопровождения СКЦ МГУ

---



+ коллеги из компании “Т-Платформы”.

*Московский государственный университет имени М.В.Ломоносова  
Научно-исследовательский вычислительный центр  
Центр коллективного пользования сверхвысокопроизводительными вычислительными ресурсами МГУ  
(Суперкомпьютерный комплекс)*

*О работе суперкомпьютерного комплекса  
Московского университета*

*12 марта 2019 г.*