



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Цифровой блок НИУ ВШЭ

СУПЕРКОМПЬЮТЕР «CHARISMA» НИУ ВШЭ

СИСТЕМА АВАРИЙНОГО ОТКЛЮЧЕНИЯ СУПЕРКОМПЬЮТЕРА КАК ПРИМЕР СКРИПТОВОЙ АВТОМАТИЗАЦИИ

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

Собрание системных администраторов суперкомпьютерных центров России
Москва, 08.06.2021



ХАРАКТЕРИСТИКИ СУПЕРКОМПЬЮТЕРА «сCHARISMA»

(COMPUTER OF HSE FOR ARTIFICIAL INTELLIGENCE AND SUPERCOMPUTER MODELLING)

- **6 место в ТОП 50 СНГ**
- Пиковая производительность: **1 Петафлопс** (1 квадриллион операций в секунду над числами с двойной точностью).
- LINPACK-производительность: **653,7 Терафлопс**.
- **40** вычислительных узлов
 - 10 узлов с **1,5 ТБ ОЗУ**, 4 GPU **V100 32 ГБ**,
 - 19 узлов с **768 ГБ ОЗУ**, 4 GPU **V100 32 ГБ**,
 - 11 узлов с **384 ГБ ОЗУ** для расчётов на **CPU**.
- **2** управляющих узла
- **116** графических процессора **NVIDIA Tesla V100 32 ГБ**
- **80** центральных процессоров **Intel Xeon Gold**
- Оперативная память: **32 ТБ**
- Дисковая память: **892 ТБ**
 - параллельная СХД на базе Lustre: **840 ТБ**
 - SSD в узлах: **2x240 ГБ** в RAID1
 - SAS HDD в управляющих узлах: **20x1,7 ТБ**
- Коммуникационная сеть: **2 x InfiniBand EDR**
(**2x100 Гбит/с**, топология **Fat Tree**)





СИСТЕМА АВАРИЙНОГО ОТКЛЮЧЕНИЯ СУПЕРКОМПЬЮТЕРА

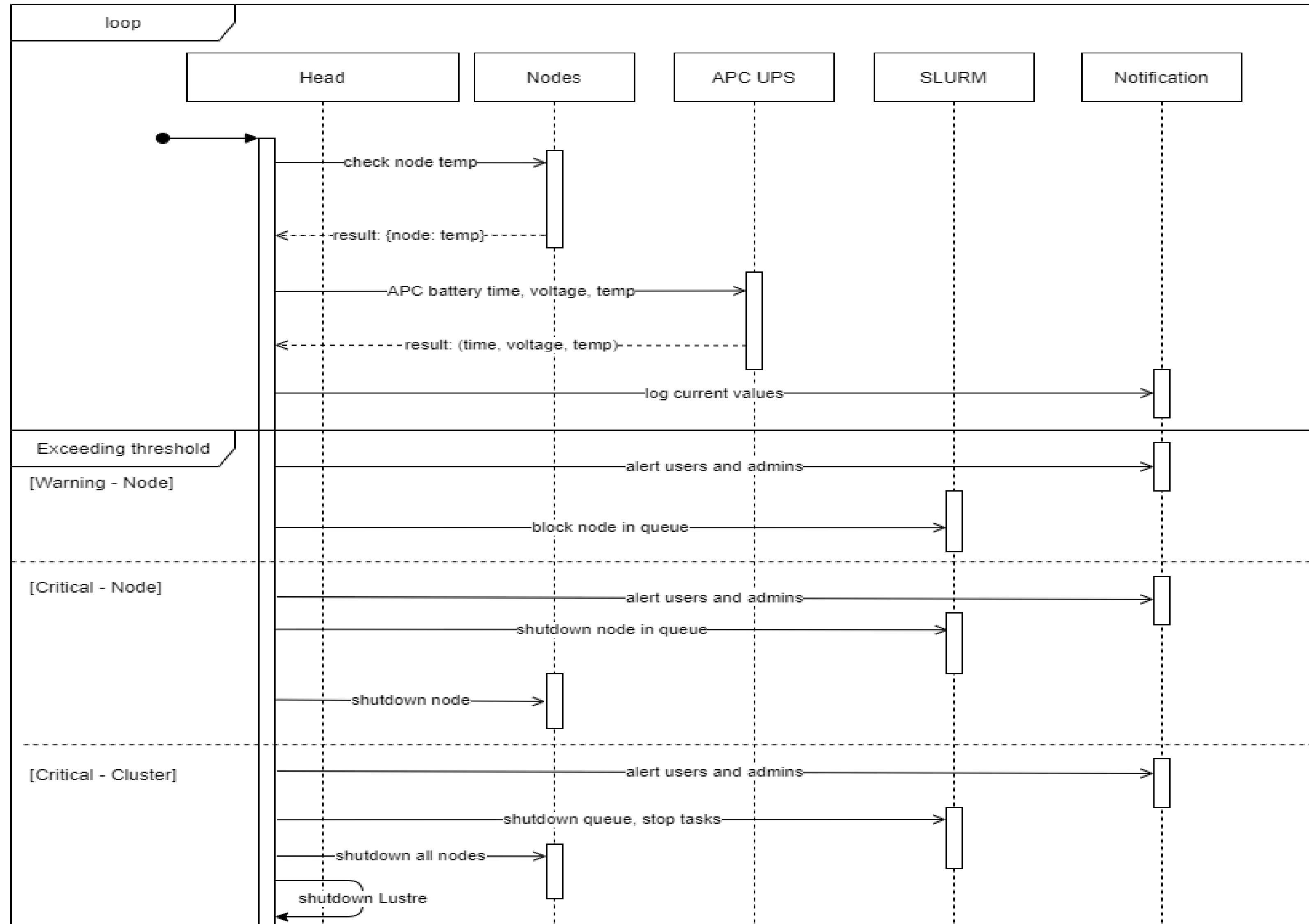
События, требующие незамедлительного реагирования:

- Выход из строя системы охлаждения вычислительного узла,
- Выход из строя системы кондиционирования воздуха в помещении,
- Переключение ИБП на режим работы от аккумуляторных батарей,
- Низкий заряд аккумуляторных батарей ИБП.





СИСТЕМА АВАРИЙНОГО ОТКЛЮЧЕНИЯ СУПЕРКОМПЬЮТЕРА





СИСТЕМА АВАРИЙНОГО ОТКЛЮЧЕНИЯ СУПЕРКОМПЬЮТЕРА

Система аварийного отключения выполняет асинхронные запросы и проверяет:

- Температуру на входе вычислительного узла (REST API RedFish) – большая температура свидетельствует об общей неисправности системы кондиционирования воздуха в помещении.
- Температуру на выходе вычислительного узла (REST API RedFish) – большая температура свидетельствует о неисправности системы охлаждения конкретного узла.
- Состояние ввода ИБП и оставшееся время работы от батарей (SNMP) – отсутствие напряжения на входе и низкий заряд аккумуляторов говорит о возможной аварии в системе электроснабжения.



СИСТЕМА АВАРИЙНОГО ОТКЛЮЧЕНИЯ СУПЕРКОМПЬЮТЕРА

При незначительном превышении заданных параметров:

- Оповещаются администраторы (Telegram, email).
- «Заболевшие» узлы блокируются в очереди (если автоматически заблокировано более 3 узлов – останавливается очередь).
- Сохраняется список пользователей, выполнение задач которых может быть прервано для последующего оповещения.

При критическом превышении заданных параметров:

- Отправляются команды на аварийное отключение вычислительных узлов с небольшой случайной задержкой между ними, для снижения резкого перепада потребляемой мощности ИБП
- Отключаются управляемые коммутаторы InfiniBand и Ethernet (кроме IPMI-коммутатора)



СИСТЕМА АВАРИЙНОГО ОТКЛЮЧЕНИЯ СУПЕРКОМПЬЮТЕРА

В дополнение к системе аварийного отключения суперкомпьютера, на вычислительных узлах работает встроенный модуль контроля температур:

Temperature Probes

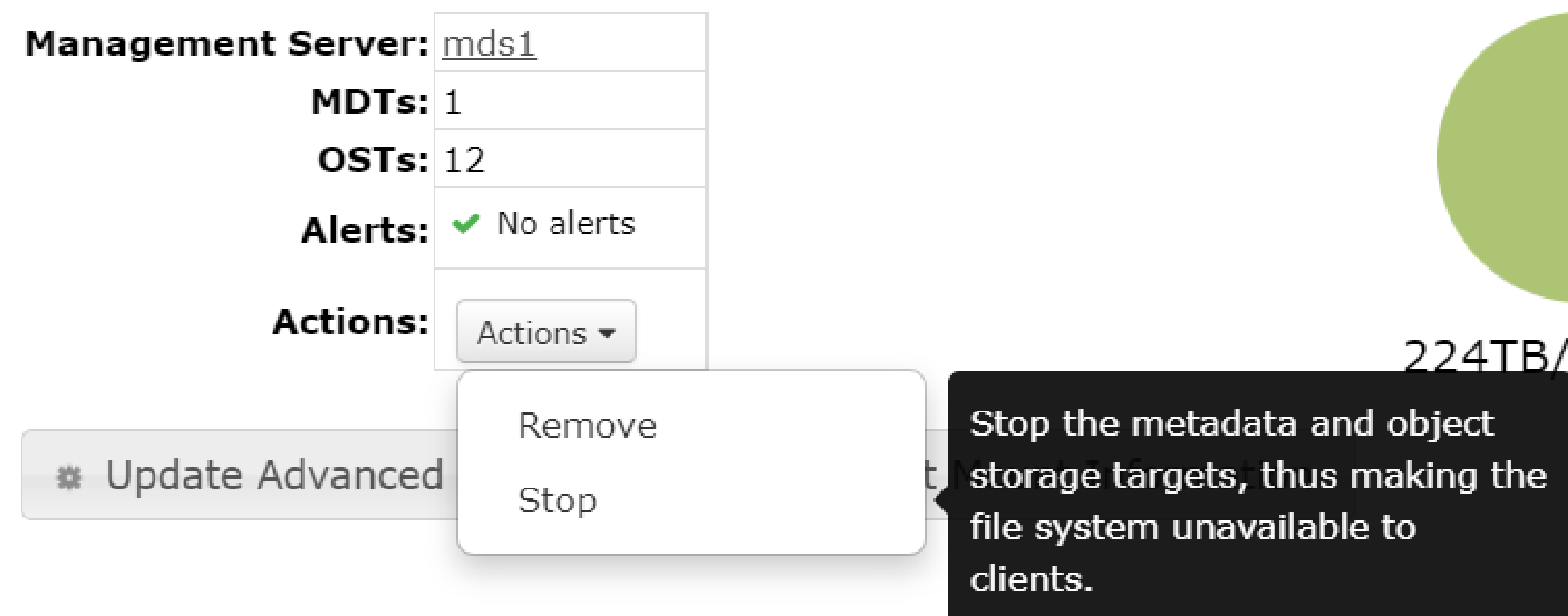
			Warning Threshold		Critical Threshold	
Status	Probe Name	Reading	Min	Max	Min	Max
✓	CPU1 Temp	68 °C (154.4 °F)	N/A	N/A	3 °C (37.4 °F)	98 °C (208.4 °F)
✓	CPU2 Temp	61 °C (141.8 °F)	N/A	N/A	3 °C (37.4 °F)	98 °C (208.4 °F)
✓	System Board Exhaust Temp	42 °C (107.6 °F)	8 °C (46.4 °F)	75 °C (167 °F)	3 °C (37.4 °F)	80 °C (176 °F)
✓	System Board Inlet Temp	16 °C (60.8 °F)	3 °C (37.4 °F) Edit	28 °C (82.4 °F) Edit	-7 °C (19.4 °F)	32 °C (89.6 °F)
✓	System Board SLOT4 Temp	55 °C (131 °F)	N/A	N/A	N/A	N/A
✓	System Board SLOT5 Temp	61 °C (141.8 °F)	N/A	N/A	N/A	N/A
✓	System Board SLOT6 Temp	45 °C (113 °F)	N/A	N/A	N/A	N/A
✓	System Board SLOT7 Temp	59 °C (138.2 °F)	N/A	N/A	N/A	N/A



СИСТЕМА АВАРИЙНОГО ОТКЛЮЧЕНИЯ СУПЕРКОМПЬЮТЕРА

В планах:

- Разработать возможность корректной остановки файловой системы Lustre и её управляющих серверов – неправильное отключение капризной ФС может привести к потере данных.
- Настроить отключение электропитания стоек.
- Поиск возможности удаленного отключения полок СХД.





РАЗРАБОТАНА СОБСТВЕННАЯ СИСТЕМА МОНИТОРИНГА

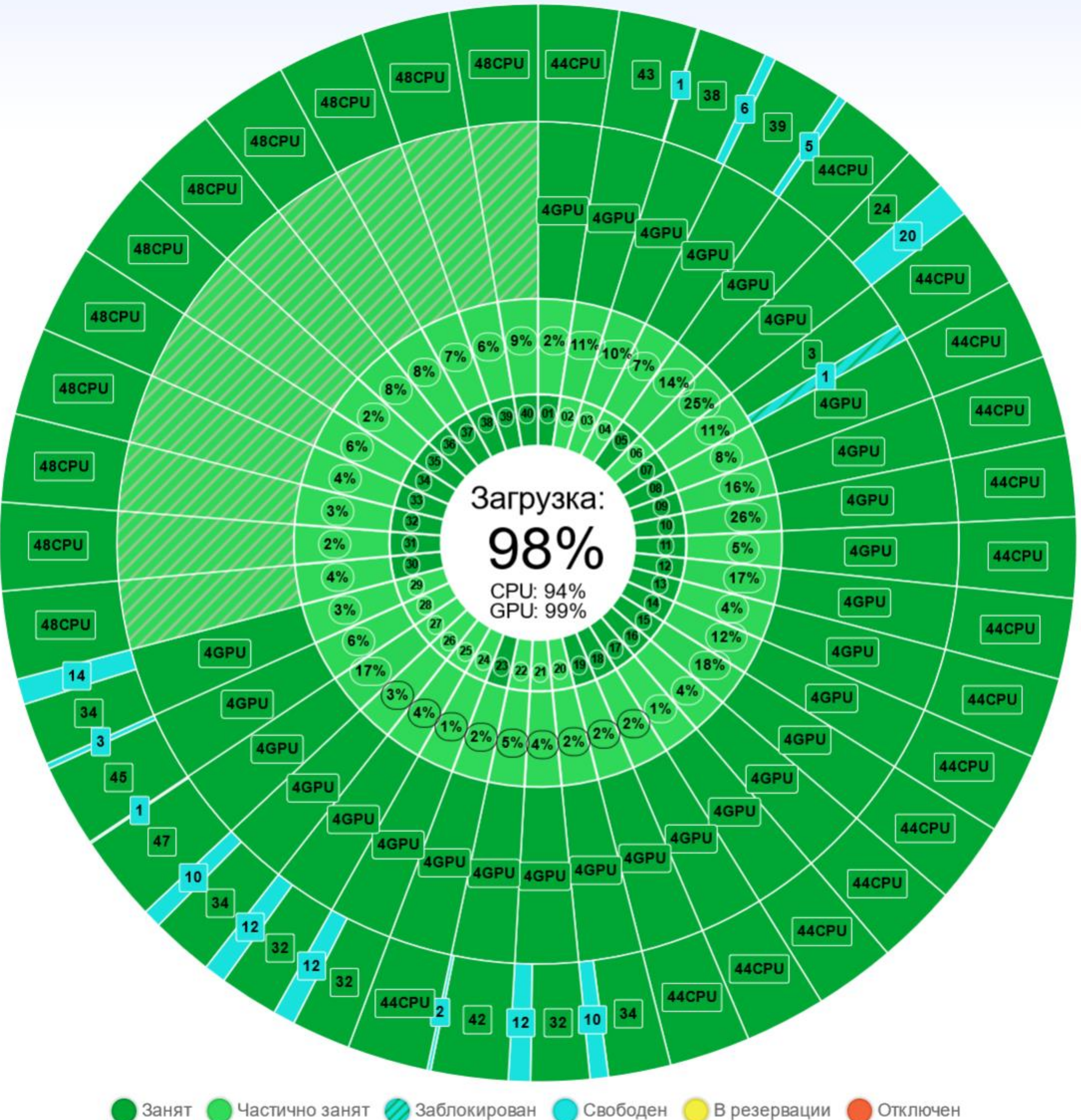
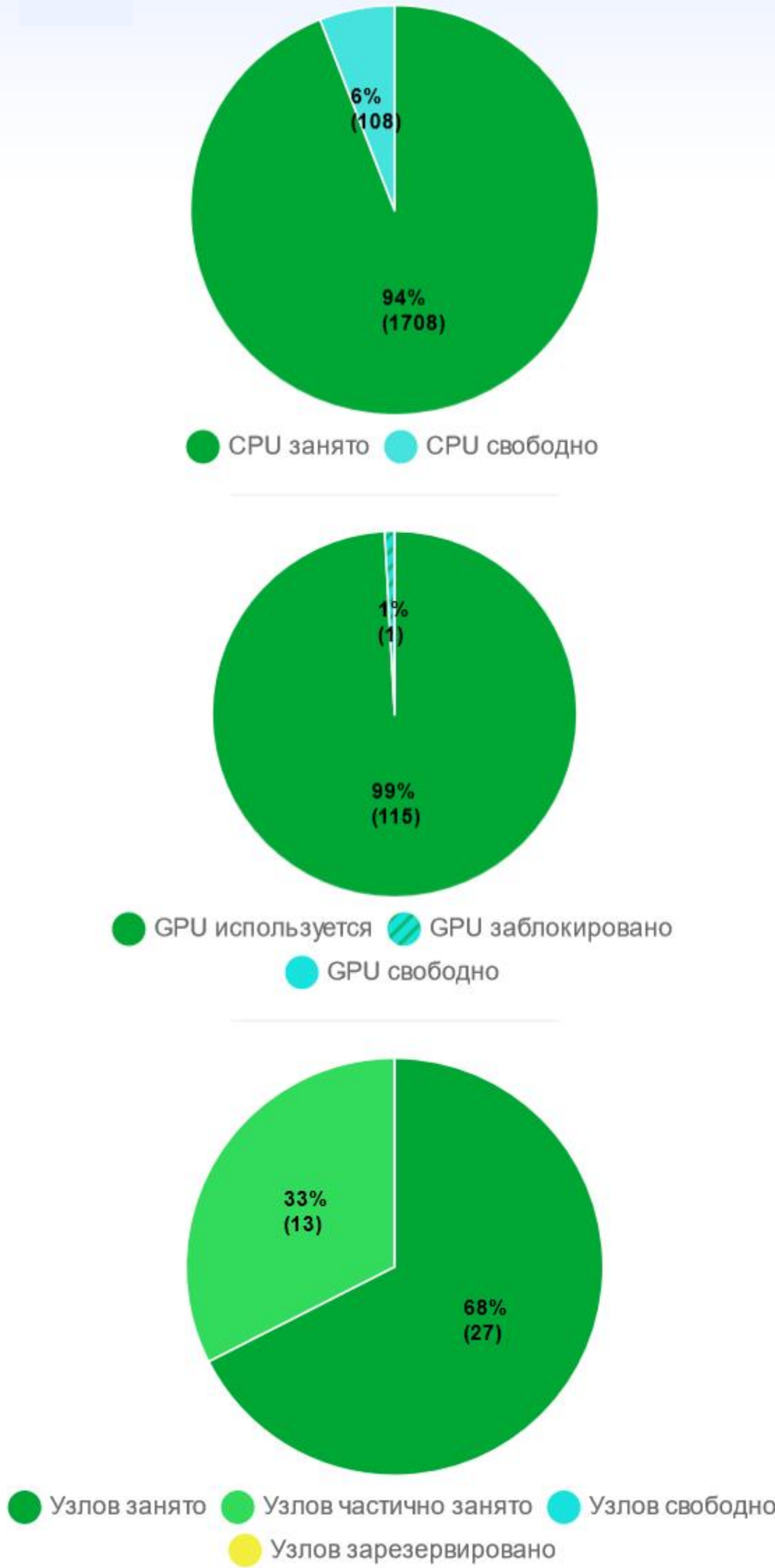
Загрузка суперкомпьютерного комплекса НИУ ВШЭ

13:40:35

Сейчас считают 40 чел.:



Загрузка CPU Загрузка GPU





HPC TASKMASTER - СИСТЕМА МОНИТОРИНГА ЭФФЕКТИВНОСТИ ЗАДАЧ НА СУПЕРКОМПЬЮТЕРЕ

Сводка данных

МОИ ВЫЧИСЛЕНИЯ

Список задач

Статистика

МОИ ПРОЕКТЫ

Список проектов

СЕРВИСЫ

Jupyter Notebook

Генератор sbatch

УПРАВЛЕНИЕ

Настройки доступа

Отдел суперкомпьютерного моделирования НИУ ВШЭ

Личный кабинет пользователя суперкомпьютерного комплекса

Костенецкий Павел Сергеевич

Сводка данных

Внимание

Личный кабинет пользователя находится в разработке. Многие функции ещё недоступны или могут работать неправильно.

Логин-сервер: ОК

Очередь задач: ОК

Файловое хранилище: ОК

Задачи пользователя

Всего: 100

Успешно завершено: 37

Выполняется: 0

Очередь задач SLURM

Выполняется: 141

Ожидает: 0

Запрошено ресурсов: 0 CPU, 0 GPU

Срок действия аккаунта

до 31.12.2099

75%

Использование квоты CPU

Израсходовано 15000 из 20000 ядро-часов

25%

Использование квоты GPU

Израсходовано 2500 из 10000 часов GPU

Дисковая квота

Всего: 1.33 TB

Занято: 14.13 MB

Доступно: 1.33 TB

Сводка данных

МОИ ВЫЧИСЛЕНИЯ

Список задач

Завершённые

Активные

Статистика

МОИ ПРОЕКТЫ

Список проектов

СЕРВИСЫ

Jupyter Notebook

Генератор sbatch

УПРАВЛЕНИЕ

Настройки доступа

Отдел суперкомпьютерного моделирования НИУ ВШЭ

Личный кабинет пользователя суперкомпьютерного комплекса

Костенецкий Павел Сергеевич

Мои задачи

Включить поиск

User	Job ID	Job Name	Start Time	End Time	Nodes	Cores	GPUs	Problems	View
rchulkevich	317449	wrap	07.06.2021 14:49:04	07.06.2021 14:51:04	1	1	0		
rchulkevich	317448	wrap	07.06.2021 14:44:21	07.06.2021 14:46:21	1	1	0		
rchulkevich	317443	bash	07.06.2021 14:32:57	07.06.2021 14:36:57	1	1	0		
vkarpichev	317442	qit_dti.sh	07.06.2021 14:14:55	07.06.2021 14:31:55	1	8	1		
vkarpichev	317441	qit_dti.sh	07.06.2021 14:14:06	07.06.2021 14:36:06	1	8	1		
ybudkov	317439	wrap	07.06.2021 13:44:18	07.06.2021 13:55:18	1	1	0		
mzmeev	317437	mzmeev_md_aida	07.06.2021 13:19:54	07.06.2021 13:56:55	1	10	1		
dsgurev	317435	AttentionGAN_training	07.06.2021 13:16:37	07.06.2021 13:17:37	1	8	4		
iakarpov	317433	ois_cls	07.06.2021 12:47:48	07.06.2021 13:59:49	1	1	4		
adkugusheva	317432	MobNet	07.06.2021 12:28:32	07.06.2021 14:52:32	1	1	1		
swix	317430	test-array	07.06.2021 12:25:39	07.06.2021 12:25:39	1	2	0		

Выделенные ядра CPU:

cn-028: 0-7

Задействование CPU (среднее по ядрам), %:

28,93

Выделенные GPU (индексы):

cn-028: 0,1,2,3

Задействование GPU (среднее), %:

83,13

Использование памяти GPU (среднее), %:

8.77 GB

Сеть InfiniBand:

Использование сети (получение данных, среднее): 0.01 GBit/s

Использование сети (отправка данных, среднее): 0.01 GBit/s

Оперативная память:

Использование ОЗУ (среднее): 8.06 GB

Файловая система Lustre:

Объем переданных данных (чтение): 0.03 GB

Объем переданных данных (запись): 0.03 GB

Перезагрузить графики

Job info (4 panels)

cn-028

Utilization of CPU core: 0

Utilization of GPU: 0

Main memory usage, %

Utilization of CPU core: 1

Memory usage of GPU: 0

Infiniband bandwidth usage, %

Utilization of CPU core: 3

Utilization of GPU: 1

Utilization of CPU core: 4

Memory usage of GPU: 1

10



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

<https://hpc.hse.ru>
pkostenetskiy@hse.ru

Телефон: +7(495)7729590 (28030)

Адрес: Москва, Покровский бульвар 11, S244