РСК
Группа компаний

Энергоэффективные и высокоплотные решения для ЦОДов

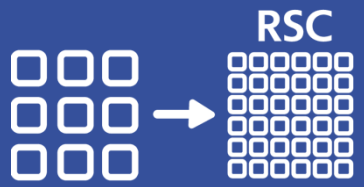
Юрий Мигаль
Руководитель департамента внедрения и эксплуатации

2023

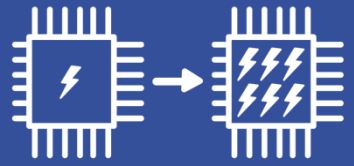
Высокопроизводительные системы с 2009 года

Разработка инновационных, энергоэффективных, высокопроизводительных и высокоплотных вычислительных систем для решения уникальных задач

Ключевые требования к HPC



Вычислительная плотность



Энергетическая плотность

Энергоэффективность

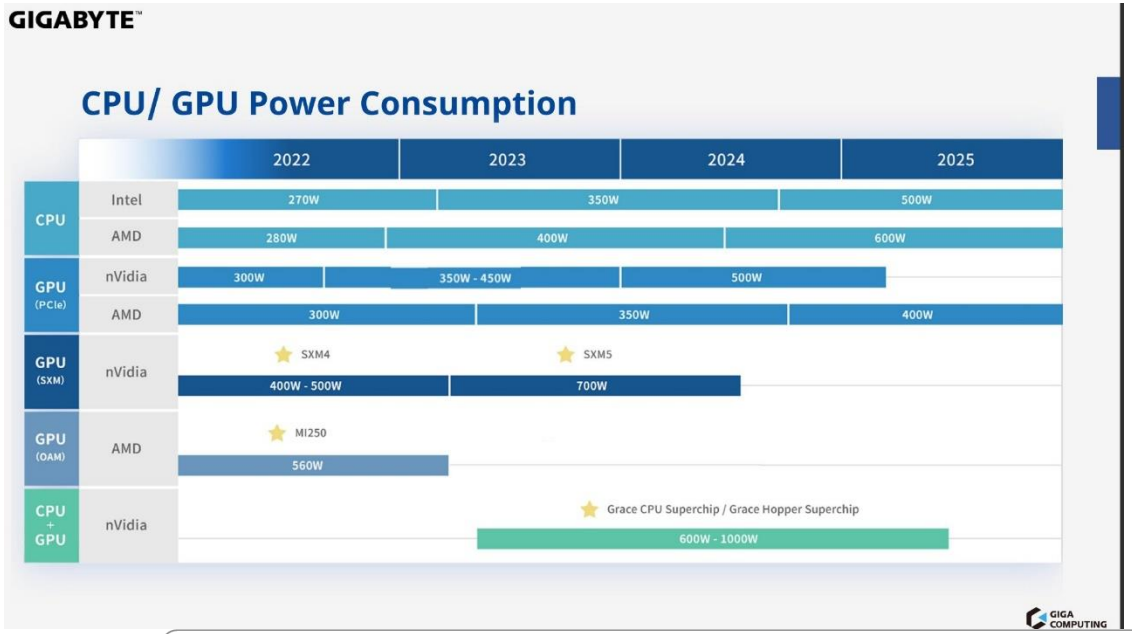
Легкость управления и обслуживания

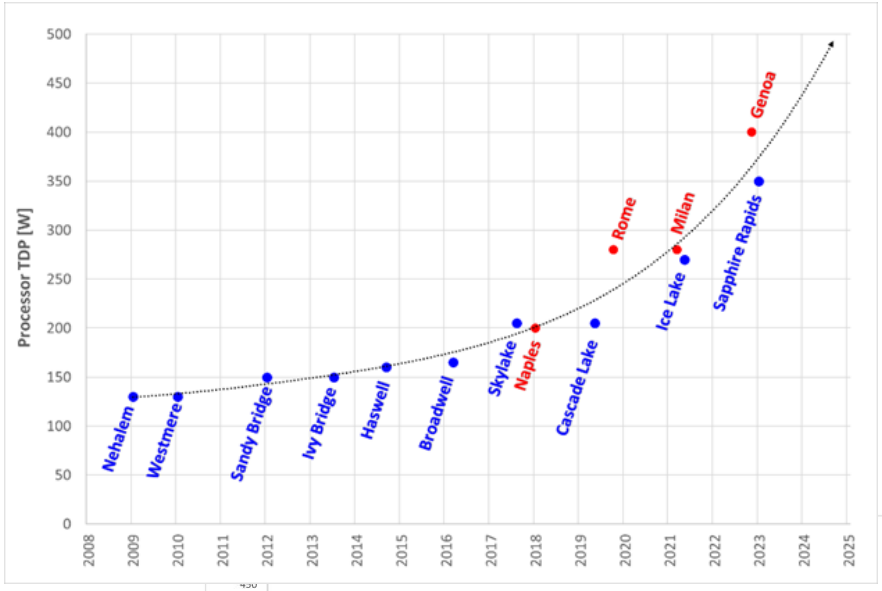
Надежность

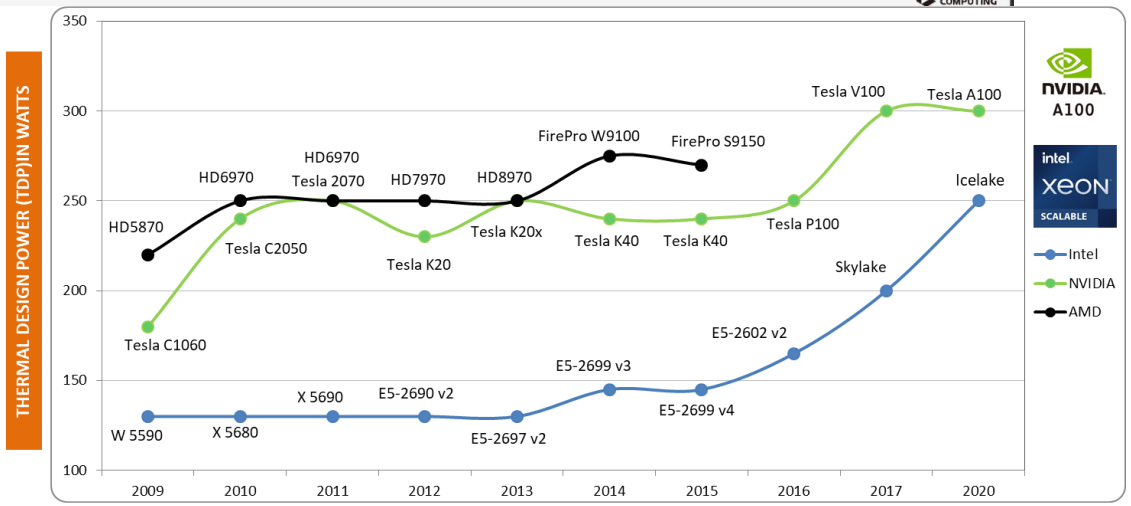
Тенденции современных ЦОД

- Увеличение мощности единицы оборудования
- Увеличение нагрузки на стойку
- «Зеленость»
- Увеличение температуры ЦОД
- Снижение PUE

Рост потребления CPU и GPU не остановить







Технологии охлаждения

Liquid Cooling Technology

~50/50 Demand Split

Immersion

Single Phase

- PA06: Zero GWP, cheaper
- FC-40 & alternatives: Higher GWP, but better cooling
- Material Compatibility

Two Phase

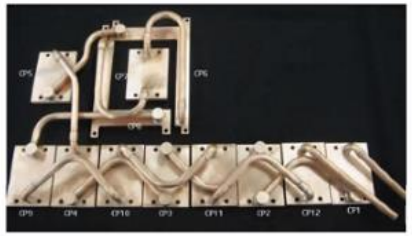
- FC-3284: Better Cooling
- High GWP
- Material Compatibility
- Heatsink design is boiling enhancement coating

Cold Plate

Single Phase

- PG25: Better Cooling
- Zero GWP
- Parallel connections to avoid pre-heat

Two Phase



- R134a, Novec7000, or other refrigerant
- Better cooling
- Enables dense systems
- Series connections ok

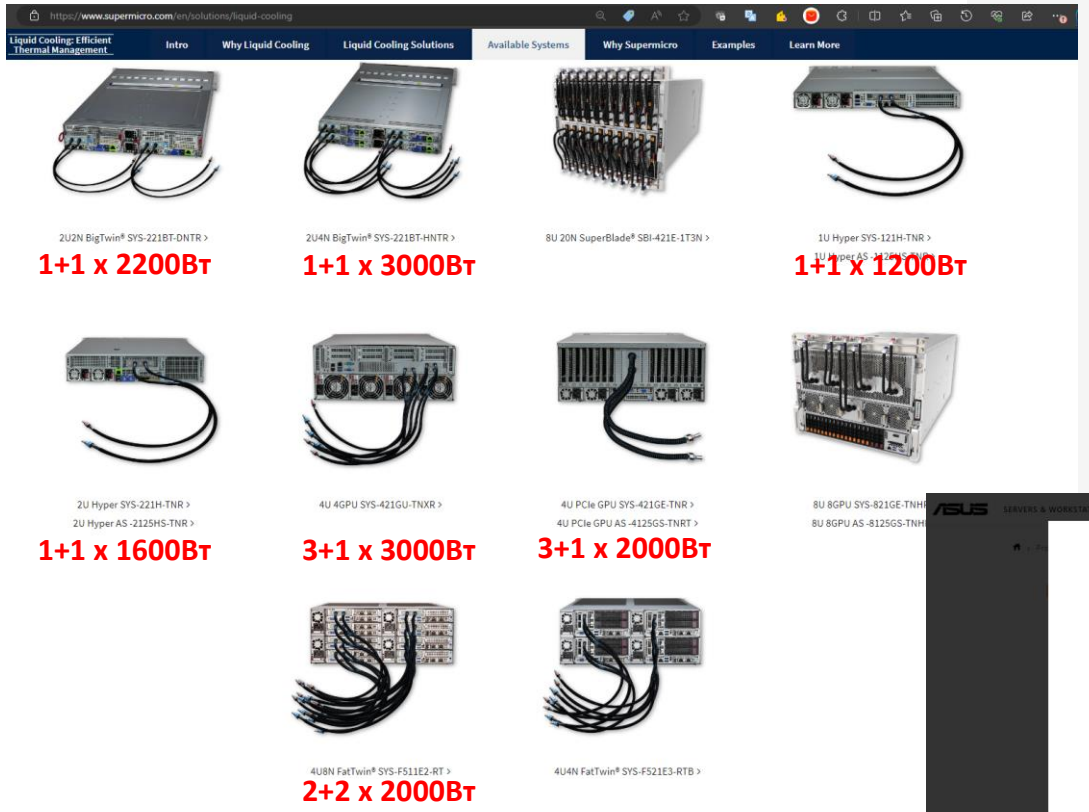
More Niche

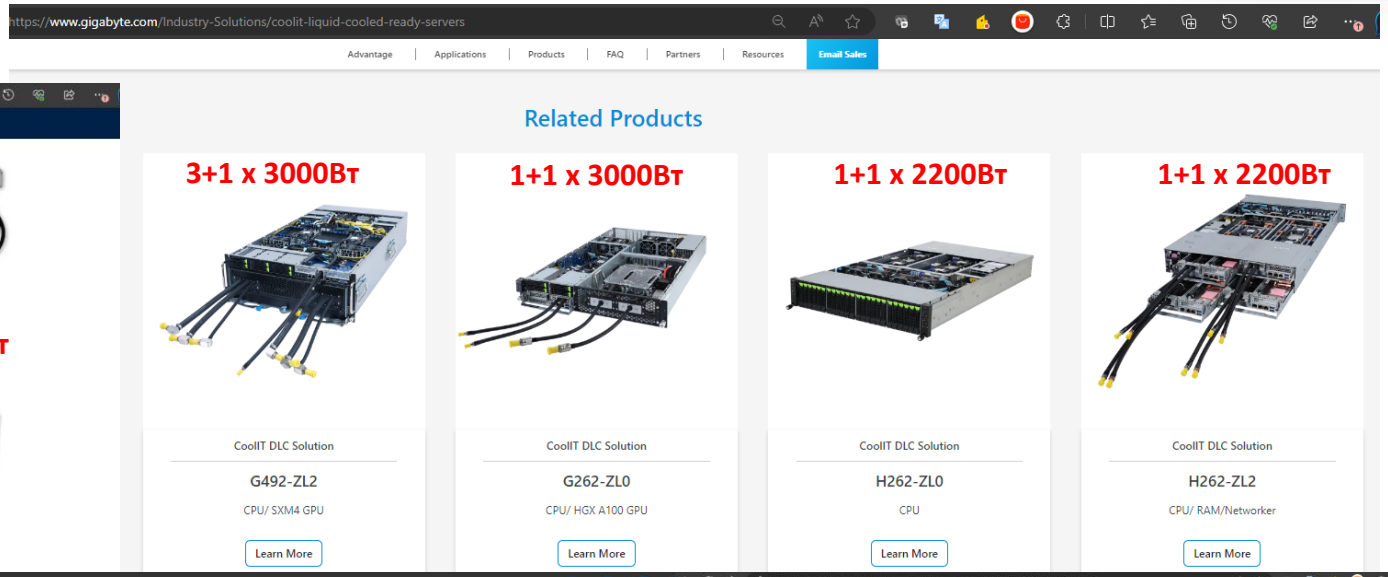
Сравнение погружного охлаждения

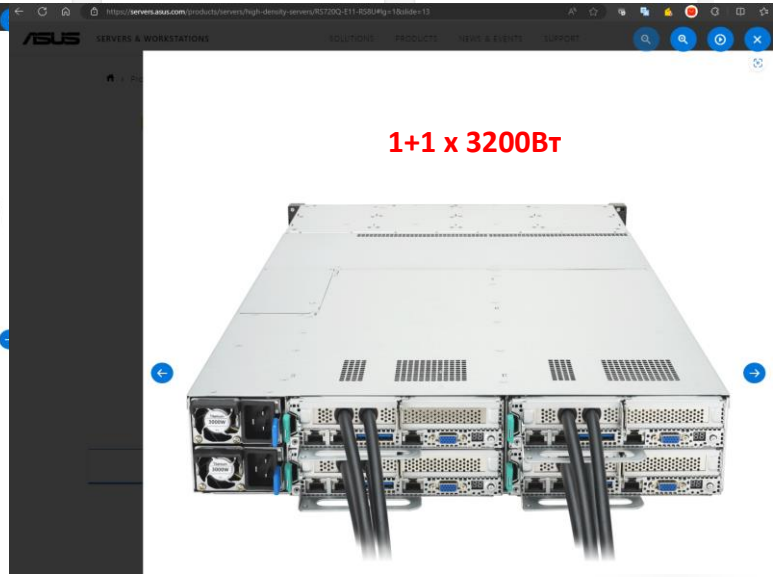
Драйверы однофазного охлаждения

	Однофазное	2-х фазное
Тип жидкости	Синтетическое углеводородное масло Фторуглерод (высокий ПГП)	Фторуглерод Фторкетон
Токсичность	Низкая токсичность	Проблема токсичности (Вытеснение кислорода)
ПГП (потенциал глобального потепления)	Низкий	Высокий
Обслуживание	Обслуживание систем в режиме реального времени	Автономное обслуживание
Стоимость	Низкая относительная стоимость	Высокая относительная стоимость
Герметичность	Отсутствие избыточного давления в ванне	Ванна под давлением
Состояние	Только жидкость	Жидкость и газ
Вязкость	Выше чем у воды	Ниже чем у воды

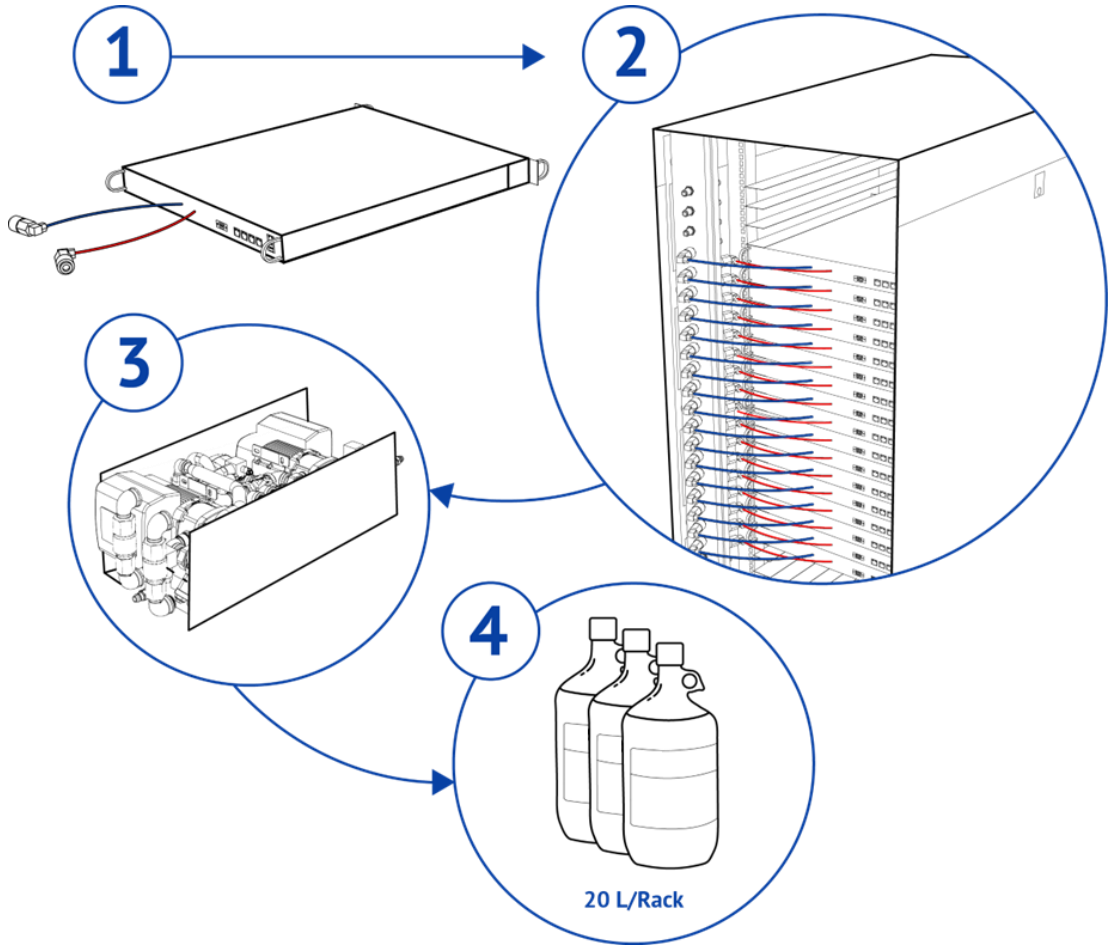
Жидкость придет и к Вам. Вопрос – КОГДА?





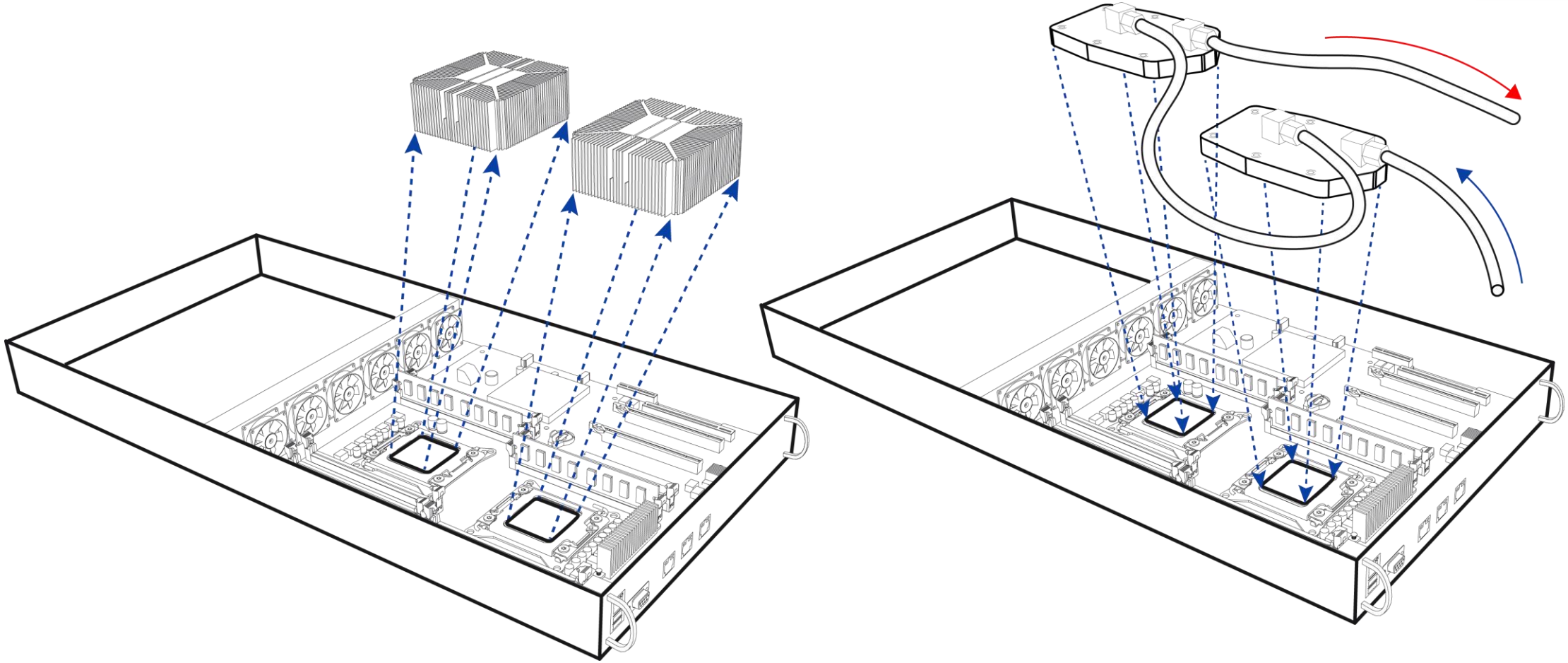


Решение РСК для ЦОД и облачных платформ – RSC HybridCoolingSolution (HCS)



- Работает с сервером **любого** производителя
- До **70%** уменьшения OPEX на охлаждение по сравнению с традиционным охлаждением воздухом
- Уменьшенное энергопотребление за счет снижения работы вентиляторов в серверах, отсутствия воздушных кондиционеров и фрикулингу (свободному охлаждению)
- Существенное сокращение использования площади ЦОД за счет более плотного размещения серверов в стойке **(до 4х раз, с 8-10кВт до 45кВт)**
- Возможность повысить температуру в серверной
- Возможность переиспользования тепловой энергии, отводимой от ЦОД

RSC HybridCoolingSolution: Установка набора для сервера

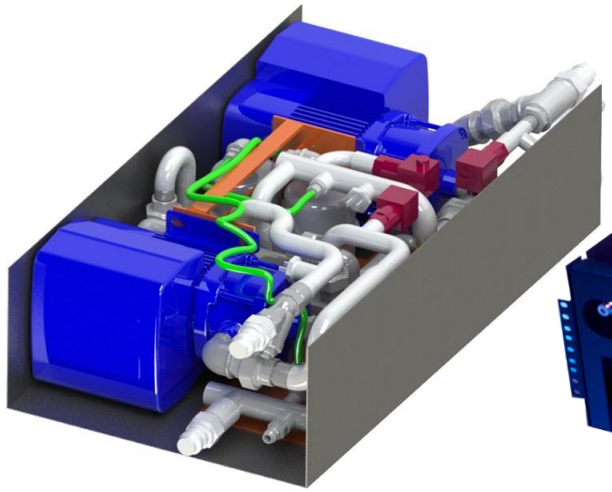


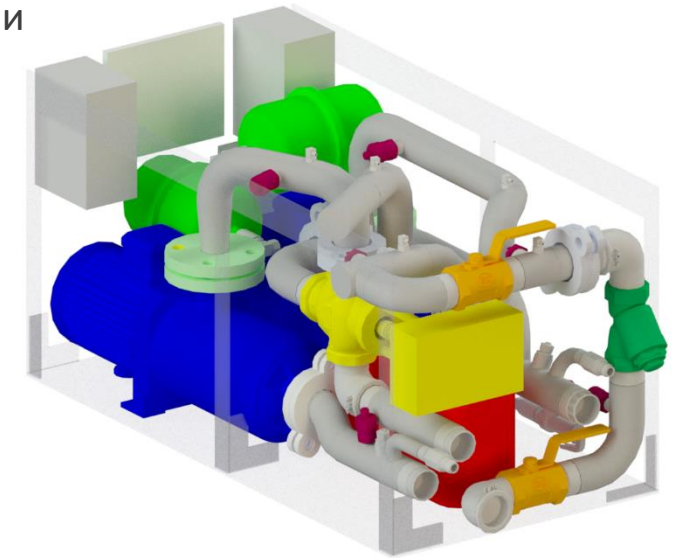
- Модули RSC HybridCoolingSolution заменяют стандартные радиаторы воздушного охлаждения и монтируются в те же отверстия на плате сервера

Узлы регулирования (насосные станции) для вычислительных систем

Более 10 лет опыта применения жидкостного охлаждения в ЦОД

Теплообменно-насосные станции обеспечивают перенос тепловой энергии и разделяют гидравлические контура с различными температурными, тепловыми, гидравлическими режимами или с различными типами жидкостей





Область применения

Работает с:

- Охлаждающие пластины РСК
- Охлаждающие пластины сторонних производителей

Отдельно стоящие узлы регулирования

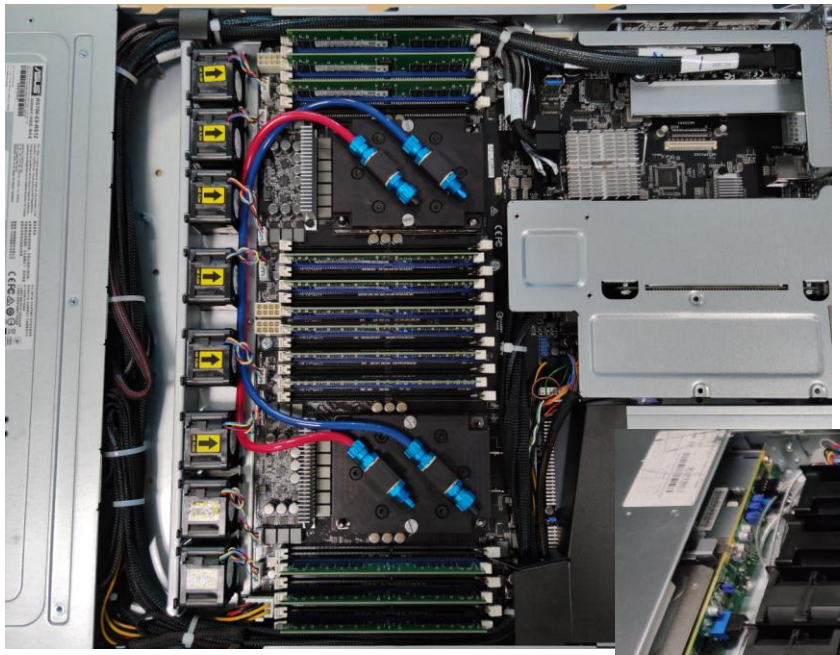
Отдельно стоящие насосные станции РСК представляют собой независимые изделия, в форм-факторе стойки.
Позволяют отвести тепловую мощность до 1000 МВт.

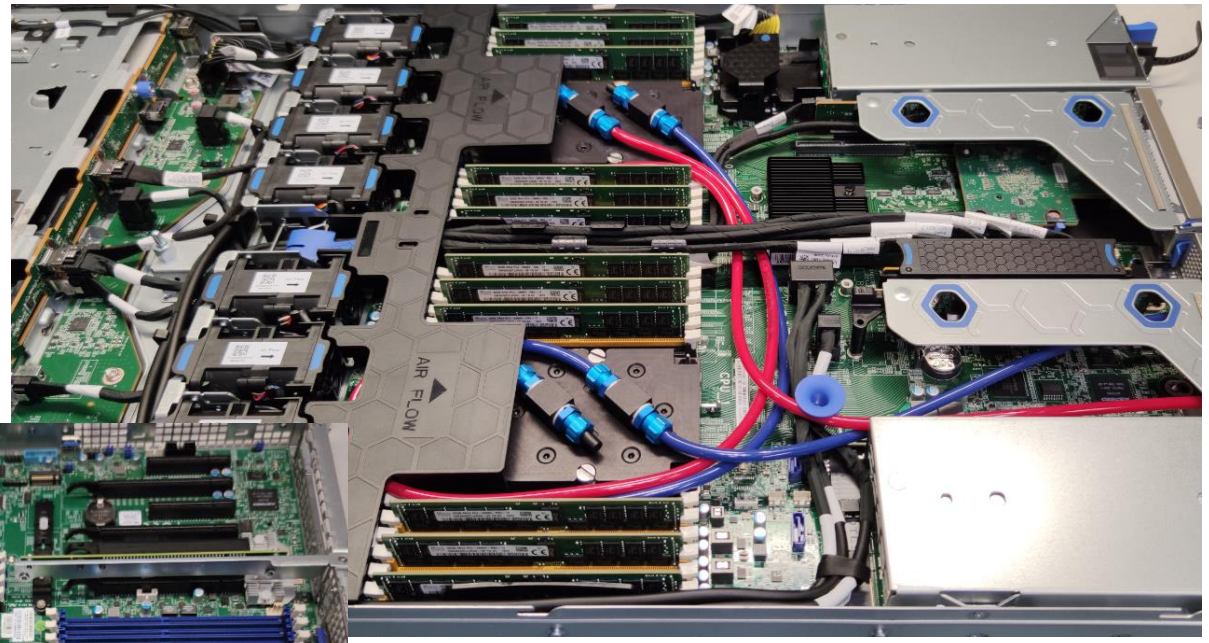
CDU 400

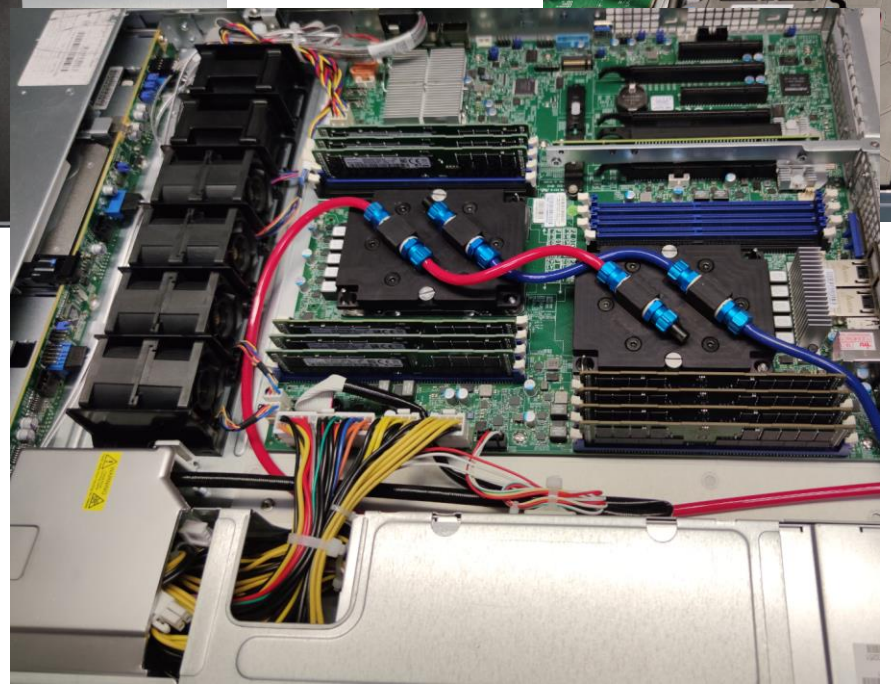
CDU-GM 400

CDU-GM 600

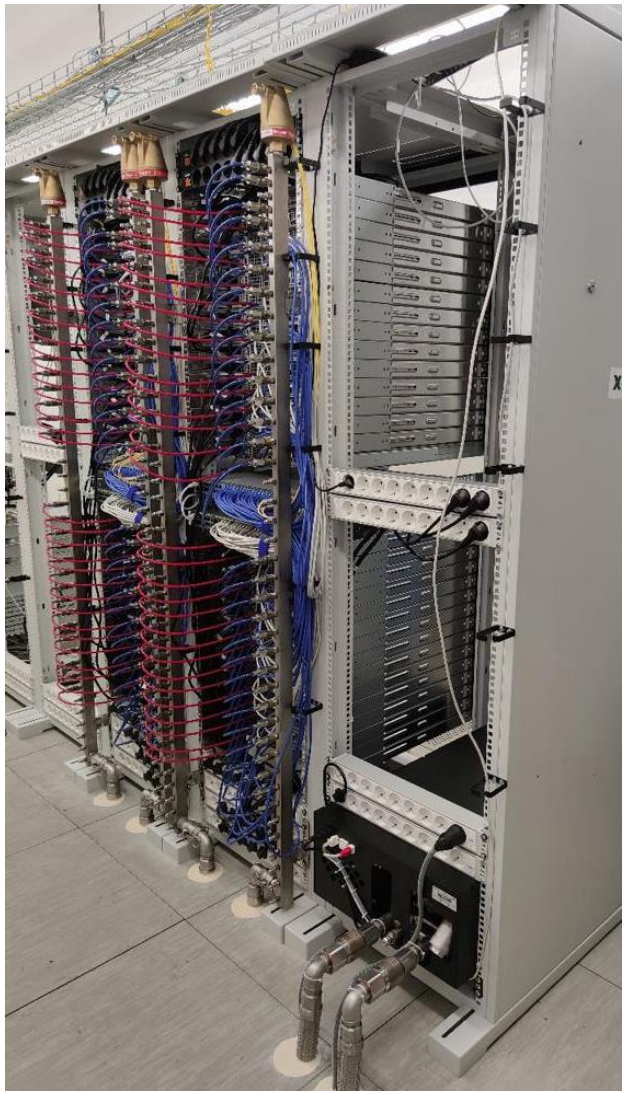
Пример инсталляции для Intel Xeon 2-го поколения



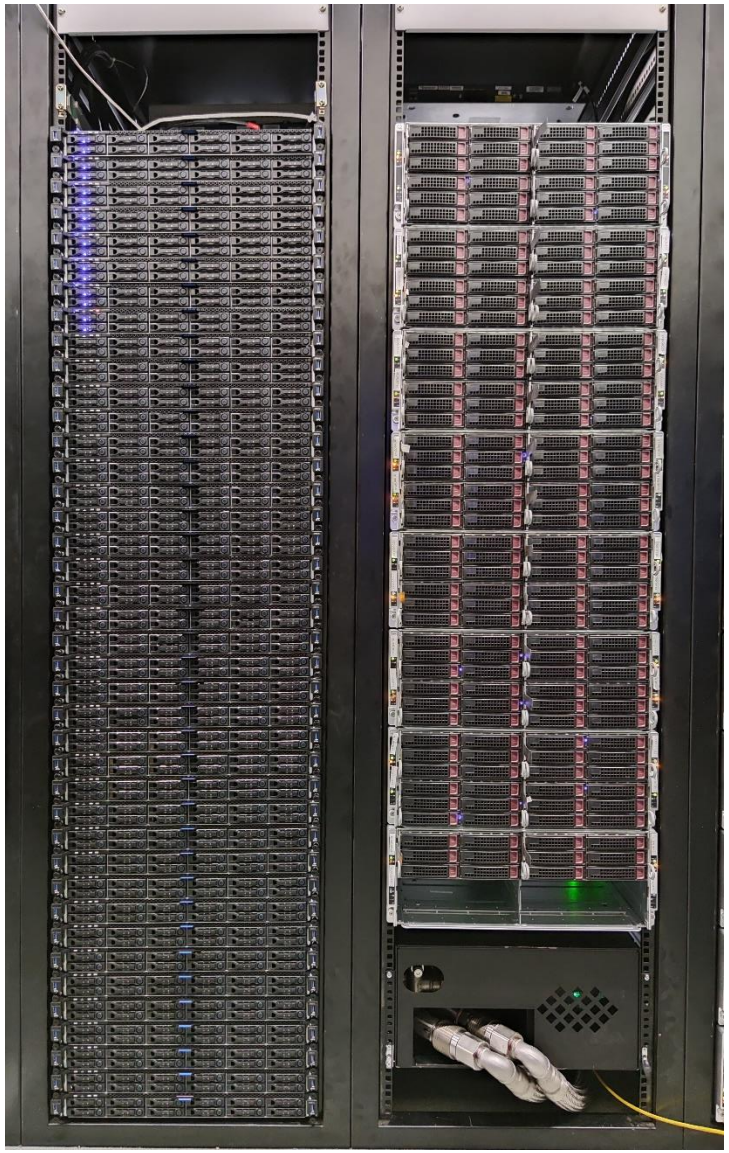


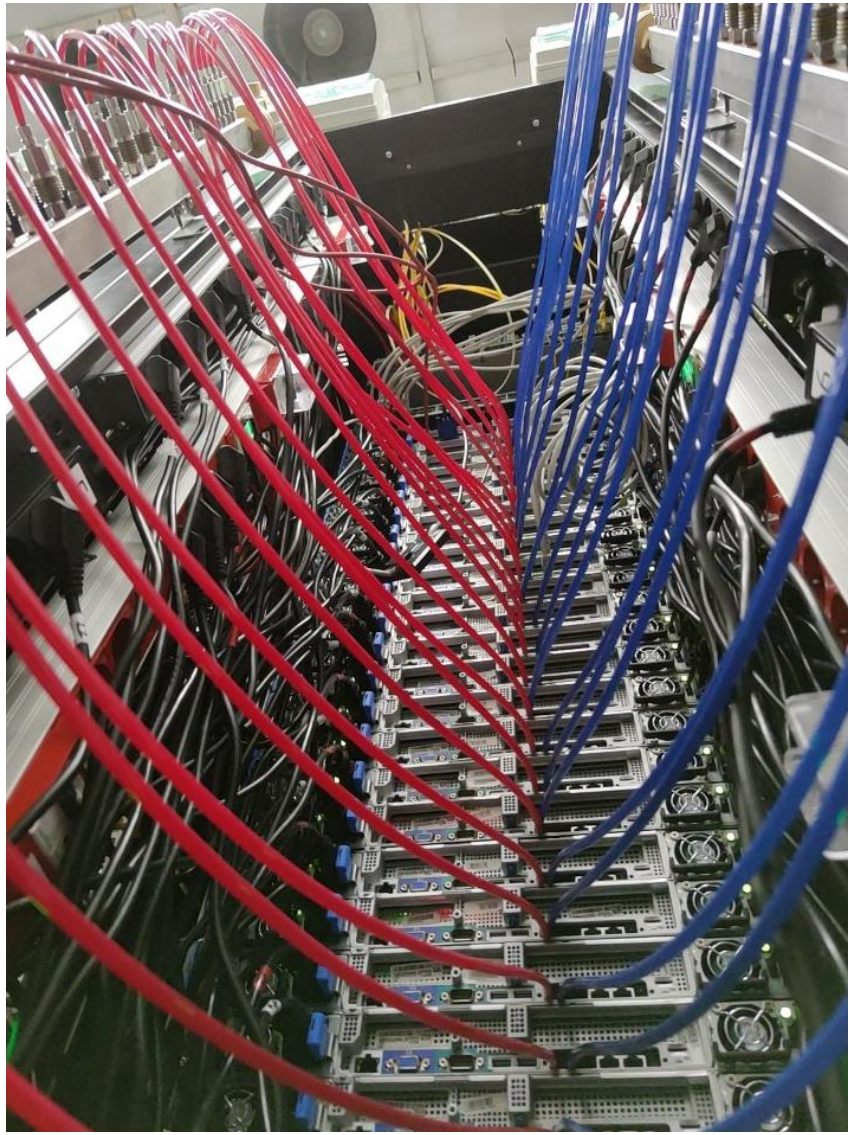


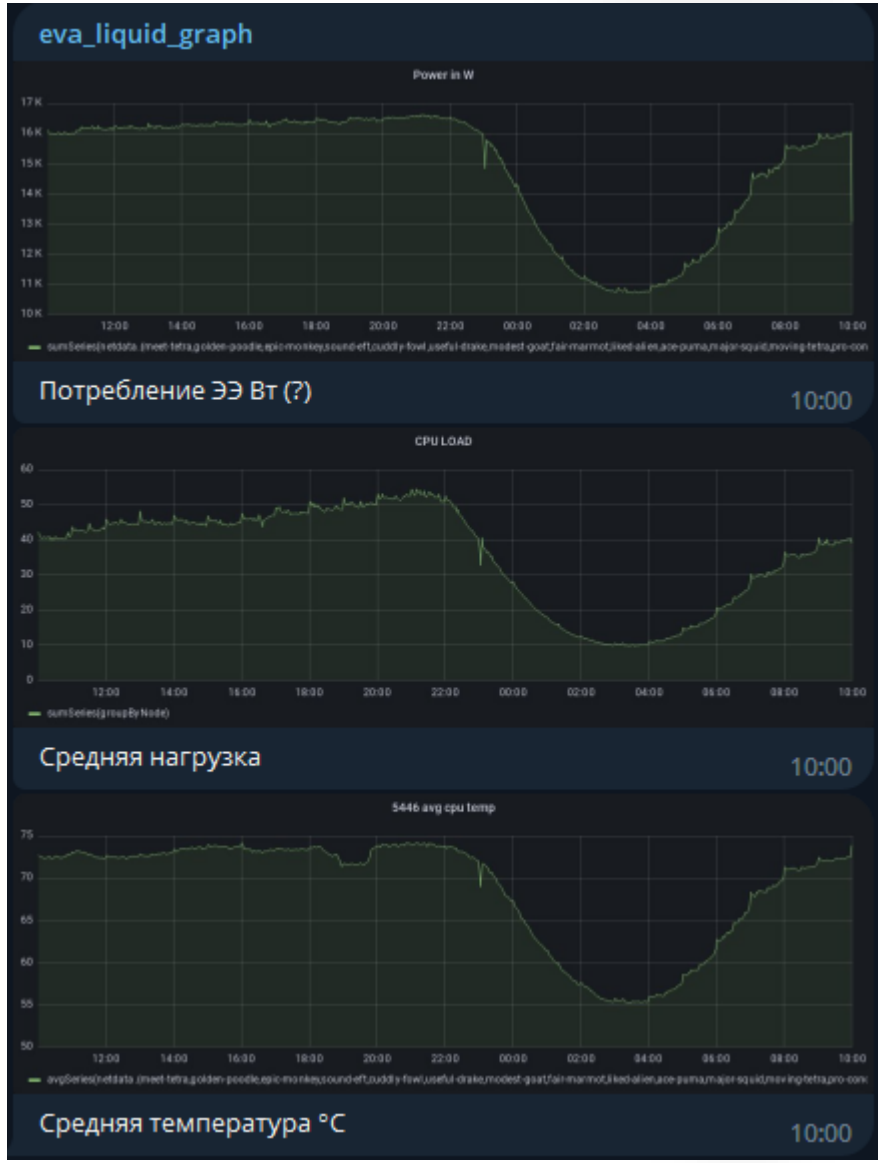
Первый этап проекта M100 VK (Москва,2020)



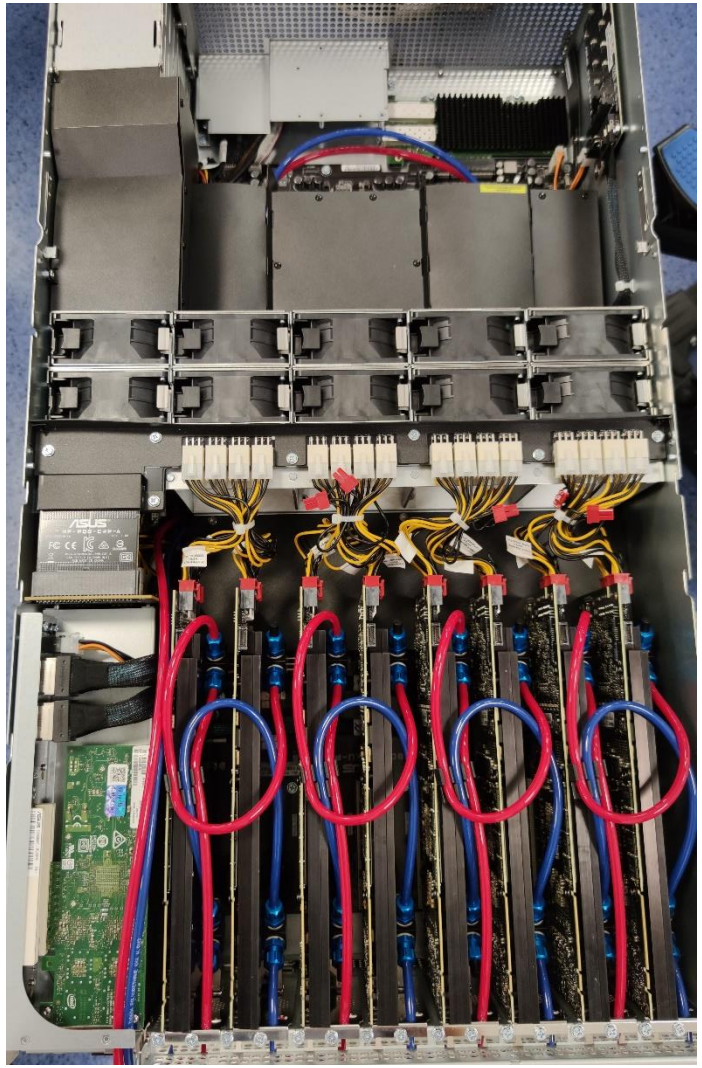
VK ИЦВА (2020)

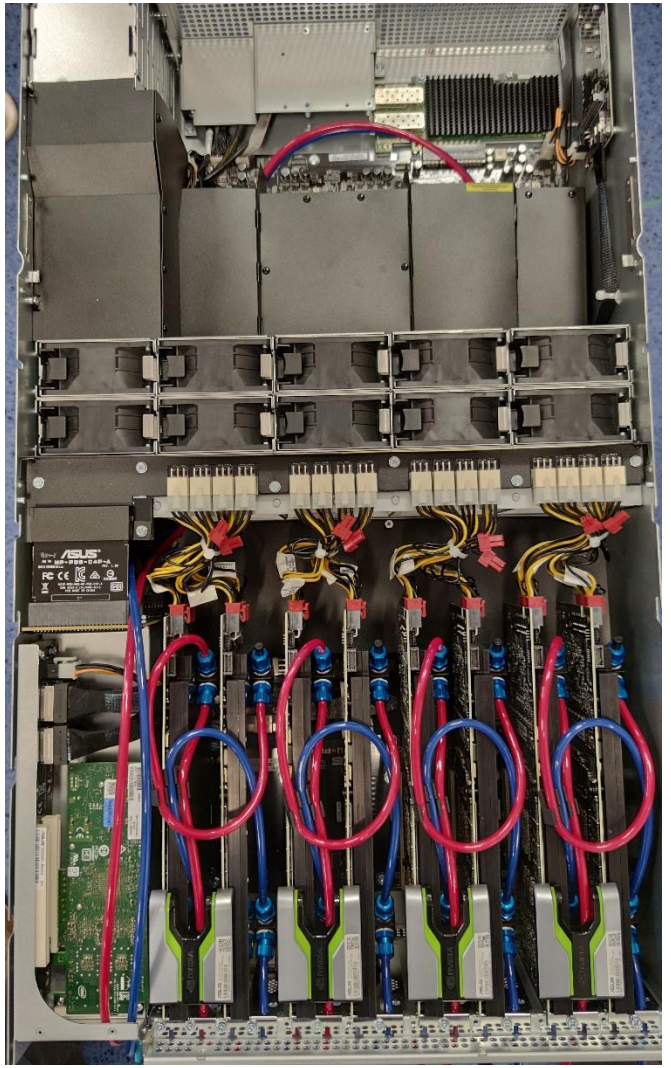




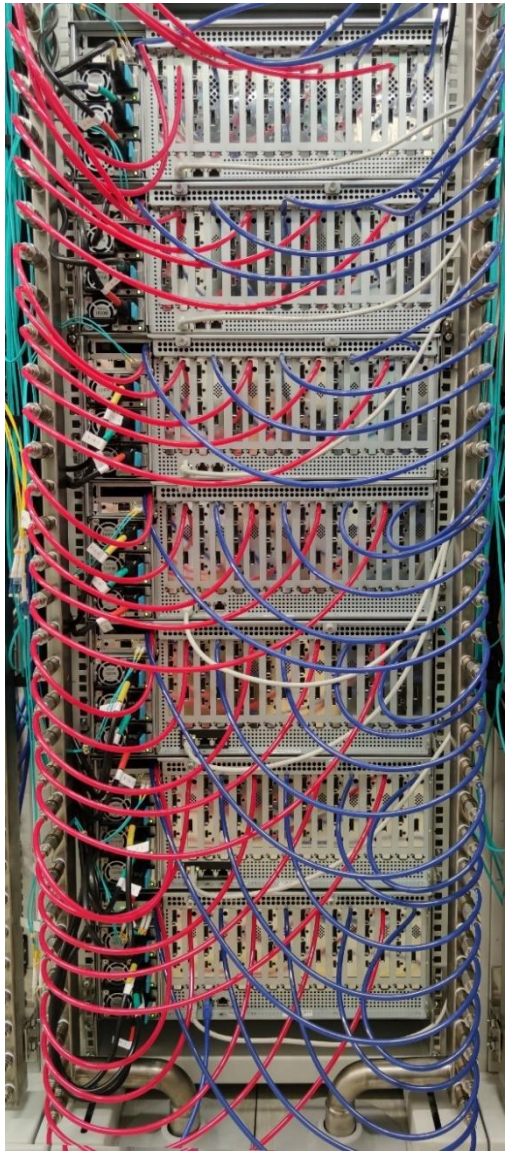

eva_liquid_graph
Power in W
Потребление ЭЭ Вт (?)
10:00
CPU LOAD
Средняя нагрузка
10:00
5446 avg cpu temp
Средняя температура °C
10:00

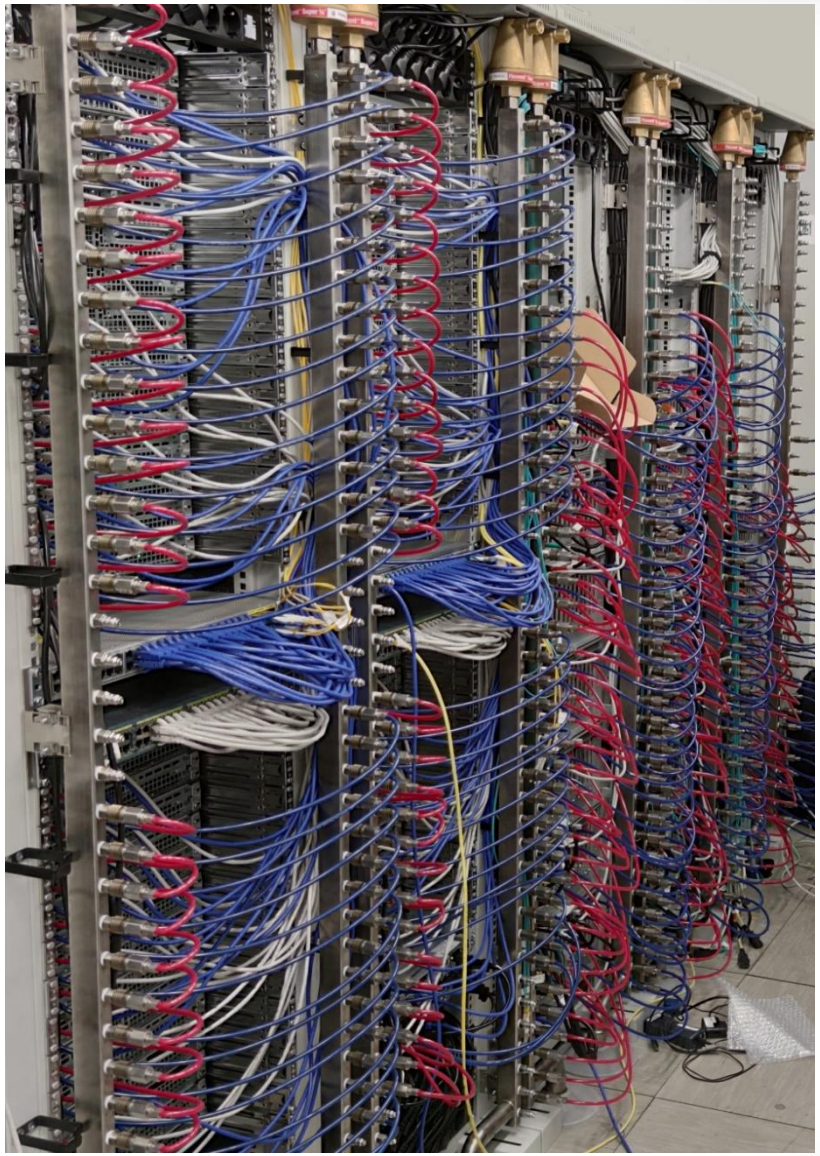
Охлаждение ускорителей RTX8000





Второй этап проекта M100 VK (Москва, 07.2021)

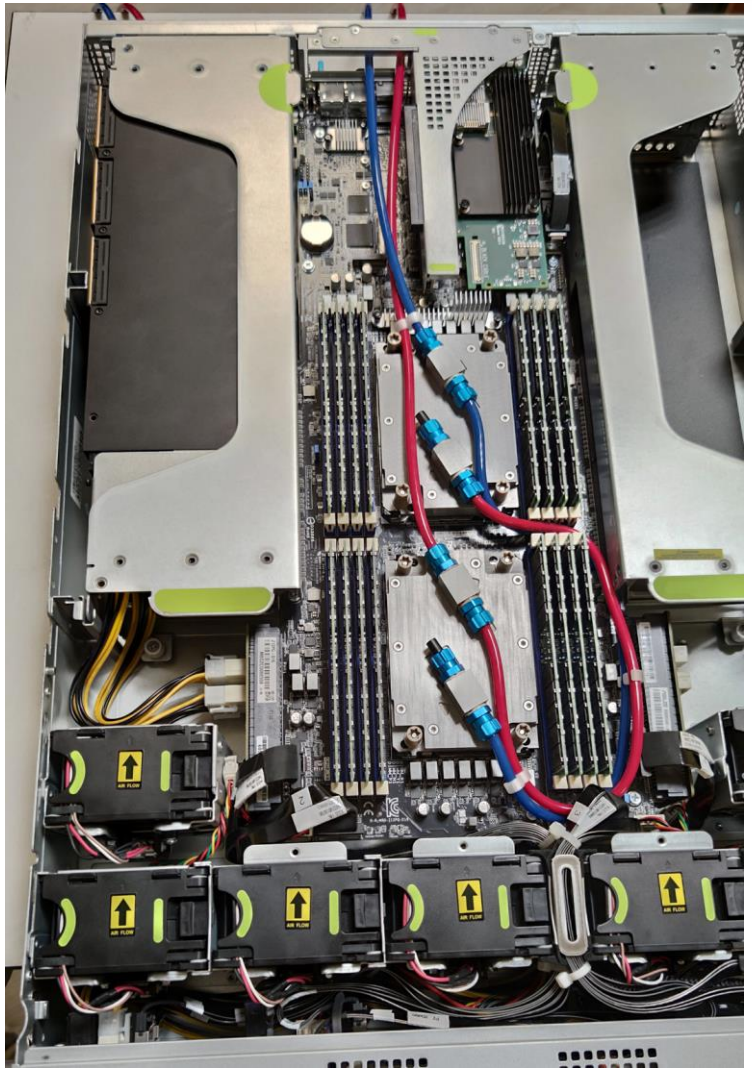




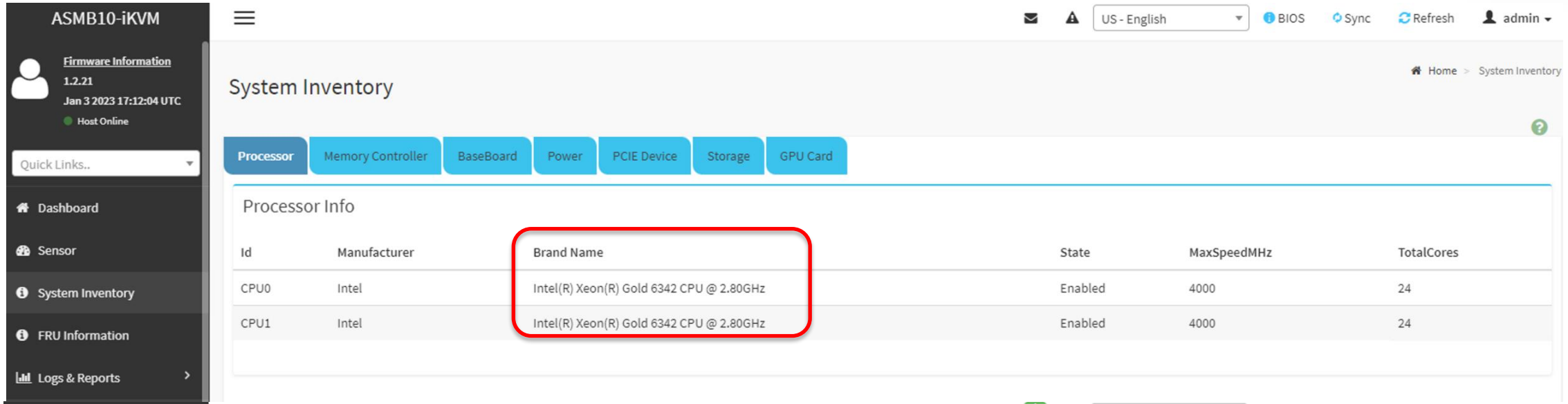
Результаты внедрения

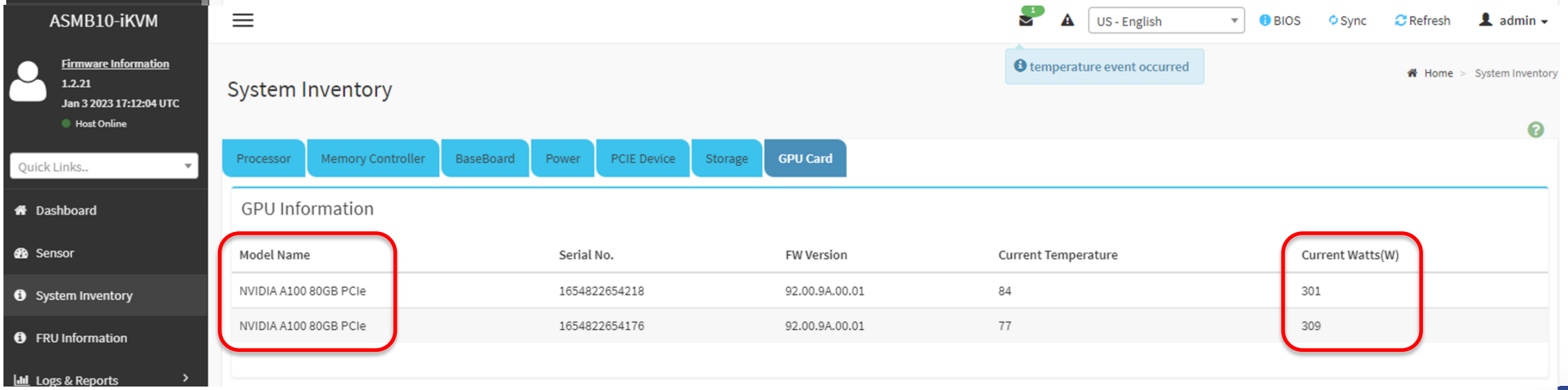
	Воздушное охлаждение	RSC ScaleStream реализация	RSC ScaleStream предельно (40U)	Примечания
Сервер 4U: 2 ЦПУ 8 ускорителей	3000 Вт	3000 Вт	3000 Вт	Без учета снижения потребления за счет вентиляторов
Охлаждается жидкостью / сервер	0 Вт	2690 Вт	2690 Вт	Ускорители 295*8=2360 ЦПУ 165*2=330
Охлаждается воздухом / сервер	3000 Вт	310 Вт	310 Вт	
Предельная мощность воздушного охлаждения на стойку	6000 Вт	6000 Вт	6000 Вт	Ограничение
Кол-во серверов в стойке	2	7	10	Стойка 42 U не более 10 серверов
Мощность потребления на стойку	6000 Вт	21000 Вт	30000 Вт	
Охлаждается жидкостью / стойку	0 Вт	18830 Вт	26900 Вт	
Охлаждается воздухом / стойку	6000 Вт	2170 Вт	3100 Вт	Воздушное охлаждение не является ограничением. 20 установленных платформ занимают 3 стойки вместо 10

Охлаждение процессоров AMD, Intel Xeon Gen4 и ускорителей A100

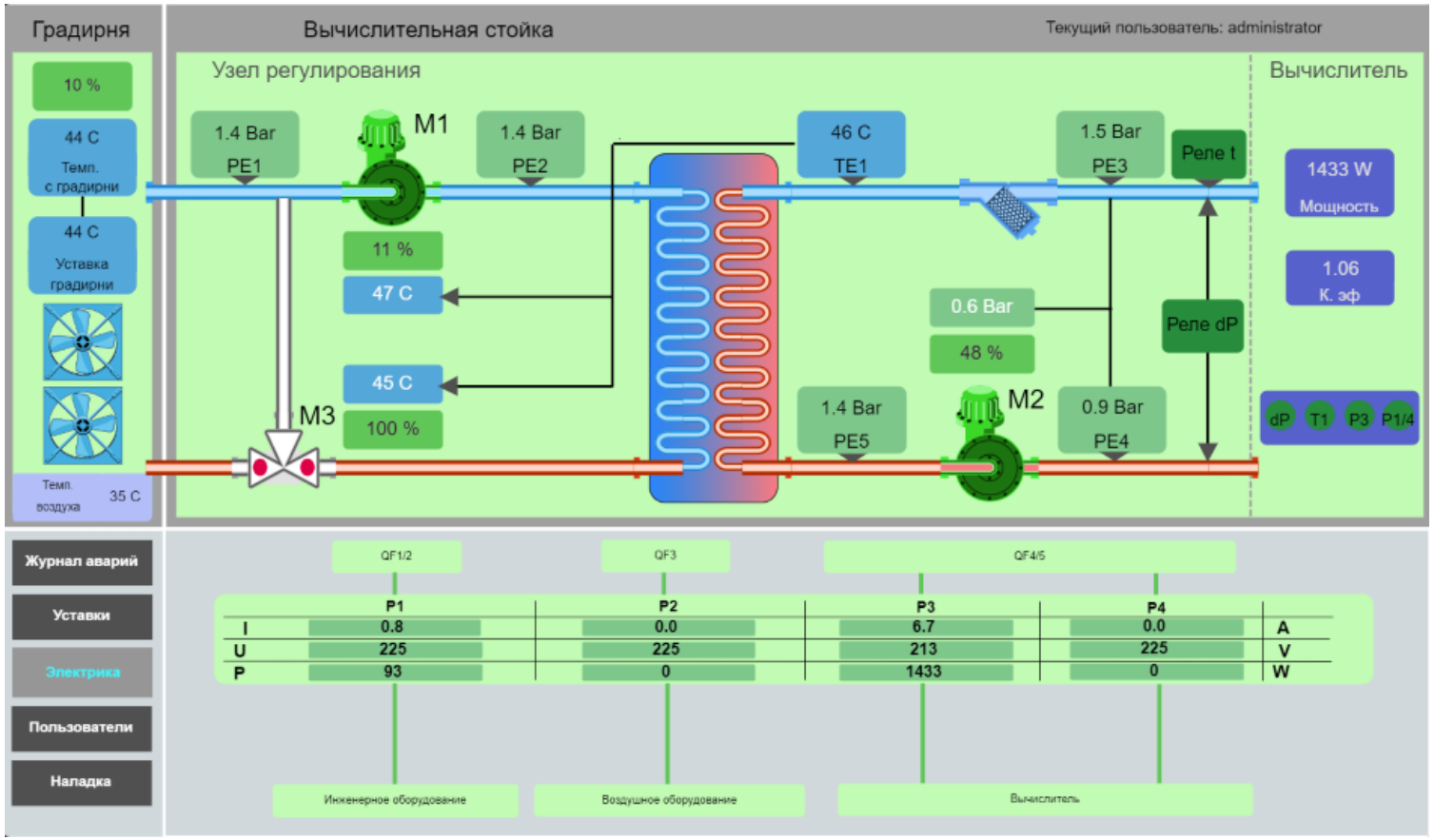


Что может дать добавление жидкости?

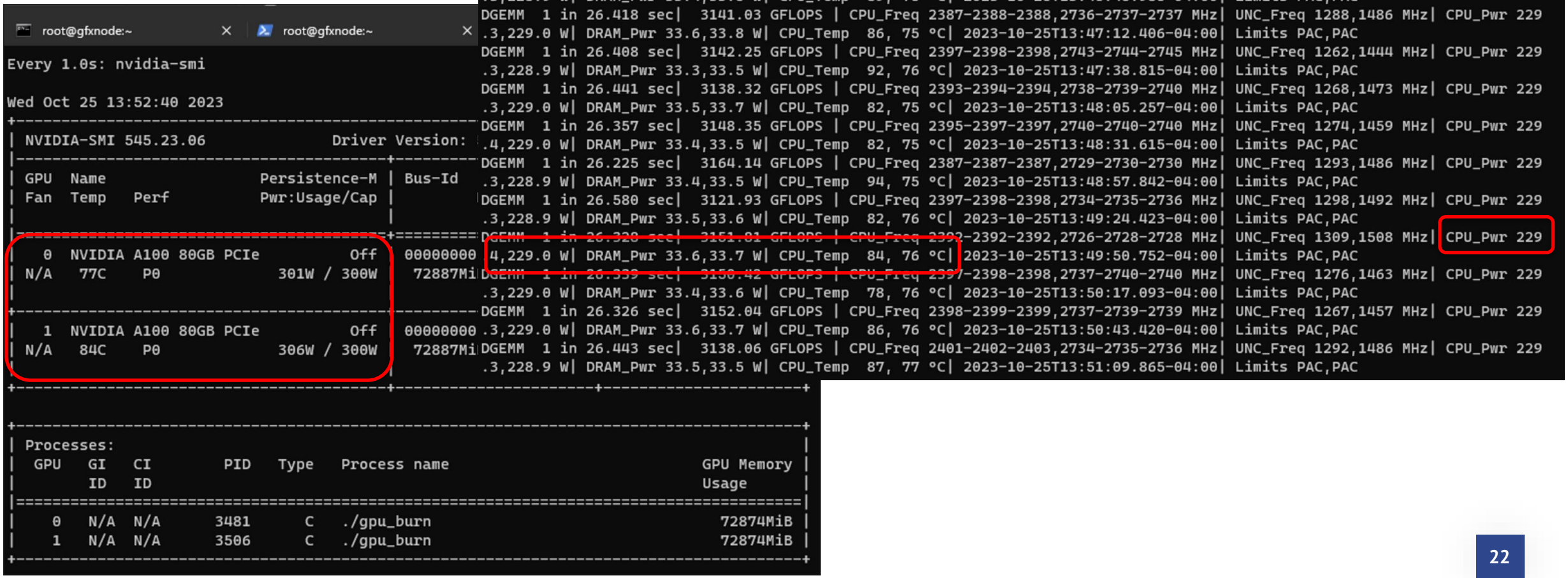




Инфраструктура эксперимента



Что будет, если нагрузить по полной?



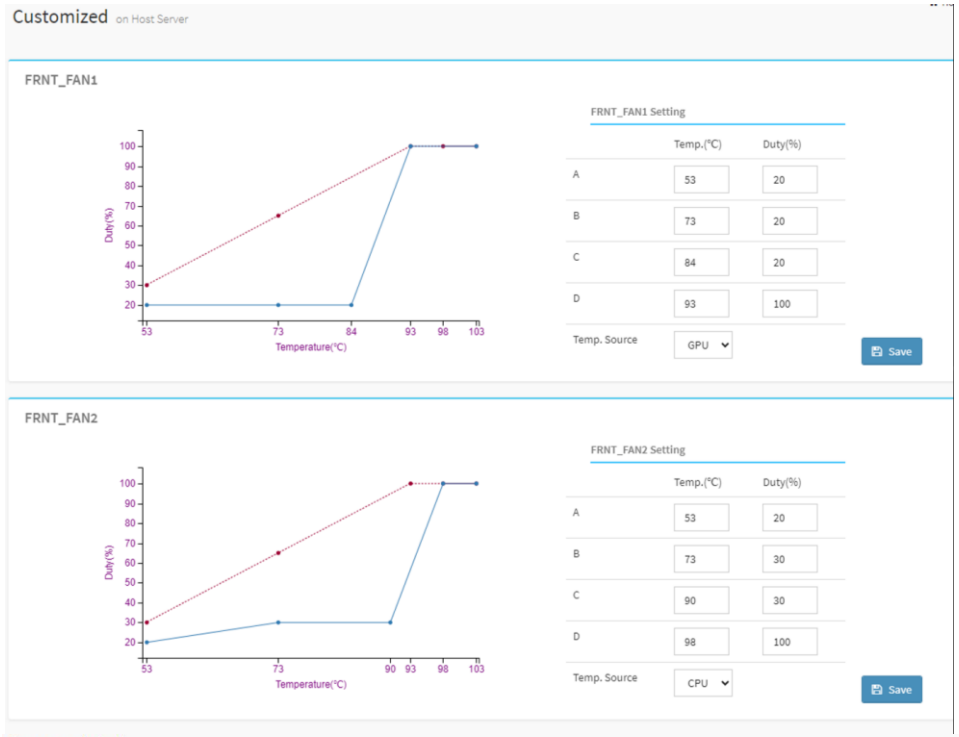
CPU1 Temperature	83 °C
CPU2 Temperature	77 °C
CPU_Power	456 Watts
DIMMA1_Temp	62 °C
DIMMB1_Temp	59 °C
DIMMC1_Temp	58 °C
DIMMD1_Temp	62 °C
DIMME1_Temp	64 °C
DIMMF1_Temp	61 °C
DIMMG1_Temp	60 °C
DIMMH1_Temp	63 °C
DIMMJ1_Temp	56 °C
DIMMK1_Temp	54 °C
DIMML1_Temp	56 °C
DIMMM1_Temp	58 °C
DIMMN1_Temp	53 °C
DIMMP1_Temp	51 °C
DIMMR1_Temp	51 °C
DIMMT1_Temp	53 °C

Потребление сервера по 220В	1408 Вт
Потребление сервера по 12В	1328 Вт
Потребление блоков питания	80 Вт
Потребление памяти	64 Вт
Потребление процессоров	456 Вт
Потребление GPU	600 Вт
Остальная система и вентиляторы	208 Вт
ИТОГО воздух	352 Вт
ИТОГО жидкость	1056 Вт

PCK

FRNT_FAN1	3360 RPM
FRNT_FAN2	4920 RPM
FRNT_FAN3	5040 RPM
FRNT_FAN4	4920 RPM
FRNT_FAN5	3360 RPM
FRNT_FAN6	3360 RPM
FRNT_FAN7	3360 RPM
Memory_Power	64 Watts
PSU1 Power In	688 Watts
PSU1 Power Out	656 Watts
PSU2 Power In	720 Watts
PSU2 Power Out	672 Watts
TR1 Temperature	33 °C
VBAT	3.12 Volts

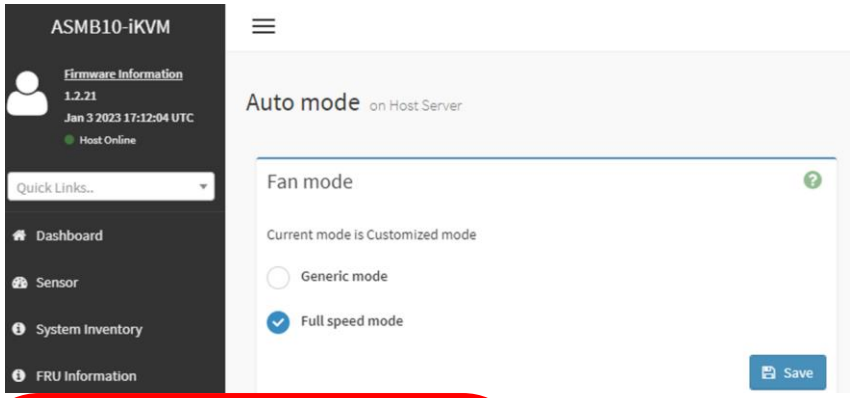
Профиль оборудования важен

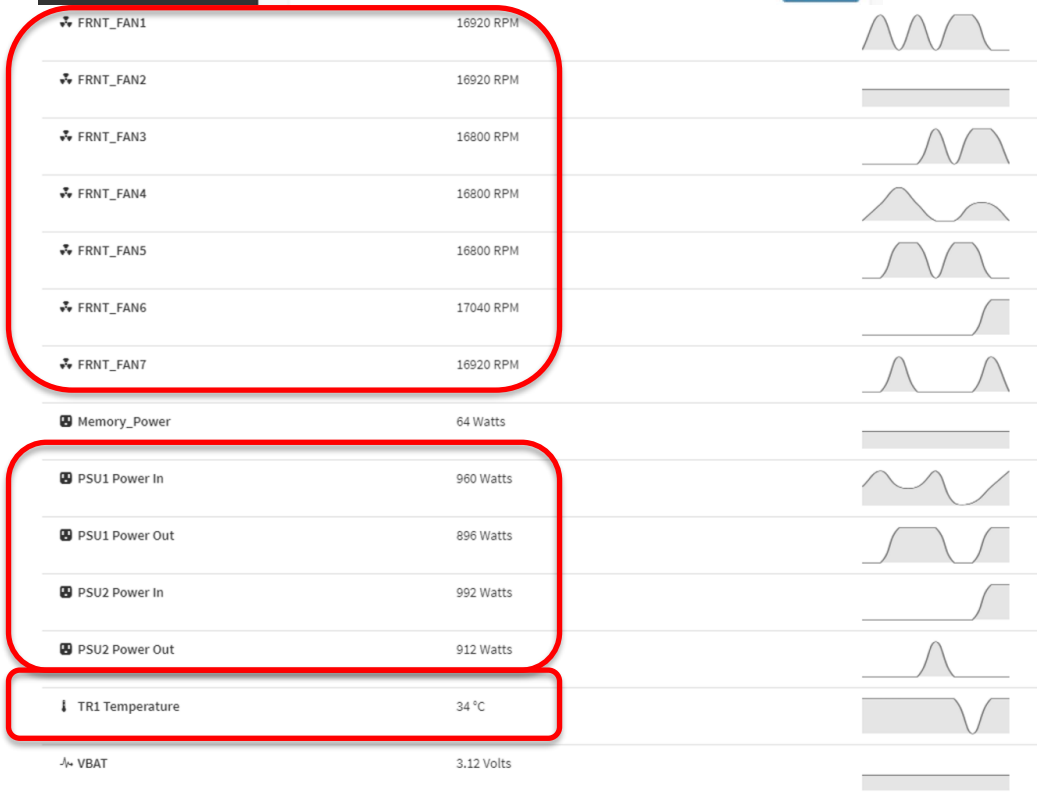


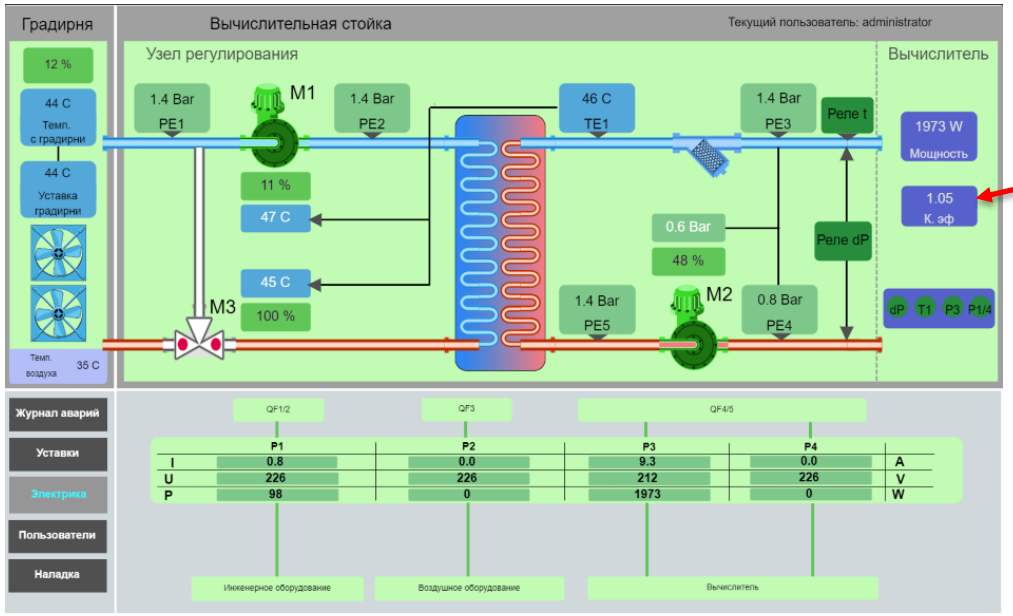
Необходимо понимание каким образом охлаждаются компоненты:

- **На вентиляторах GPU остаются лишь цепи питания**
- **На вентиляторах CPU остаются VRM и внешние сетевые адаптеры**
- **Память имеет низкую энергетическую плотность и низкие требования к охлаждению**
- **Диски SSD не имеют проблем с охлаждением**

А если только воздух?







PUE стал лучше?
Причины?
Вентиляторы в сервере входят в ИТ нагрузку!
Но это же система охлаждения!!!

(1973+98)/1973=1,05
(1973+98)/1473=1,41

Расчетная нагрузка	*Вт/сервер*	*6 кВт/стойку*	*10 кВт/стойку*
Кол-во серверов потребляющих по 220В	1 952	3	5
ИТОГО воздух	1 952	5 856	9 760
Затраты на воздушное охлаждение (0.3)	586	1 757	2 928
Общее потребление		**7 613**	**12 688**

Жидкость поможет!

Расчетная нагрузка	*Вт/сервер*	*6 кВт/стойку*	*10 кВт/стойку*
Кол-во серверов потребляющих по 220В	1 408	4	7
ИТОГО воздух	352	1 408	2 464
ИТОГО жидкость	1 056	4 224	7 392
Затраты на воздушное охлаждение (0.3)	106	422	739
Затраты на жидкостное охлаждение (0.1)	106	422	739
Общее потребление		**6 477**	**11 334**

Мировые рекорды

Компонуемая архитектура «РСК Торнадо»
на базе Intel Xeon Scalable 3-го поколения

967,45 ТФлопс**
Вычислительная плотность
на шкаф*

604,66 ТФлопс/м³
Производительность на объем

3,67 ТБ/с
Пропускная способность
распределенной системы
хранения

130 кВт
Энергетическая
плотность на шкаф

* шкаф 42U 80x100 см
** для решений на базе Intel® Xeon®

Дезагрегированная компонуемая инфраструктура

Вычислительные узлы

с поддержкой процессоров Intel, AMD и «Эльбрус»

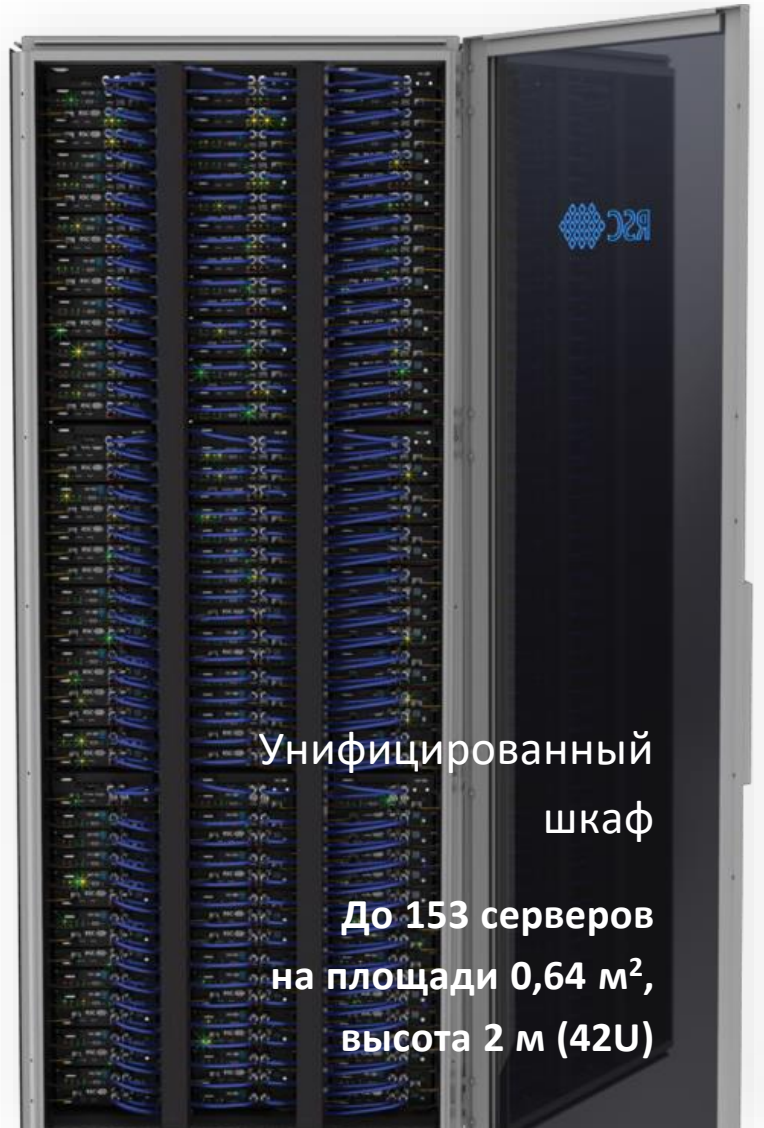
Гиперконвергентные узлы

с устройствами хранения NVMe

Модули избыточного питания

Програмный стек управления

RSC BasIS Software Platform



Унифицированный шкаф

До 153 серверов на площади 0,64 м², высота 2 м (42U)

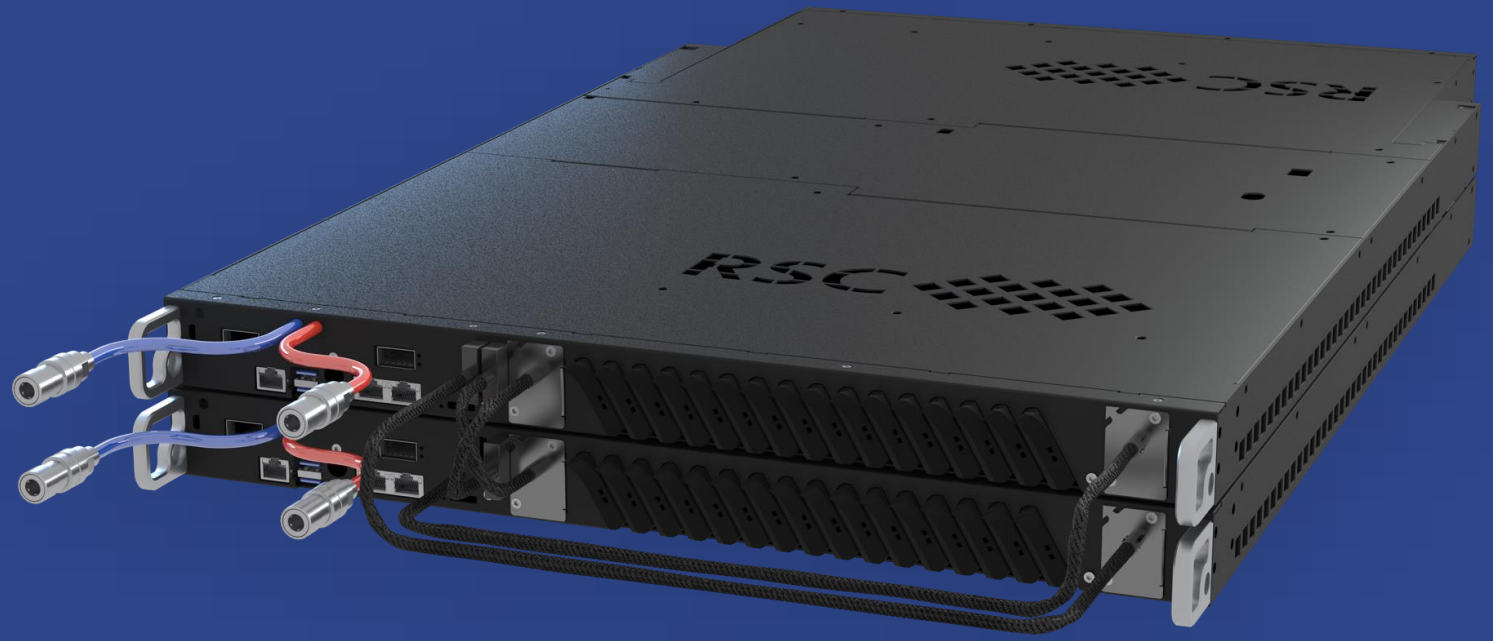
Вычислительный узел «РСК Торнадо» на базе Intel Xeon Scalable 3-го поколения

- **100% охлаждение «горячей водой»** позволяет достичь рекордной энергоэффективности (PUE < 1,04)
- **2х процессора Intel® Xeon® Scalable 3-го поколения:** Platinum 83xx (TDP 270W) или Gold 63xx (TDP 205W)

- **8** модулей энергонезависимой памяти Intel® Optane DC Persistent Memory (до 4 ТБ)
- **5 NVMe SSD** диска E1.S/M.2 (2 Hot Swap) до 16 ТБ хранения
- Intel® Omni-Path 100 Gb/s, Infiniband 100/200Gb/s, High speed Ethernet
- 2 порта 10Gb/s Ethernet

СХД РСК Tornado AFS рекордно большого объема

E1.L Intel® Data Center SSDs
в форм-факторе EDSFF



Сверхвысокая емкость – 1 ПБ
в одном сервере (1U)
Надежное объединение 2 серверов
в СХД емкостью **2 ПБ (2U)**

Первое решение на 100% жидкостном охлаждении с высочайшей плотностью на базе 32x Intel EDSFF SSDs, двух процессоров Intel® Xeon® Scalable и памяти Intel® Optane™ DC Persistent Memory

100% охлаждение «горячей водой» позволяет достичь рекордной энергоэффективности (CUE < 1,04)

Мировые рекорды

РСК Торнадо AFS с новым Intel SSD P5316

41.3 ПБ

Рекордная емкость теплой системы хранения на шкаф

1 ТБ/с

Пропускная способность распределенной системы хранения

50 кВ

Энергетическая плотность на шкаф

шкаф 42U 60x120 см

Система хранения «по запросу»

«РСК БазИС» позволяет создавать системы хранения данных «по запросу»:

Кластерная файловая система Lustre

Стандарт «де-факто» в мире суперкомпьютеров

Новая высокопроизводительная объектная система хранения DAOS

Разработана «с чистого листа» для поддержки высокоскоростных фабрик, устройств NVMe и Storage Class Memory

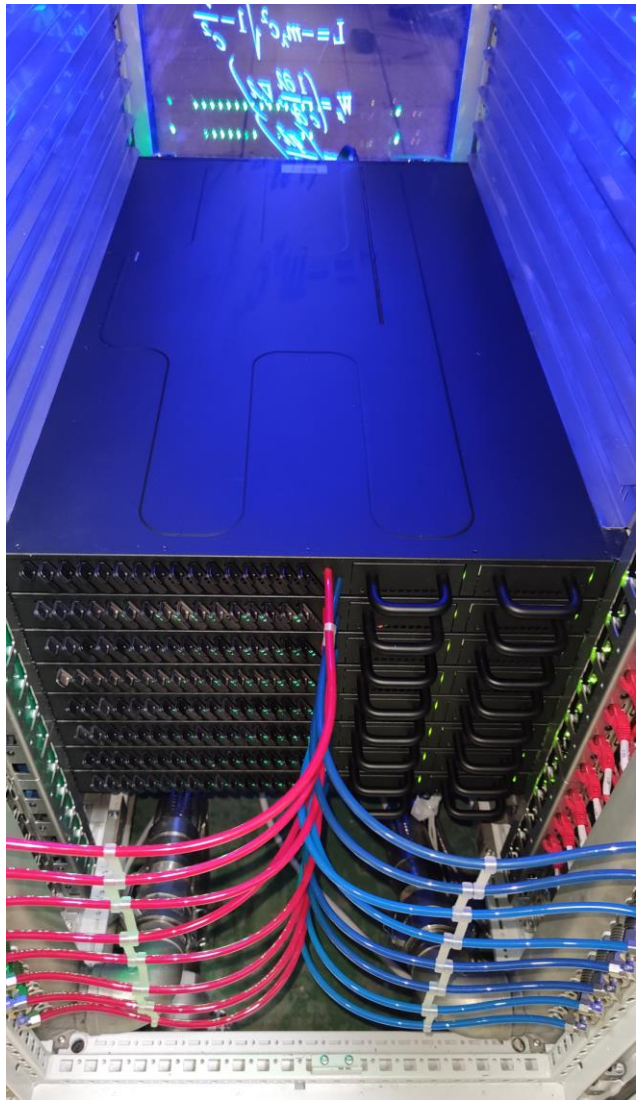
Предоставляет современные высокопроизводительные методы работы с данными:

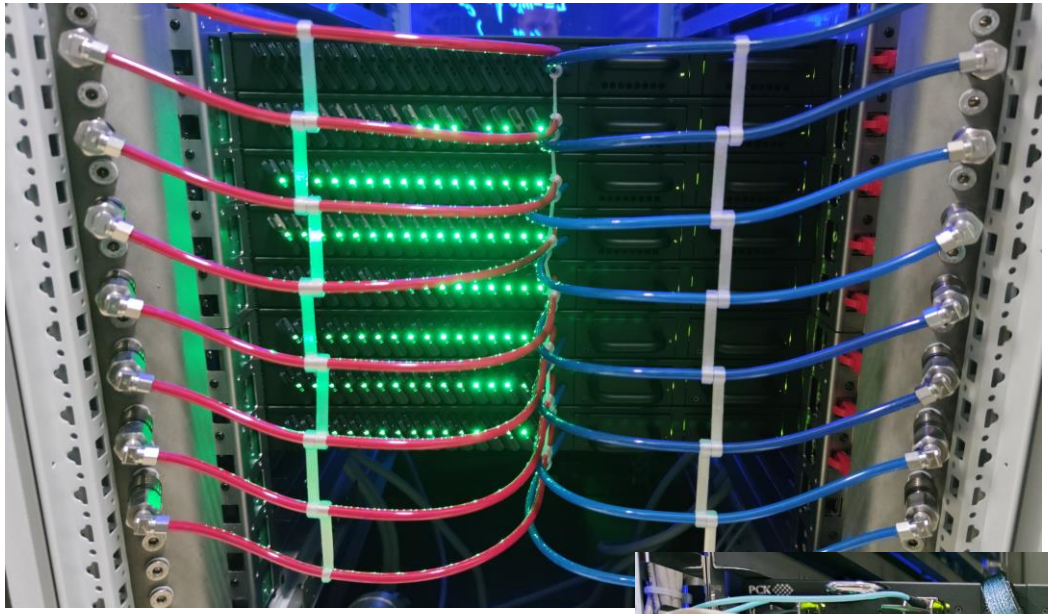
HDF5	Apache Spark	MPI-IO	TensorFlow	NoSQL	S3	POSIX

Система хранения «по запросу» РСК БазИС

Группа компаний РСК получила престижную награду Russian DC Awards 2020 в номинации «Лучшее ИТ-решение для ЦОДа», победив с проектом «Высокопроизводительная система хранения для суперкомпьютера», реализованном в 2020 году в Объединенном институте ядерных исследований (ОИЯИ) в Дубне.

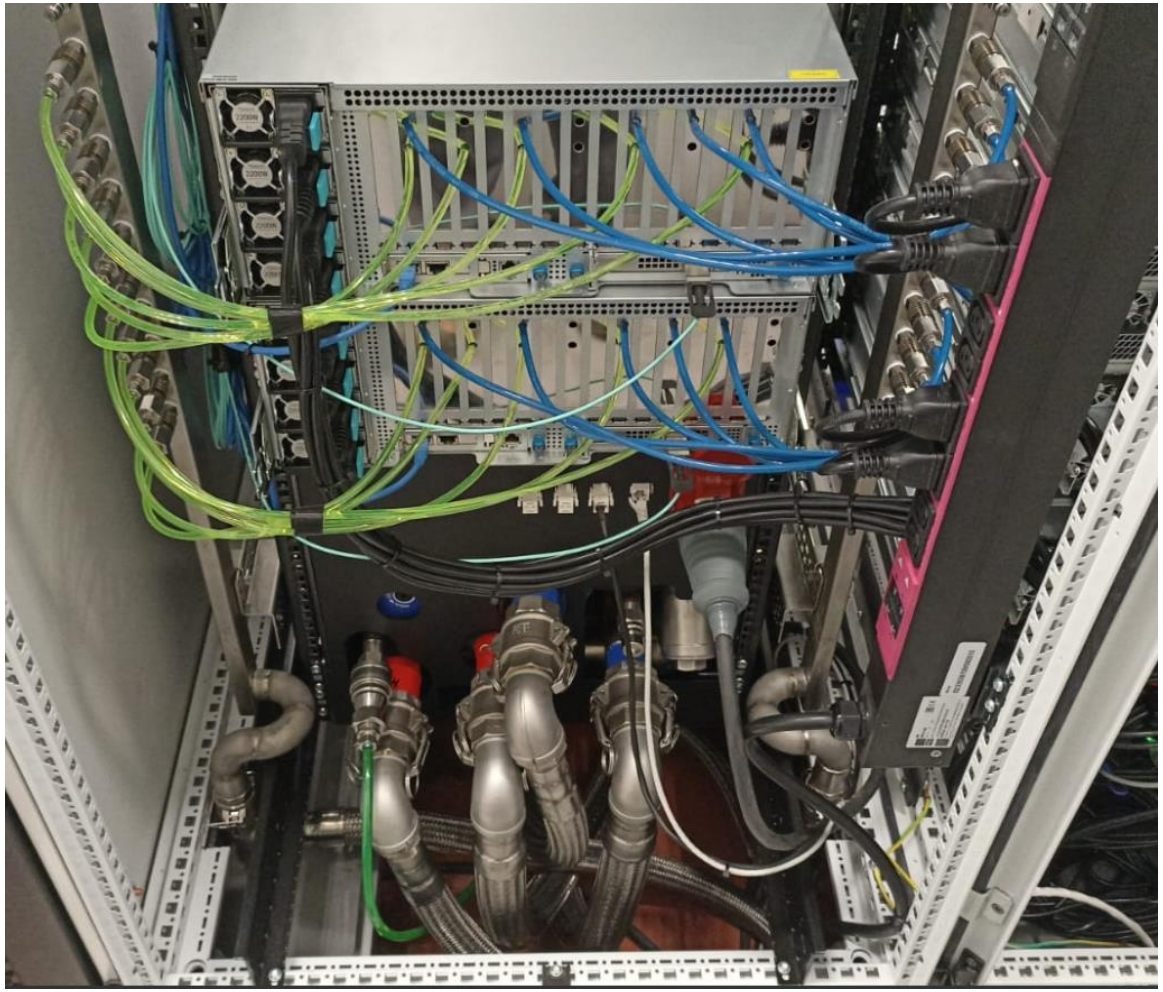
СХД 8ПБ в ОИЯИ





МикроЦОД

XXXX



Реализованные проекты

Южно-Уральский государственный университет в 2010году. В 2013 году после модернизации занимал 127-е место в списке Топ500 (ноябрь 2013 г).

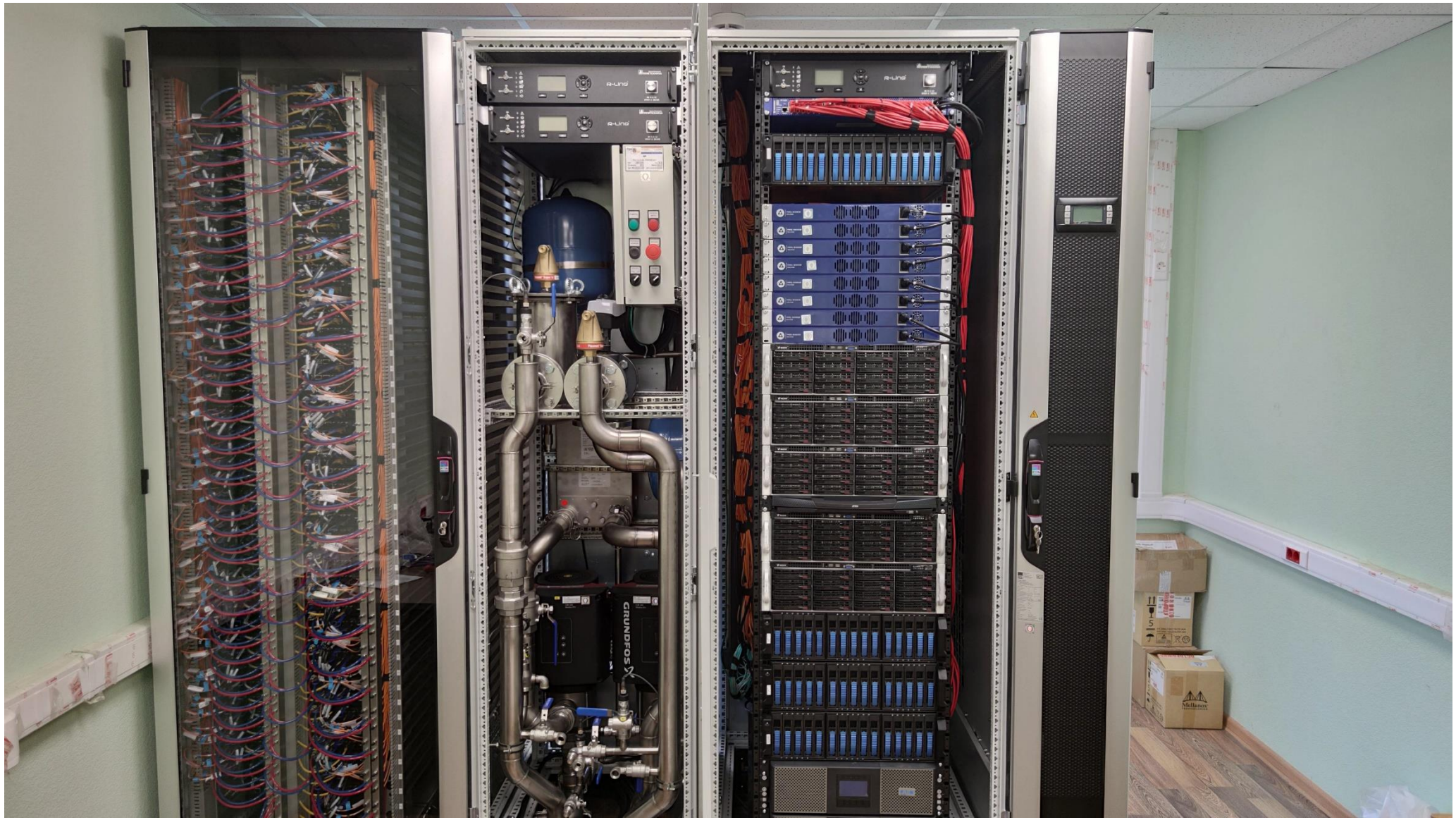
Межведомственный суперкомпьютерный центр РАН

Реализованные проекты

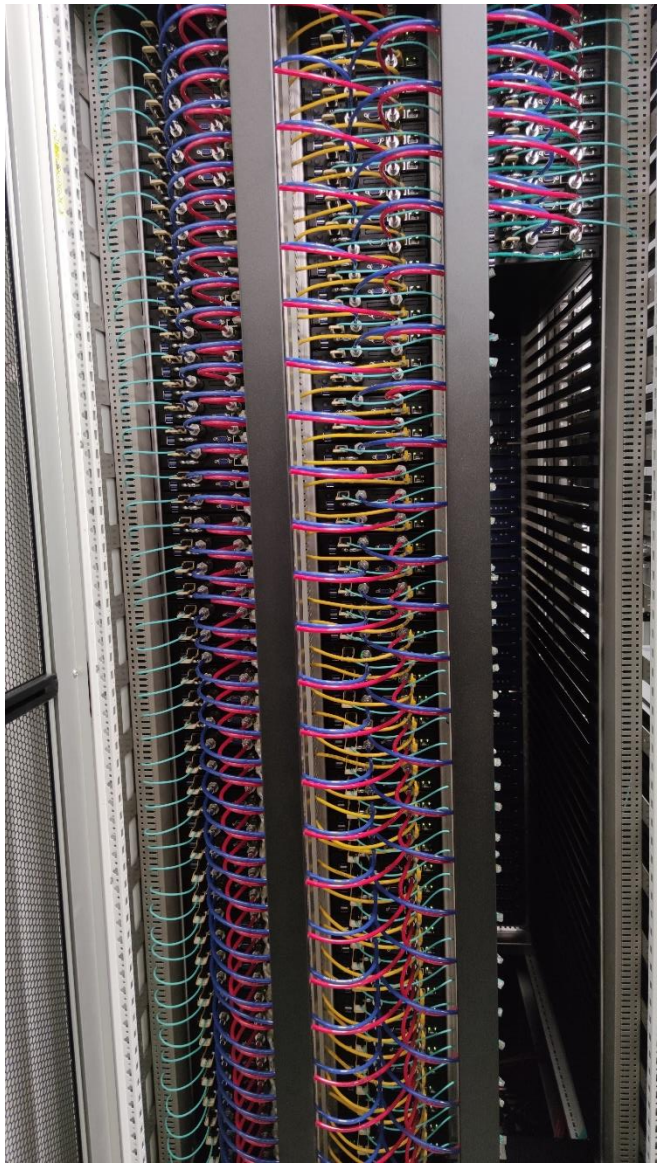
Объединенный институт ядерных исследований (ОИЯИ), Дубна

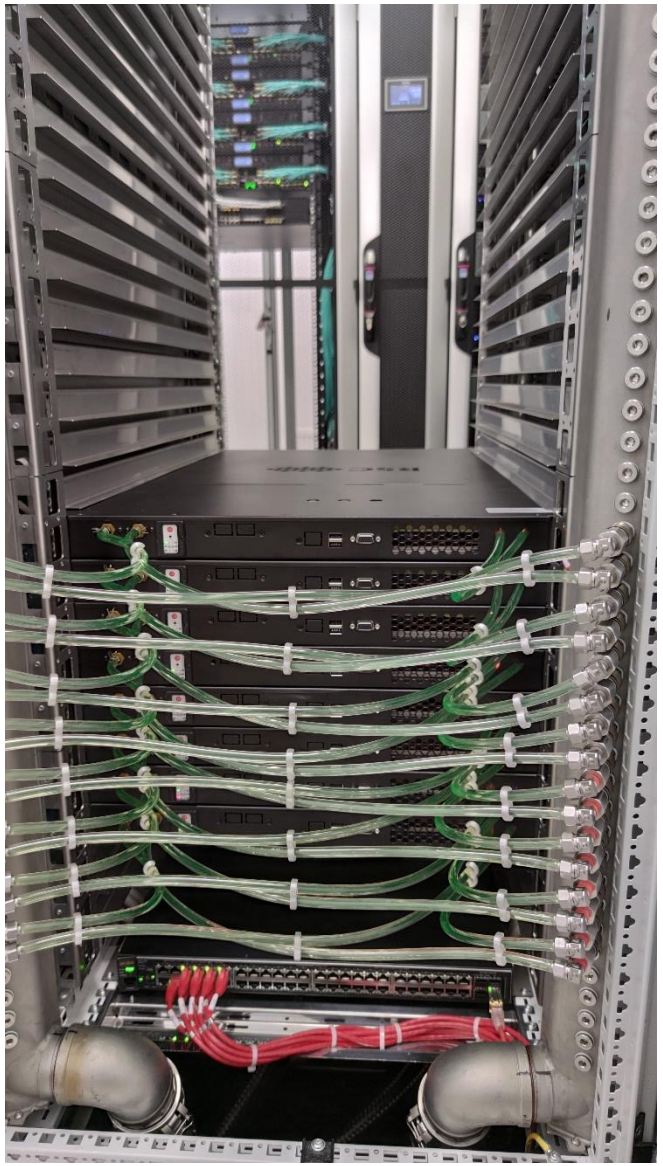
Санкт-Петербургский политехнический университет имени Петра Великого (СПбПУ)

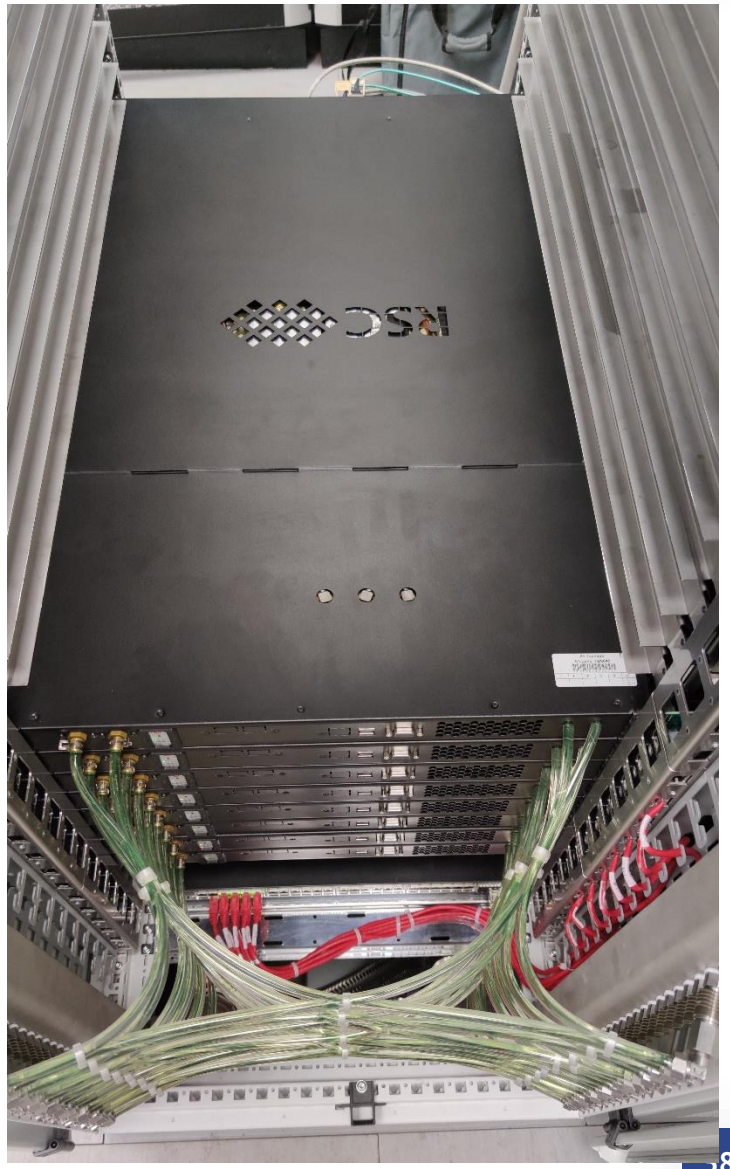
Реализованные проекты. Автономный ЦОД ЦИАМ.



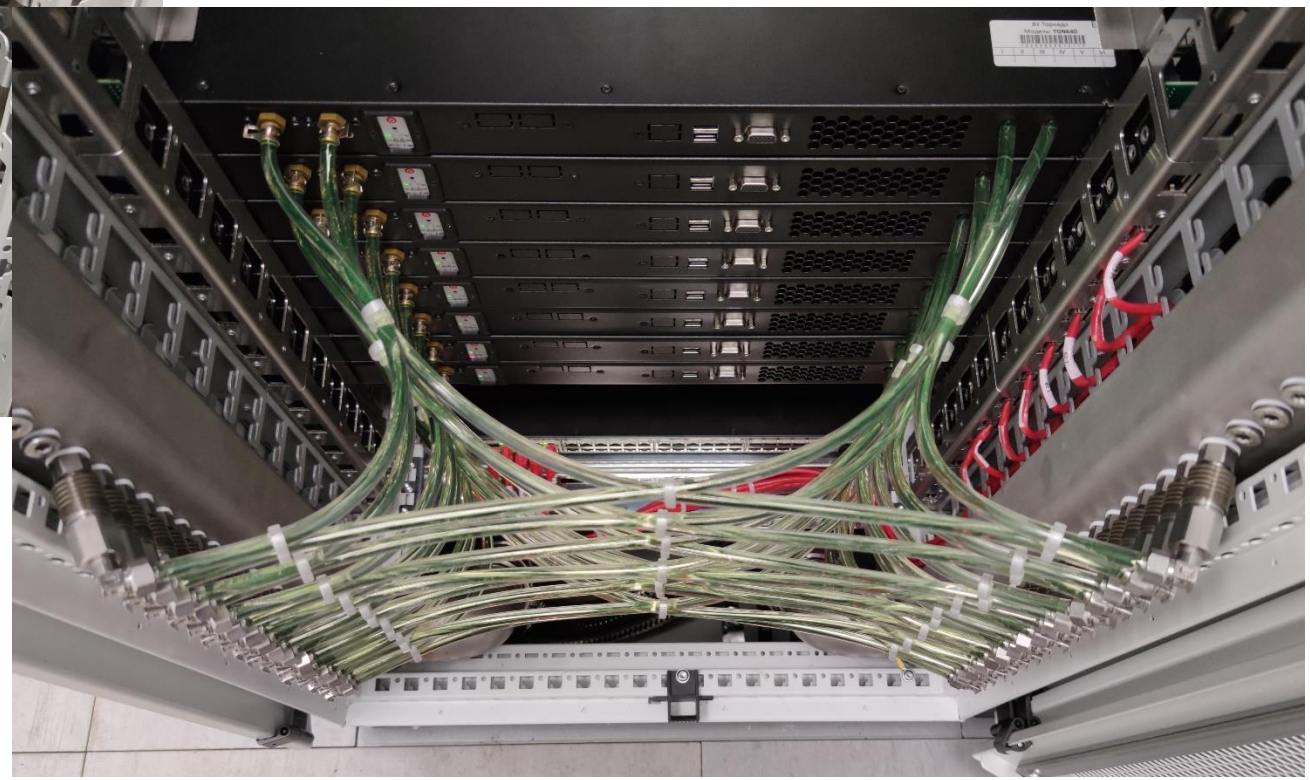
Вычислитель, ЦИАМ.



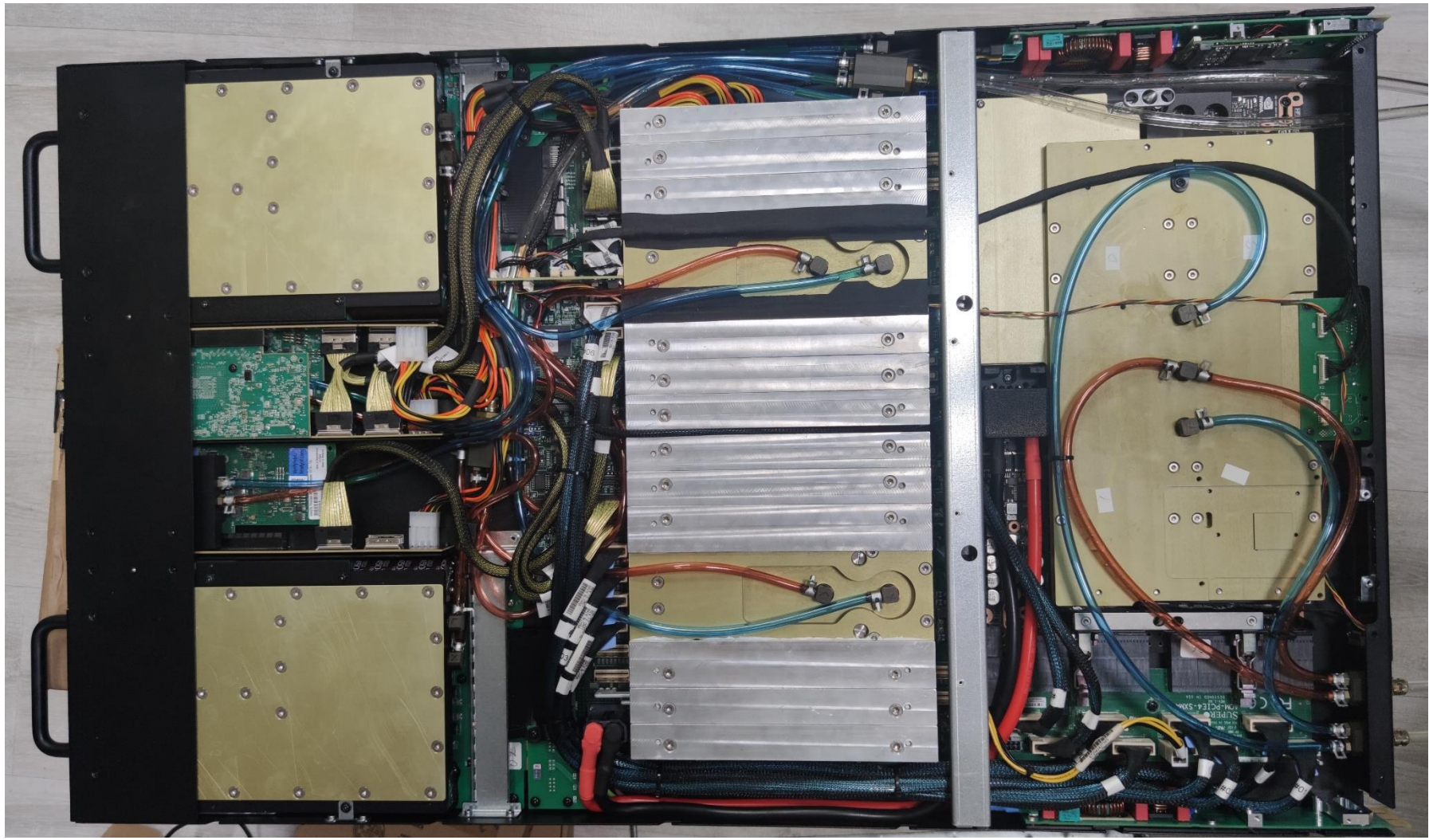




Сервера с ускорителями A100, ЦИАМ.



Сервер с ускорителями 4 x A100 в формате SXM



Спасибо!

РСК

rscgroup.ru
migal@rsc-systems.ru