РСК
Группа компаний

Реальная фантастика vs фантастическая реальность
или
Энергонагруженность - от фантастики к реальности.

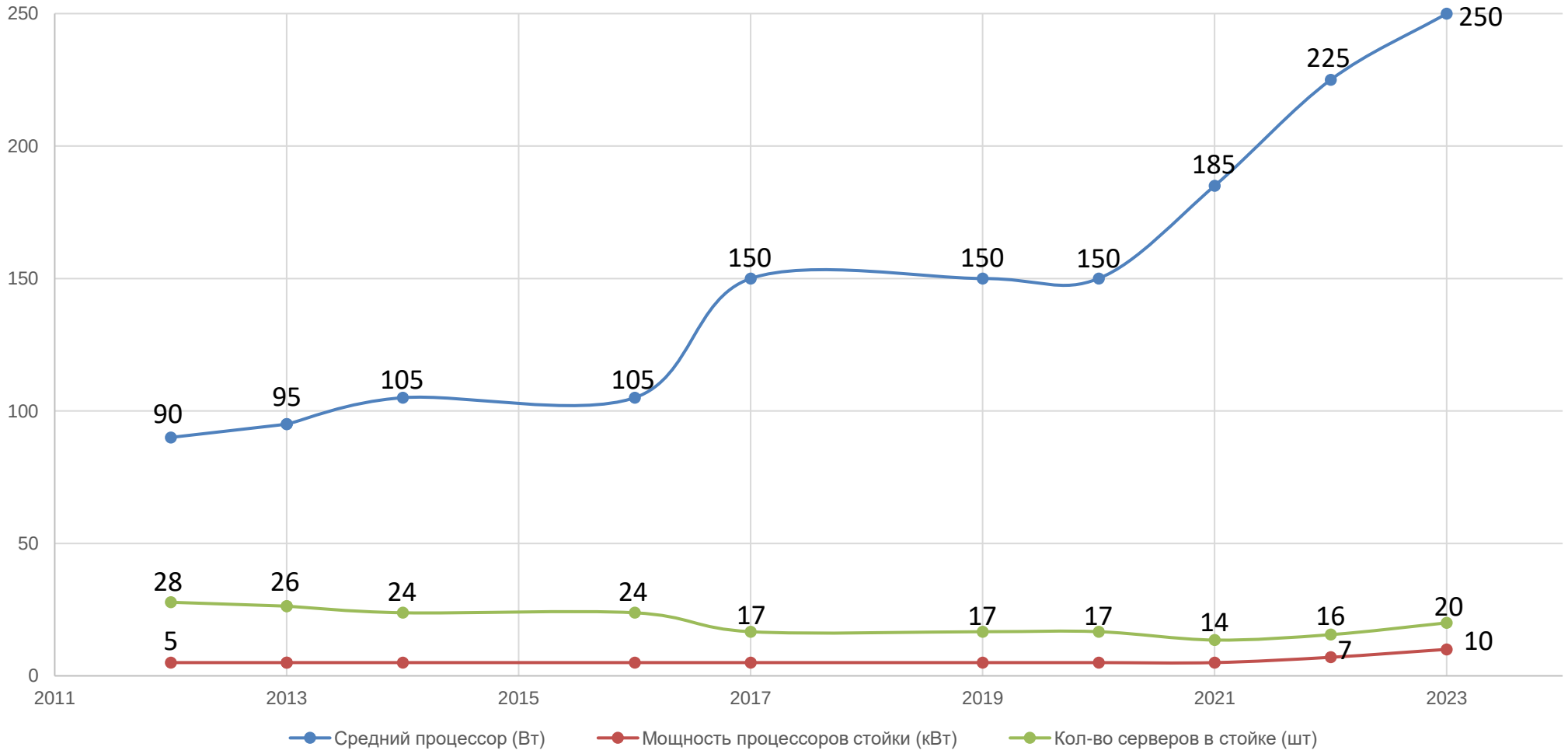
Юрий Мигаль
Руководитель департамента внедрения и эксплуатации

2024

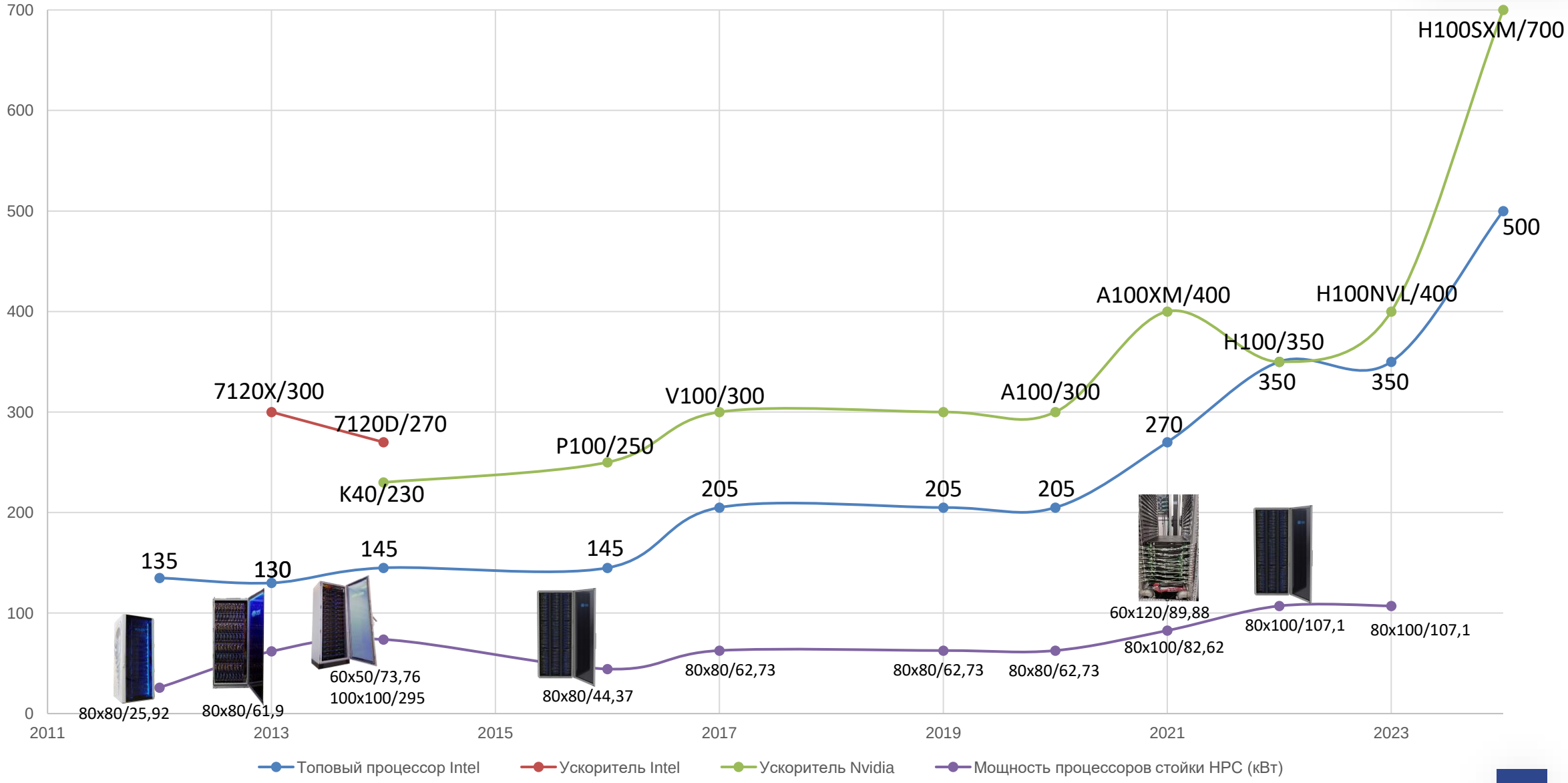
Высокопроизводительные системы с 2009 года

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

Реальная фантастика – «обычный» ЦОД



Фантастическая реальность – технологии из HPC



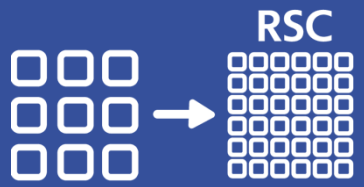
Первая десятка ТОП500. «Облака» в тройке лидеров.

Rank	Previous Rank	First Rank	Name	Computer	Manufacturer	Year	Segment	Total Cores	Accelerator /Co-Processor Cores	Power (kW)	Accelerator/Co-Processor	Interconnect	Continent
1	1	1	Frontier	HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	HPE	2021	Research	8 699 904	8 138 240	22 786,00	AMD Instinct MI250X	Slingshot-11	North America
2	2	2	Aurora	HPE Cray EX - Intel Exascale Compute Blade, Xeon CPU Max 9470 52C 2.4GHz, Intel Data Center GPU Max, Slingshot-11	Intel	2023	Research	9 264 128	8 159 232	38 698,36	Intel Data Center GPU Max	Slingshot-11	North America
3	3	3	Eagle	Microsoft NDv5, Xeon Platinum 8480C 48C 2GHz, NVIDIA H100, NVIDIA Infiniband NDR	Microsoft Azure	2023	Industry	2 073 600	1 900 800		NVIDIA H100	NVIDIA Infiniband NDR	North America
4	4	1	Supercomputer Fugaku	Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D	Fujitsu	2020	Research	7 630 848		29 899,23	None	Tofu interconnect D	Asia
5	5	3	LUMI	HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	HPE	2023	Research	2 752 704	2 566 080	7 106,82	AMD Instinct MI250X	Slingshot-11	Europe
6		6	Alps	HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip, Slingshot-11	HPE	2024	Research	1 305 600	844 800	5 194,00	NVIDIA GH200 Superchip	Slingshot-11	Europe
7	6	4	Leonardo	BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64 GB, Quad-rail NVIDIA HDR100 Infiniband	EVIDEN	2023	Academic	1 824 768	1 714 176	7 493,74	NVIDIA A100 SXM4 64 GB	Quad-rail NVIDIA HDR100 Infiniband	Europe
8	8	8	MareNostrum 5 ACC	BullSequana XH3000, Xeon Platinum 8460Y+ 32C 2.3GHz, NVIDIA H100 64GB, Infiniband NDR	EVIDEN	2023	Research	663 040	591 360	4 158,90	NVIDIA H100 64GB	Infiniband NDR	Europe
9	7	1	Summit	IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband	IBM	2018	Research	2 414 592	2 211 840	10 096,00	NVIDIA Volta GV100	Dual-rail Mellanox EDR Infiniband	North America
10	9	9	Eos NVIDIA DGX SuperPOD	NVIDIA DGX H100, Xeon Platinum 8480C 56C 3.8GHz, NVIDIA H100, Infiniband NDR400	Nvidia	2023	Vendor	485 888	439 296		NVIDIA H100 SXM5 80GB	Infiniband NDR400	North America

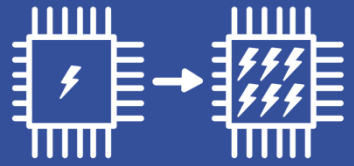
НАШИ!!!

Rank	Previous Rank	First Rank	Name	Computer	Site	Manufacturer	Country	Year	Accelerator/ Co-Processor Cores	Accelerator/Co-Processor	System Model
42	35	19	Chervonenkis	YANDEX Y4N-GA1-TY25-ZB0, AMD EPYC 7702 64C 2GHz, NVIDIA A100 80GB, Infiniband	Yandex	YANDEX, NVIDIA	Russia	2021	168 480	NVIDIA Tesla A100 80G	YANDEX Y4N-GA1-TY25-ZB0
69	57	36	Galushkin	YANDEX Y4N-GA1-TY25-ZB0, AMD EPYC 7702 64C 2GHz, NVIDIA A100 80GB, Infiniband	Yandex	YANDEX, NVIDIA	Russia	2021	117 504	NVIDIA Tesla A100 80G	YANDEX Y4N-GA1-TY25-ZB0
79	63	40	Lyapunov	Inspur NF5488A5, AMD EPYC 7662 64C 2GHz, NVIDIA A100 40GB, Infiniband	Yandex	NVIDIA, Inspur	Russia	2020	114 048	NVIDIA Tesla A100 40G	Inspur NF5488A5
83	66	43	Christofari Neo	NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100 80GB, Infiniband	SberCloud	Nvidia	Russia	2021	85 536	NVIDIA Tesla A100 80G	NVIDIA DGX A100
142	118	29	Christofari	NVIDIA DGX-2, Xeon Platinum 8168 24C 2.7GHz, Mellanox InfiniBand EDR, NVIDIA Tesla V100	SberCloud	Nvidia	Russia	2019	96 000	NVIDIA Tesla V100	NVIDIA DGX-2
406	369	130	Lomonosov 2	T-Platform A-Class Cluster, Xeon E5-2697v3 14C 2.6GHz,Intel Xeon Gold 6126, Infiniband FDR, Nvidia K40m/P-100	Moscow State University - Research Computing Center	T-Platforms	Russia	2014	40 960	NVIDIA Tesla K40m	T-Platforms A-Class Cluster
472	432	240	MTS GROM	NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100 40GB, Infiniband	#CloudMTS	Nvidia	Russia	2021	17 280	NVIDIA Tesla A100 40G	NVIDIA DGX A100

Ключевые требования к высокопроизводительным системам (HPC)



Вычислительная плотность



Энергетическая плотность

Энергоэффективность

Легкость управления и обслуживания

Надежность

Тенденции современных ЦОД

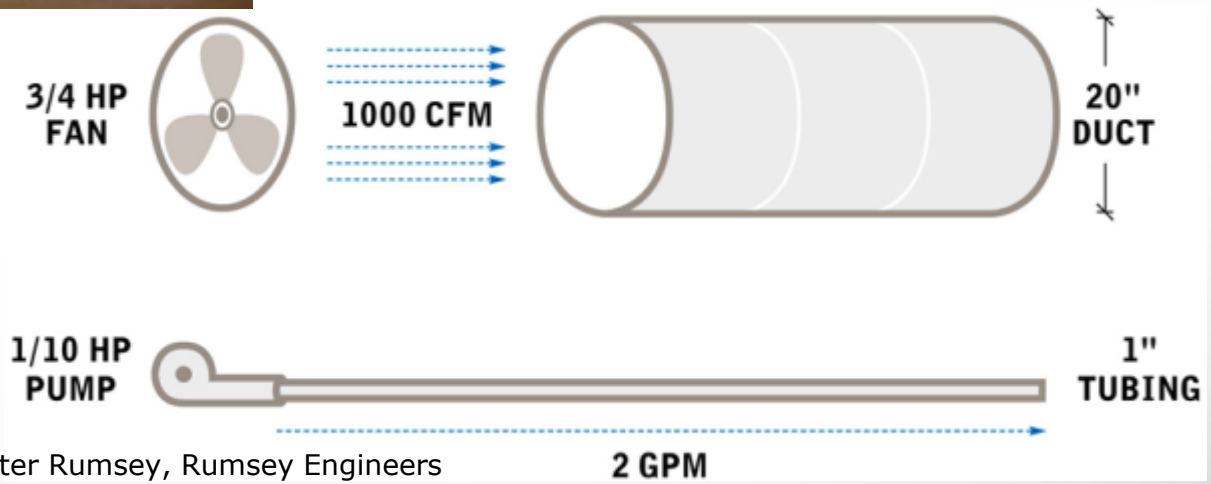
- Увеличение мощности единицы оборудования
- Увеличение нагрузки на стойку
- «Зеленость»
- Увеличение температуры ЦОД
- Снижение PUE

Почему жидкостное охлаждение?

- Воздушные вентиляторы менее эффективны
- Высокое аэродинамическое сопротивление высоконагруженных стоек
- Сложности с режимом свободного охлаждения

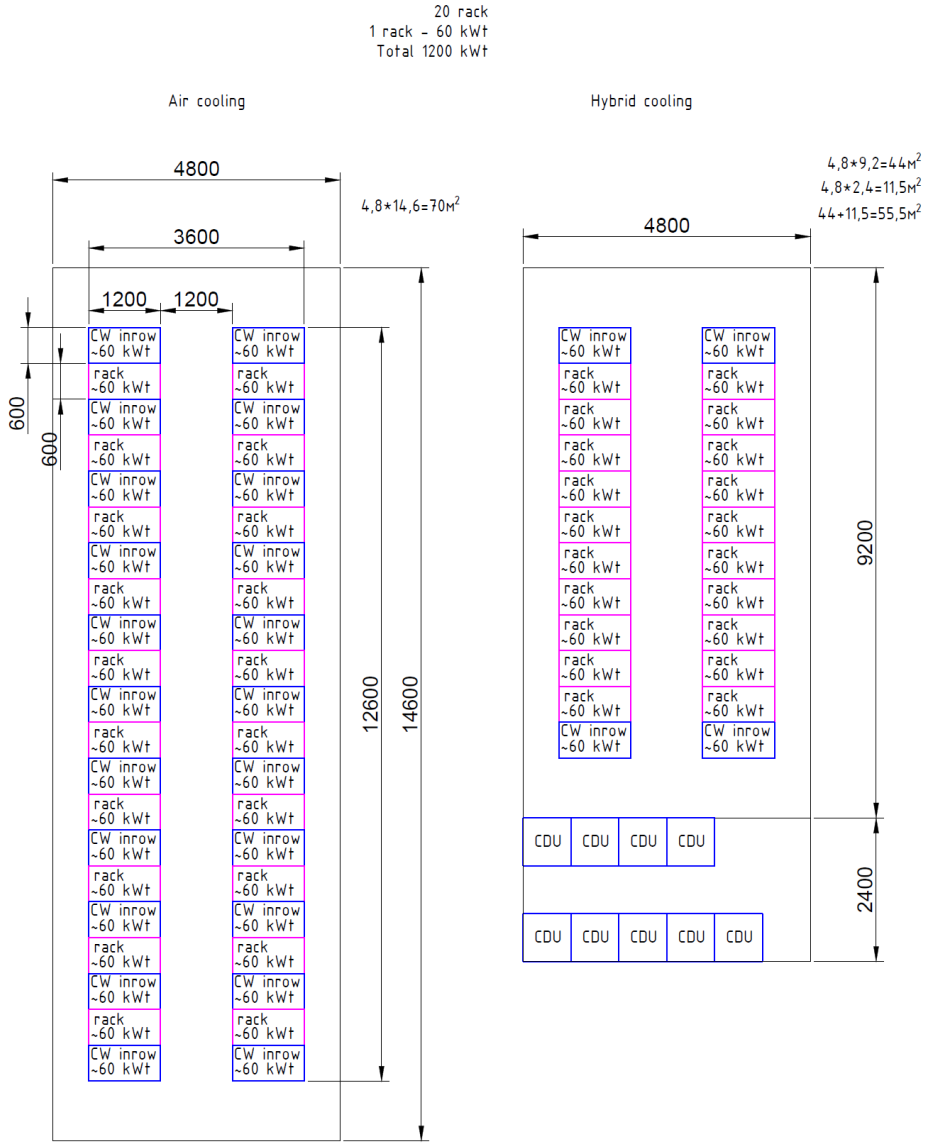
- Повышение плотности в HPC инсталляциях

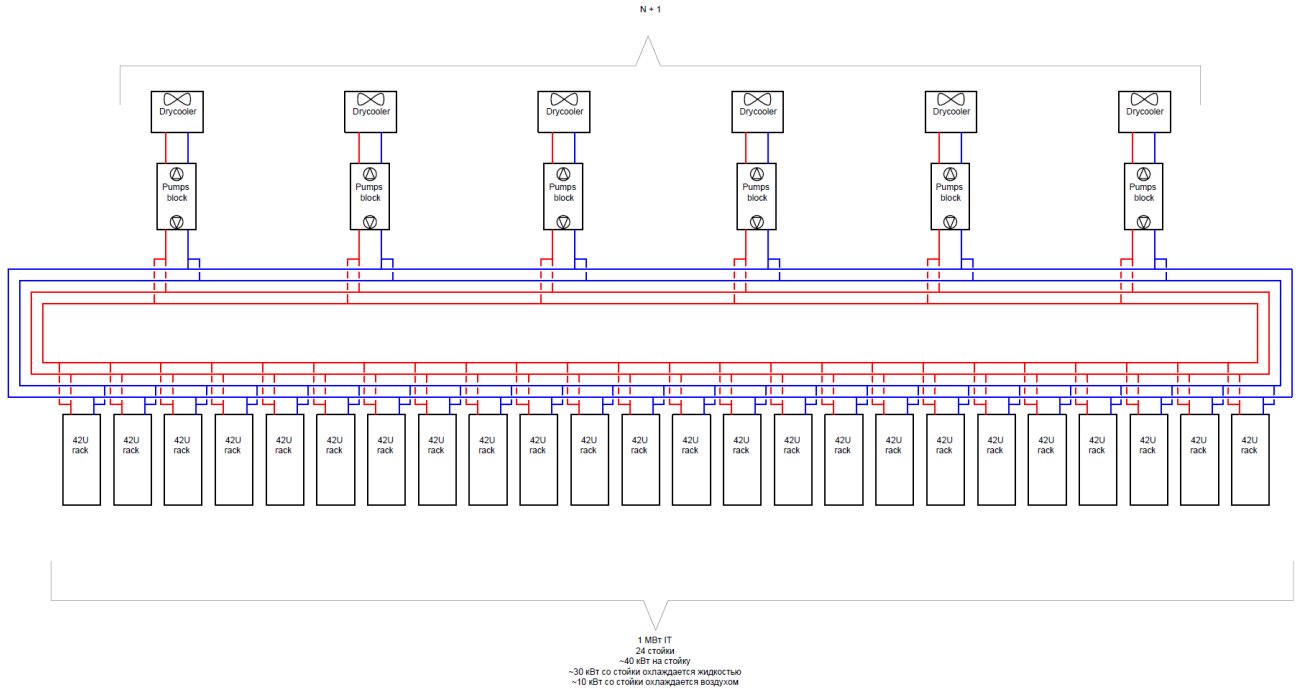
=> 80kW+ в шкафу для ExaScale



Source: Overview of Liquid Cooling Systems, Peter Rumsey, Rumsey Engineers
Available at: hightech.lbl.gov/presentations/Dominguez/5_LiquidCooling_101807.ppt

От фантастики к реальности.

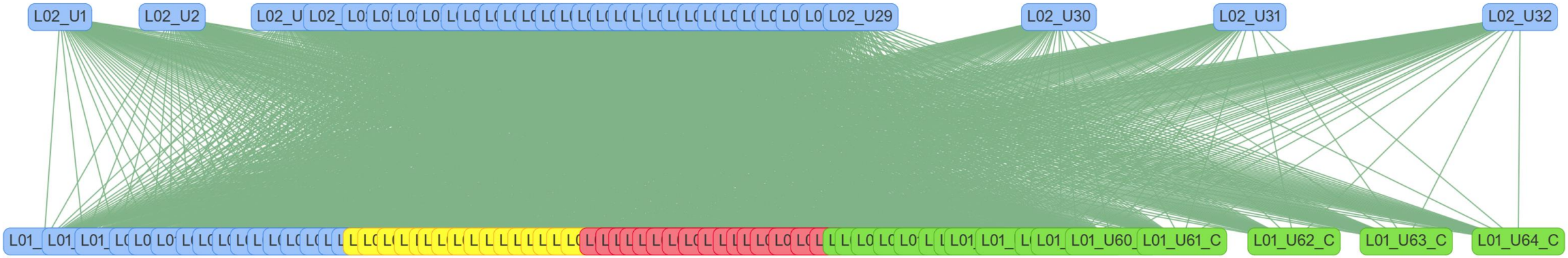


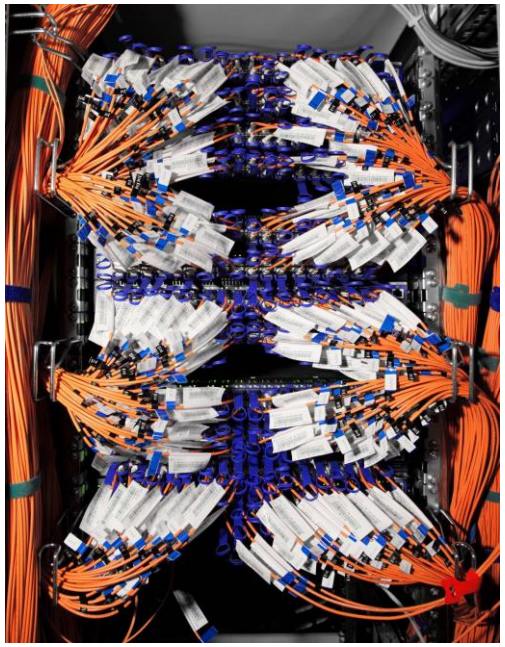


Высоконагруженные стойки с гибридным охлаждением:

- **Уменьшение занимаемой площади**
- **Уменьшение количества оборудования инфраструктуры**
- **Снижение CAPEX**
- **Снижение PUE**
- **Снижение OPEX**
- **Возможность сертификации Tier III**

Почему «плотность» имеет значение

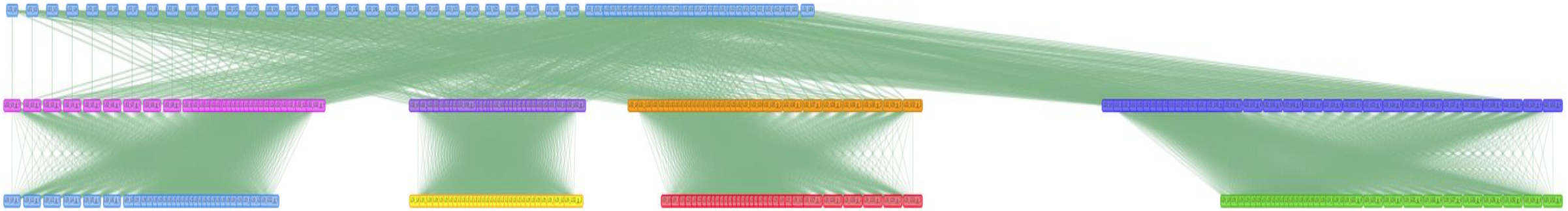




Сеть накладывает свои ограничения:

- **«Дерево без блокировок» – самое простое до 2048 портов (256 современных серверов)**
- **Каждый коммутатор 1,6кВт**
- **«Корень» – 32 коммутатора и 51,2кВт воздушного охлаждения**

А больше можно?

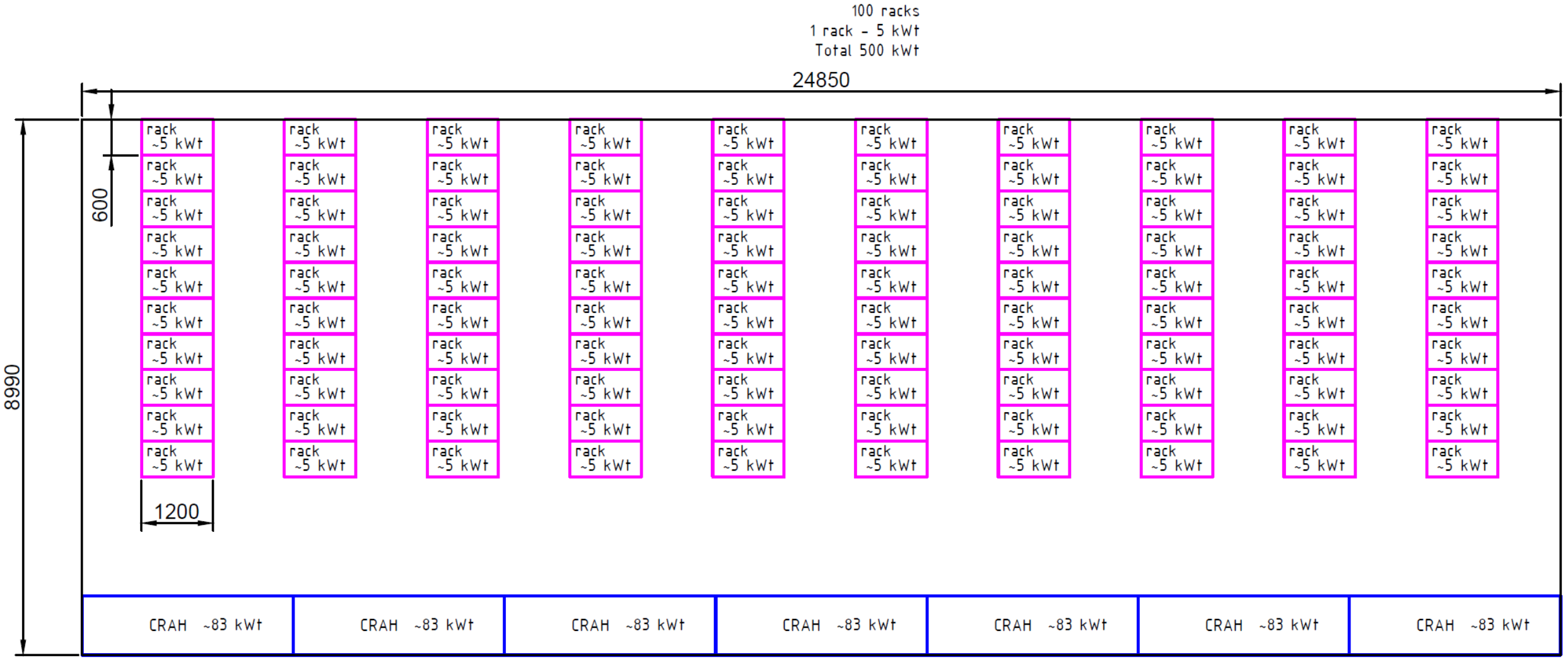


Любой каприз за Ваши деньги!

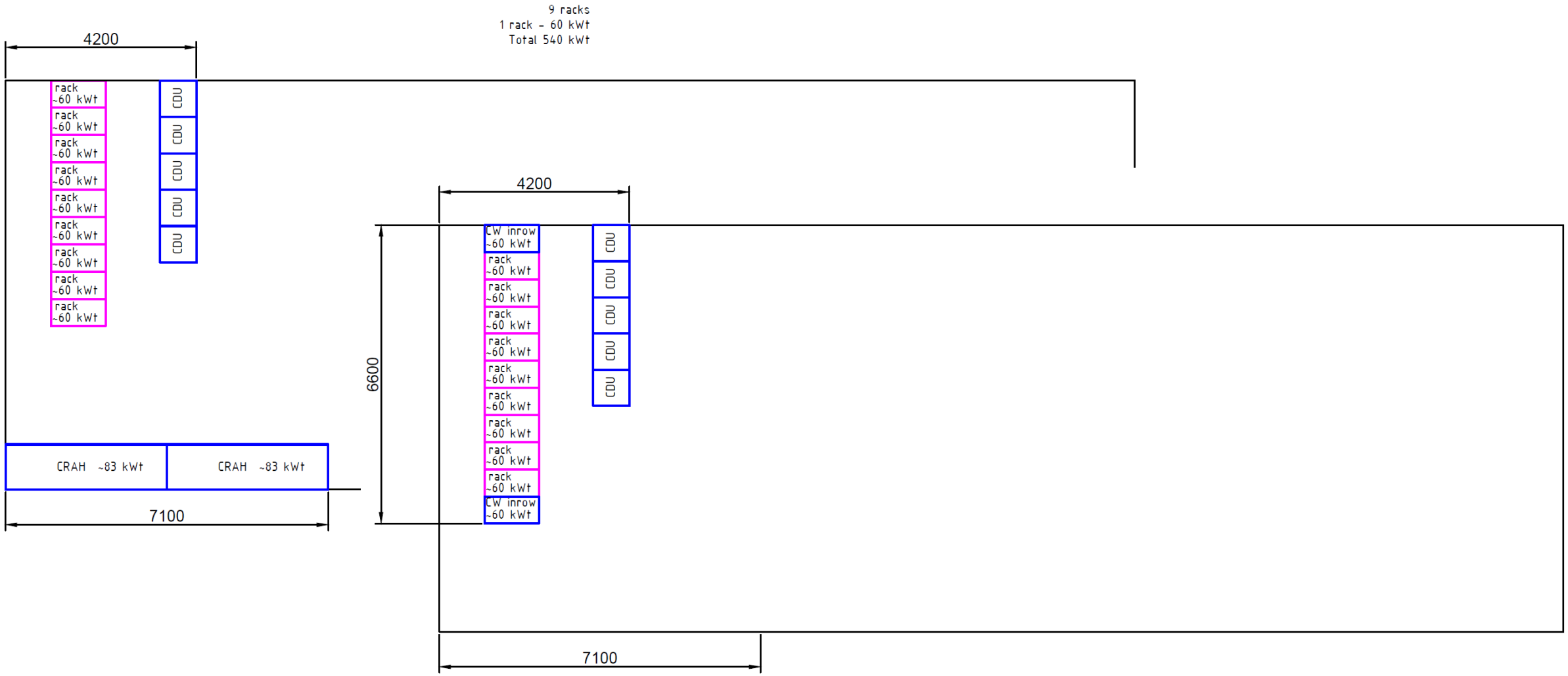
4096 портов выглядят так:

- **1-й и второй уровни – 128 коммутаторов каждый**
- **3-й и уровень («Корень») – 64 коммутатора**

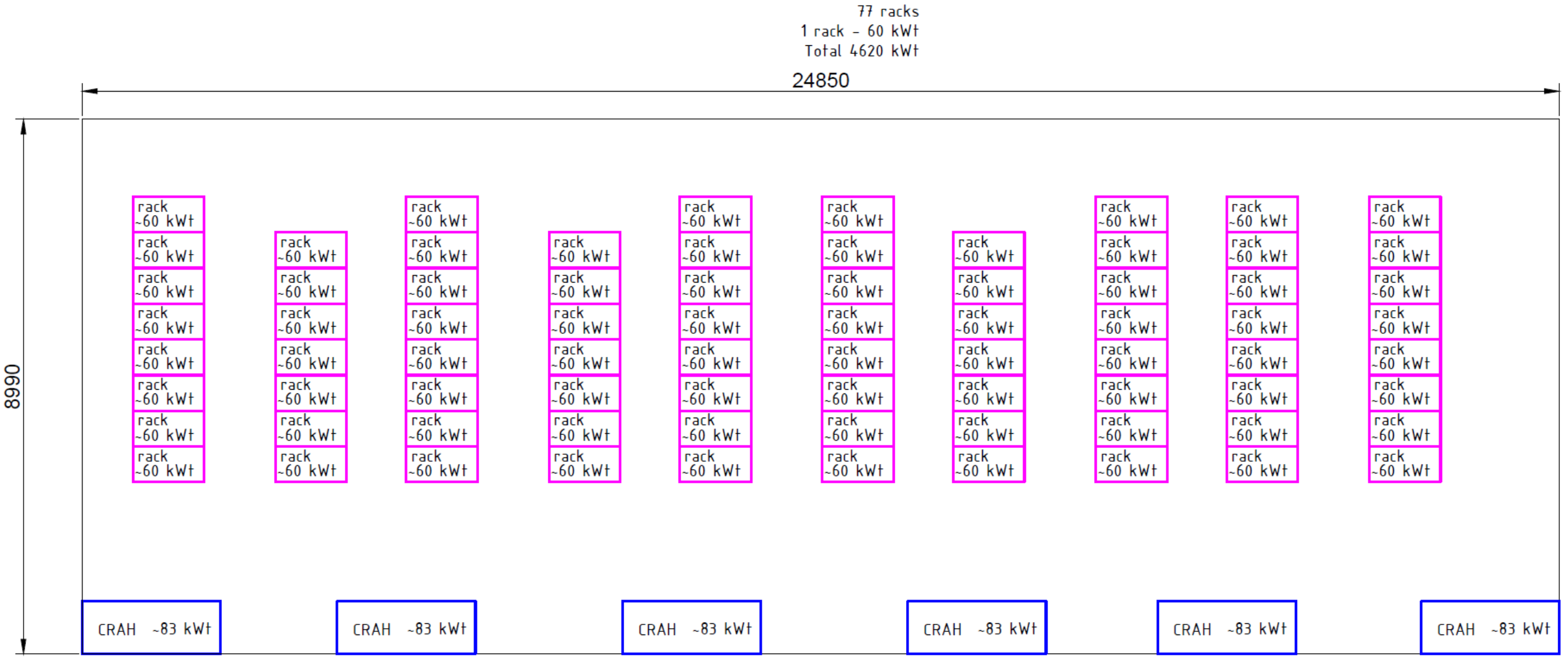
Любая цель должна быть достижима!



Любая цель должна иметь смысл!



Не останавливайся перед препятствиями!



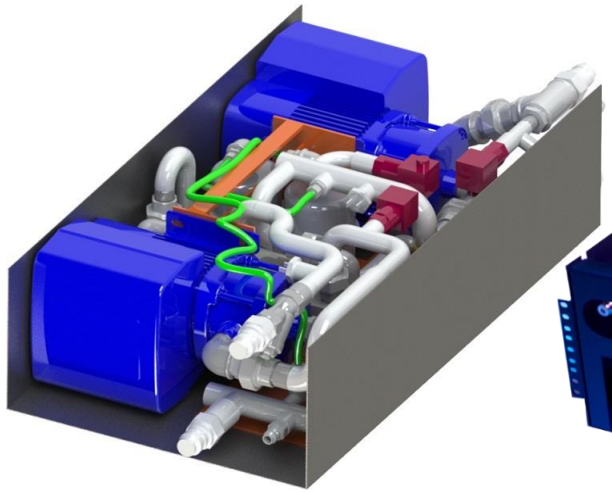
Нынешняя реальность

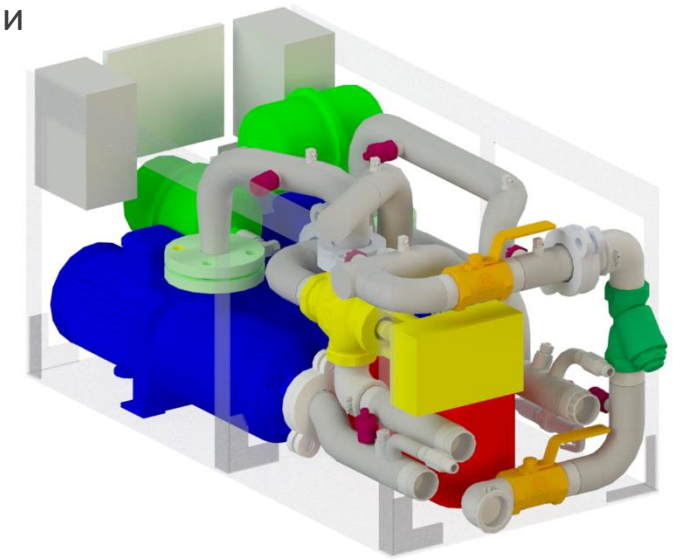
	Вт/ сервер	Сервер/ стойку	Стойка	Жидкость/ сервер	Жидкость/ стойку	Вентиляторы/ сервер	Вентиляторы/ стойку	Воздух/ сервер	Воздух/ стойку	Стойка + жидкость
Сервер 1U, 2 ЦПУ	1100 - 1200 Вт	42	50,4 кВт	ЦПУ: 2*350=700	29,4 кВт	200 Вт	8,4 кВт	300 Вт	12,6 кВт	42 кВт
Сервер 2U, 2 ЦПУ, 4 ГПУ	2900 Вт	21	60,9 кВт	ЦПУ: 2*350=700 ГПУ: 4*350=1400 Итого: 2100	44,1 кВт	500 Вт	10,5 кВт	300 Вт	6,3 кВт	50,4 кВт
Сервер 4U, 2 ЦПУ, 8 ГПУ	4800 Вт	10	40 кВт	ЦПУ: 2*350=700 ГПУ: 8*350=2800 Итого: 3500	35 кВт	1000 Вт	10 кВт	300 Вт	3 кВт	38 кВт
Сервер 5U, 2 ЦПУ, HGX 8	10200 Вт	8	96 кВт	ЦПУ: 2*350=700 ГПУ: 8*700=5600 NVLink, NVMe: 600 Итого: 6900	55,2 кВт	2000 Вт	16 кВт	1,3 кВт	10,4 кВт	65,6 кВт
Сервер 8U, 2 ЦПУ, HGX 8	10200 Вт	5	60 кВт	ЦПУ: 2*350=700 ГПУ: 8*700=5600 NVLink, NVMe: 600 Итого: 6900	34,5 кВт	2000 Вт	10 кВт	1,3 кВт	6,5 кВт	41 кВт

Все технологии доступны уже сегодня!

Более 10 лет опыта применения жидкостного охлаждения в ЦОД

Теплообменно-насосные станции обеспечивают перенос тепловой энергии и разделяют гидравлические контуры 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 (насосные станции) для вычислительных систем

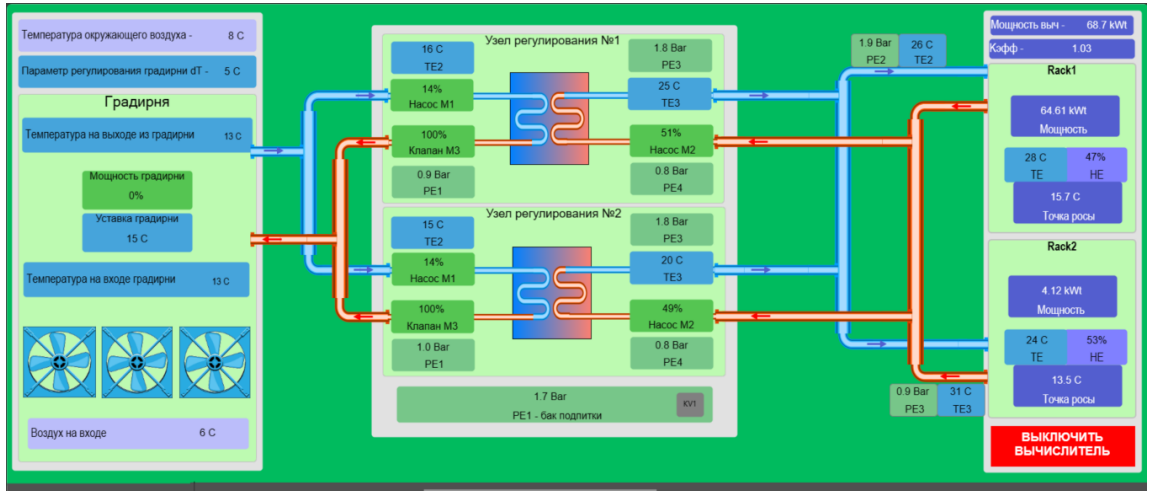
Отдельно стоящие насосные станции РСК представляют собой независимые изделия, в форм-факторе стойки.
Позволяют отвести тепловую мощность до 1 МВт.

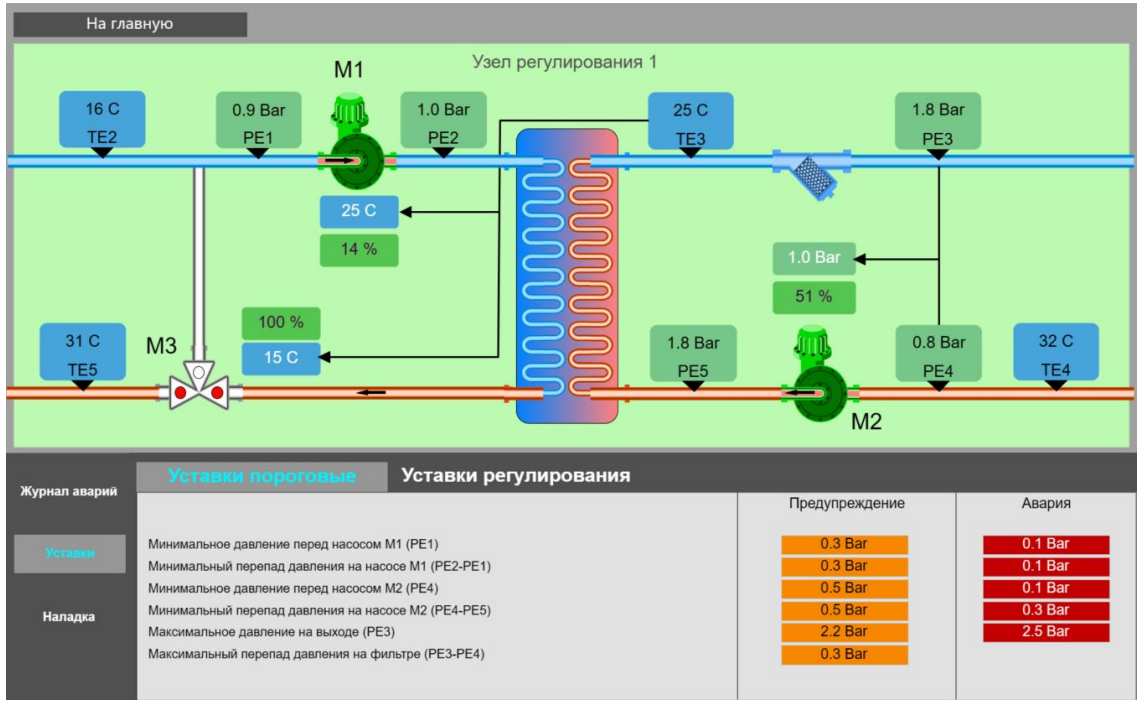
CDU 400

CDU-GM 400

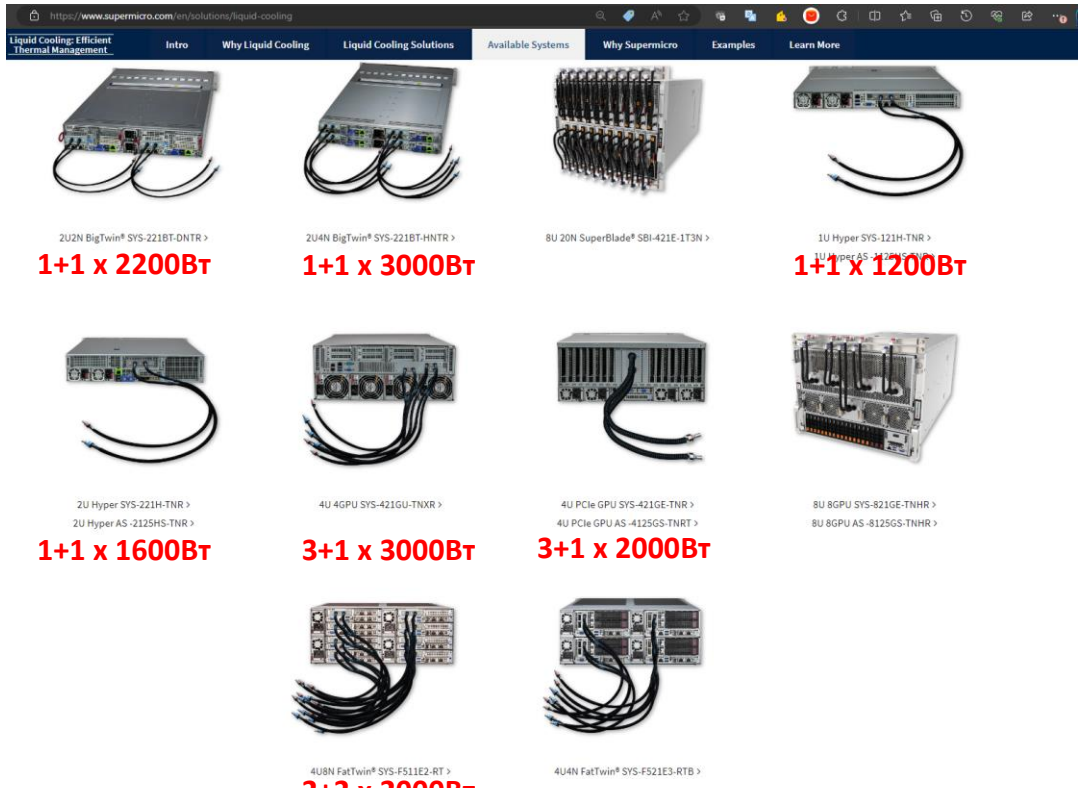
CDU-GM 600

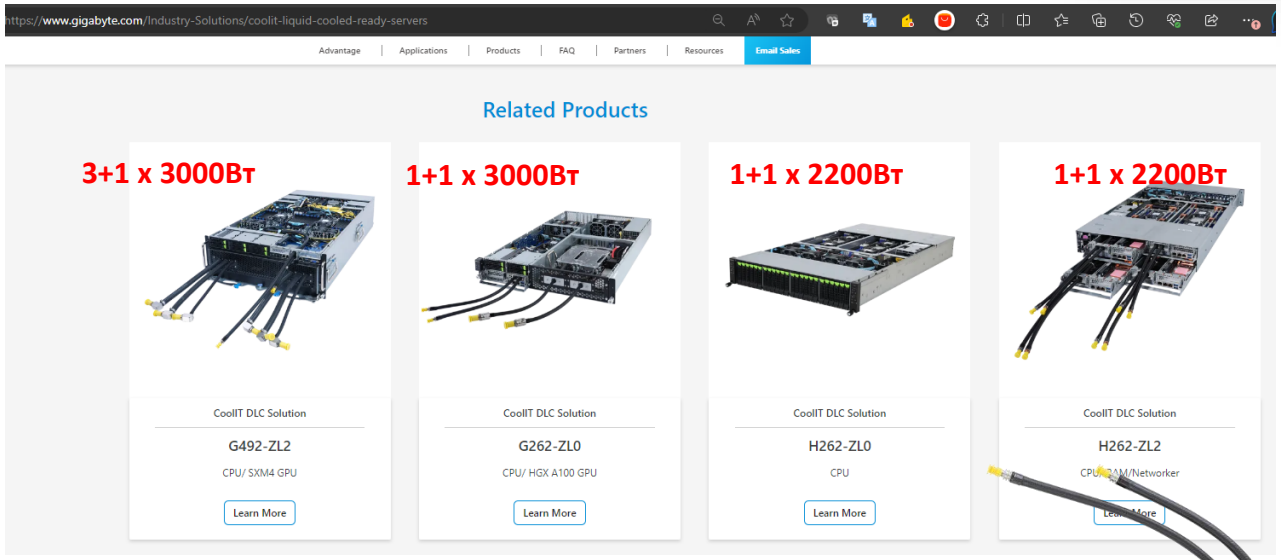
Мониторинг и управление инфраструктурой

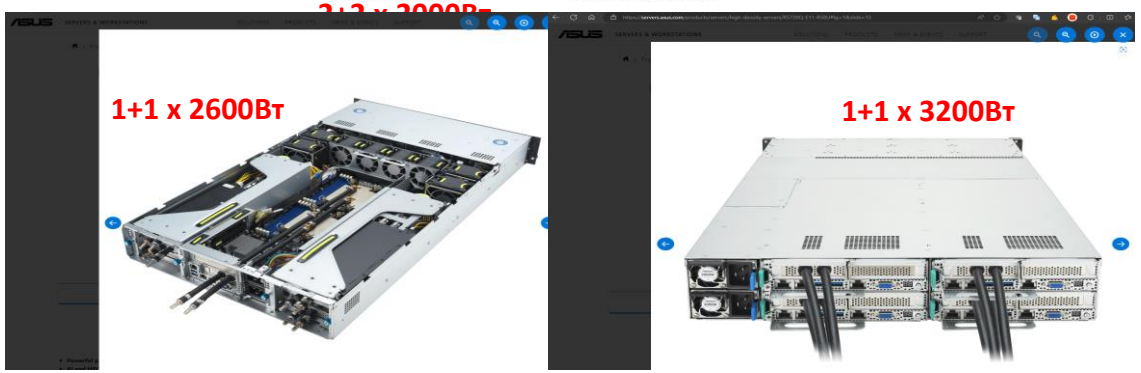


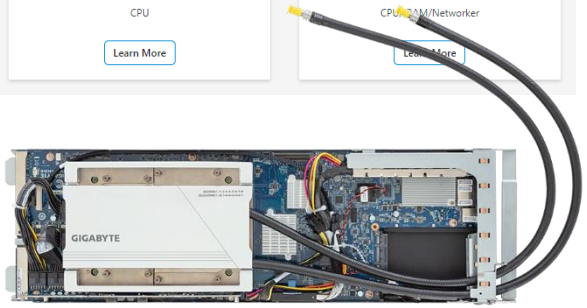


Если? – тут не уместно! Вопрос только – КОГДА?

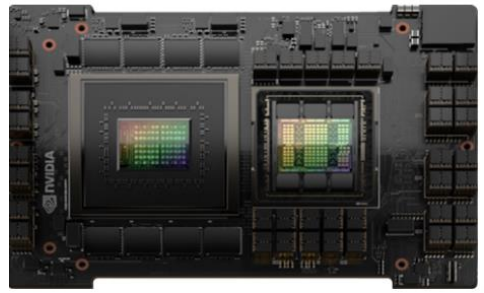




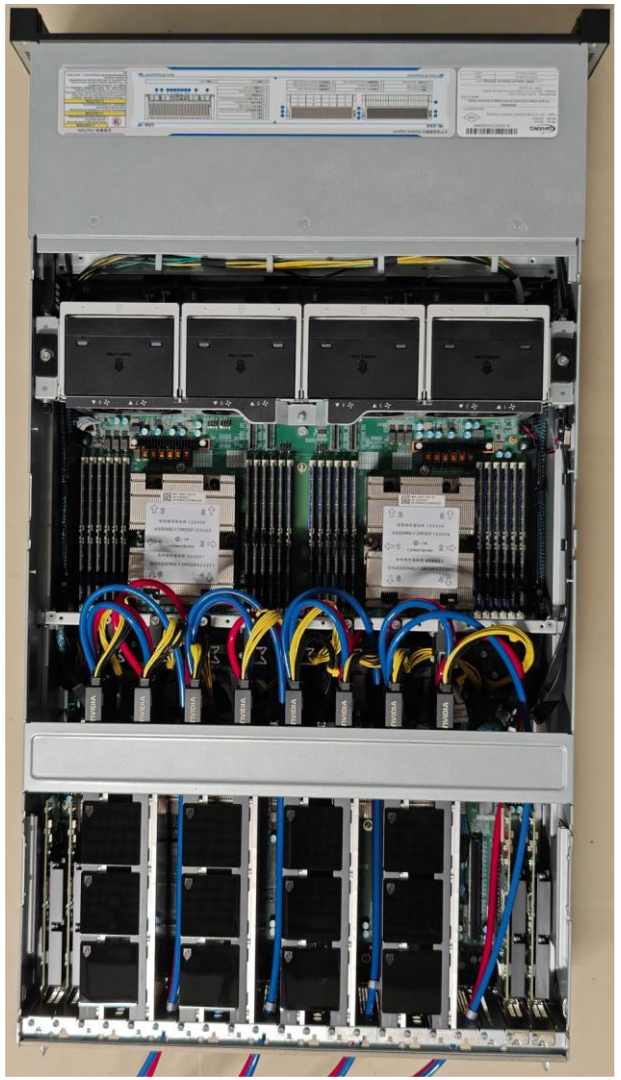




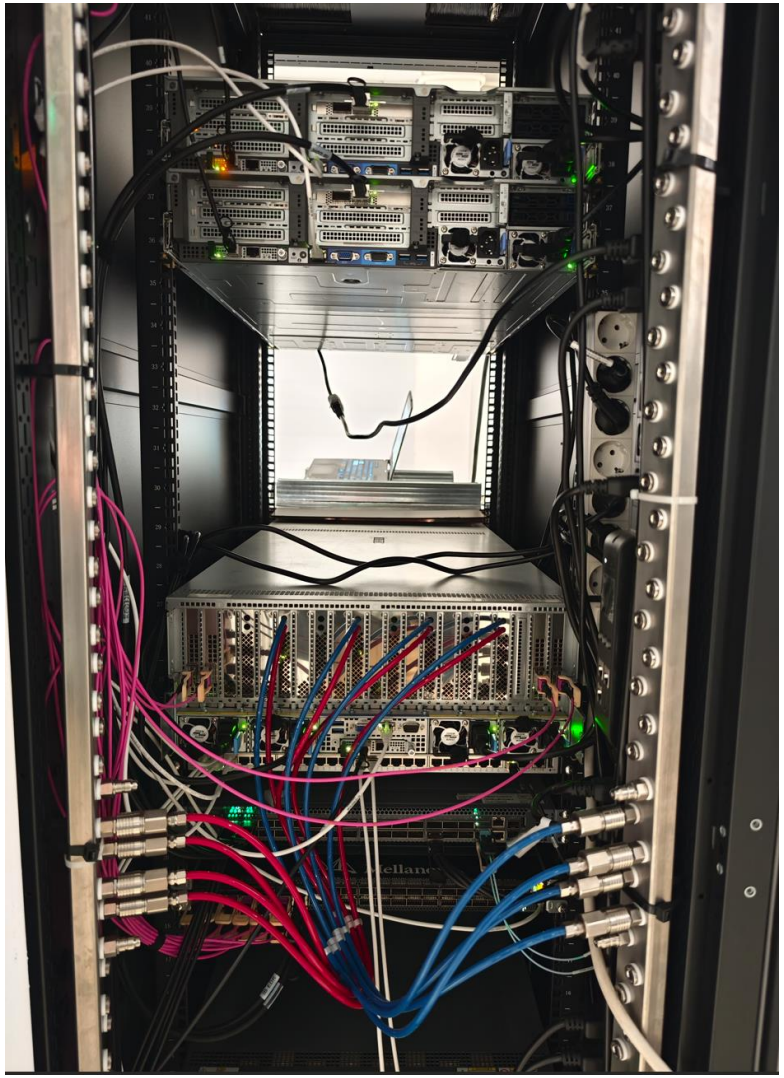
NVIDIA GH200 Grace Hopper Superchip



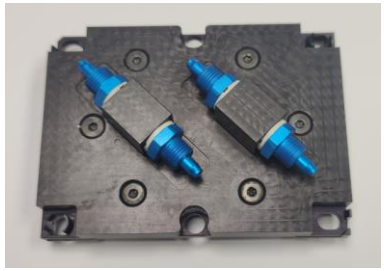
Пока Вы думаете – остальные пробуют!

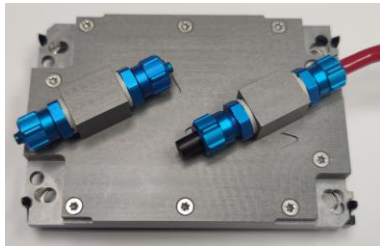


Пробуют на своей территории!



Занимайте очередь.





Спасибо!
www.rscgroup.ru
migal@rsc-tech.ru