



# Software Engineering Conference Russia

14-15 ноября, 2019. Санкт-Петербург

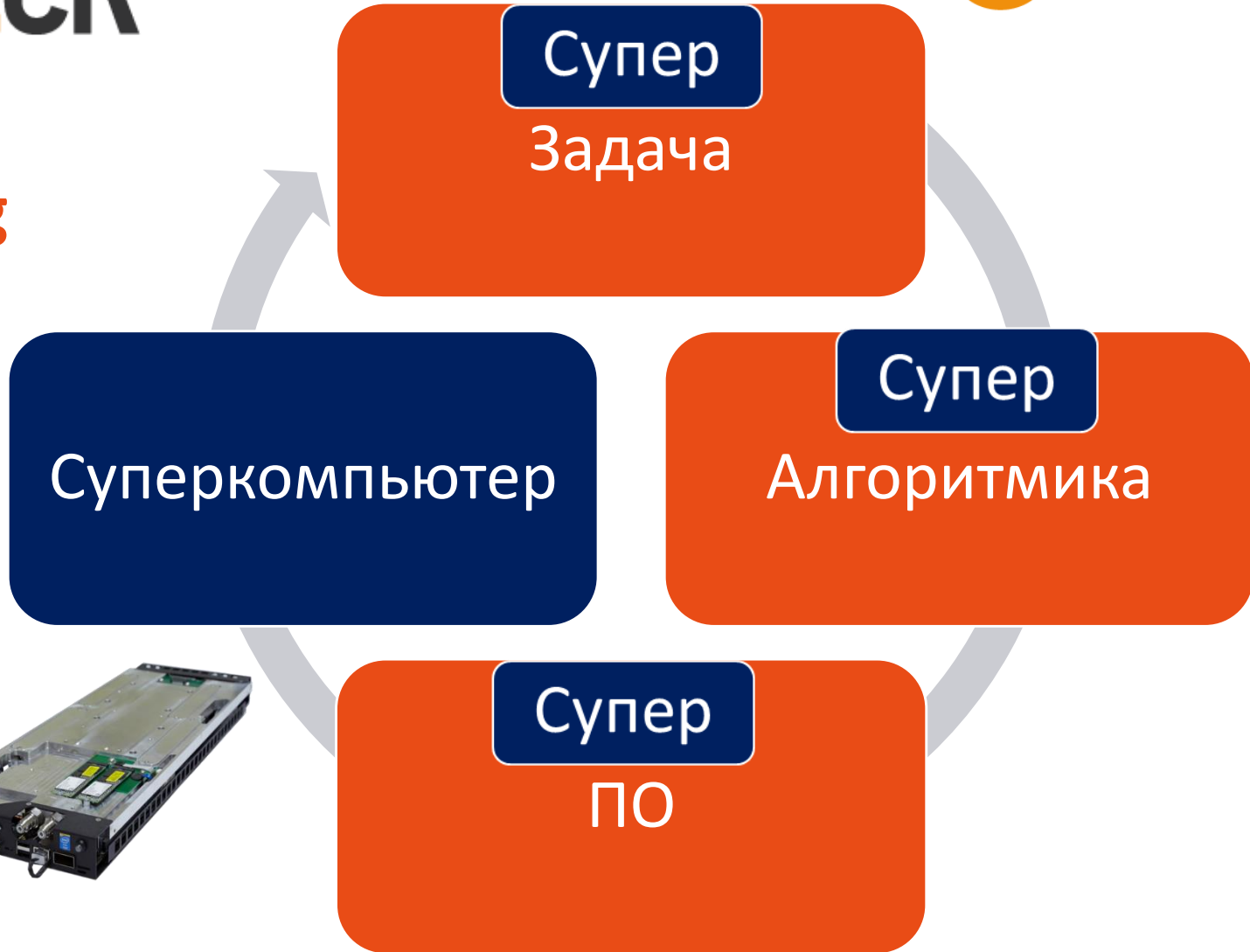
## Тенденции развития вычислительных узлов современных суперкомпьютеров

**Игорь Одинцов**

РСК



Software  
Engineering  
Conference  
Russia



# Современные суперкомпьютерные задачи

Производительность суперкомпьютеров		
Название	год	флопсы
флопс	1941	$10^0$
килофлопс	1949	$10^3$
мегафлопс	1964	$10^6$
гигафлопс	1987	$10^9$
терафлопс	1997	$10^{12}$
петафлопс	2008	$10^{15}$

Предметная область	
Аэродинамика	10 Petaflops
Лазерная оптика	20 Petaflops
Молекулярная динамика	200 Petaflops
Аэродинамический дизайн	1 Exaflops
Вычислительная космология	10 Exaflops
Турбулентность в физике	100 Exaflops
Вычислительная химия	1 Zettaflops

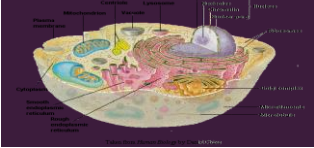
Классические мощные вычислительные задачи

Задачи из специфических предметных областей

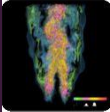
Обработка больших данных

Искусственный интеллект (нейронные сети)

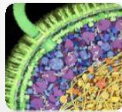
**Климат**  
предсказание локального влияния глобальных климатических сдвигов



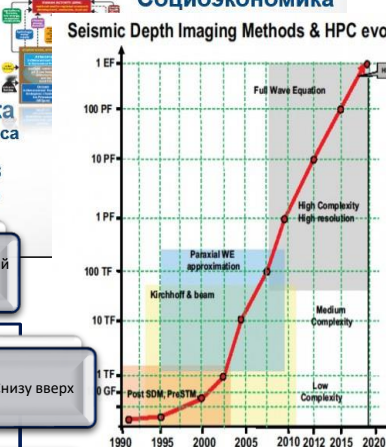
**Энергетика**  
- качественно новые ядерные реакторы  
- индивидуальные катализаторы  
- новые источники топлива




**Биология**  
- 100 микросекунд сворачивания белка за секунды вместо 3 лет на петафлопе  
- конструирование нано-механизмов с нуля



**Социэкономика**  
Seismic Depth Imaging Methods & HPC evolution




**Астрофизика**  
1 расчет процесса образования сверхновой за 8 недель при 20% реальной эффективности



**Искусственный интеллект**

- Сверху вниз
- Снизу вверх
- Экспертные системы
- Машинное обучение
- Глубокое обучение



Нет смысла описывать происходящее поэтому напишу: "У нас все хорошо,..."

# Современные суперкомпьютеры

5	«Политехник - РСК Торнадо»	734	623: CPU: 2x Intel Xeon E5-2697v3, 64 GB RAM	715.94
		1468	56: CPU: 2x Intel Xeon E5-2697v3, 64 GB RAM	1015.1
		128	Acc: 2x NVIDIA Tesla K40	
			36: CPU: 2x Intel Xeon E5-2697v3, 128 GB RAM	
			8: CPU: 2x Intel Xeon E5-2697v3, 128 GB RAM	
		Acc: 1x NVIDIA K1		
		8: CPU: 2x Intel Xeon E5-2697v3, 128 GB RAM		
		Acc: 1x NVIDIA K2		
		3: CPU: 2x Intel Xeon E5-2697v3, 256 GB RAM		

Суперкомпьютерный  
центр,  
Санкт-Петербургский  
политехнический  
университет

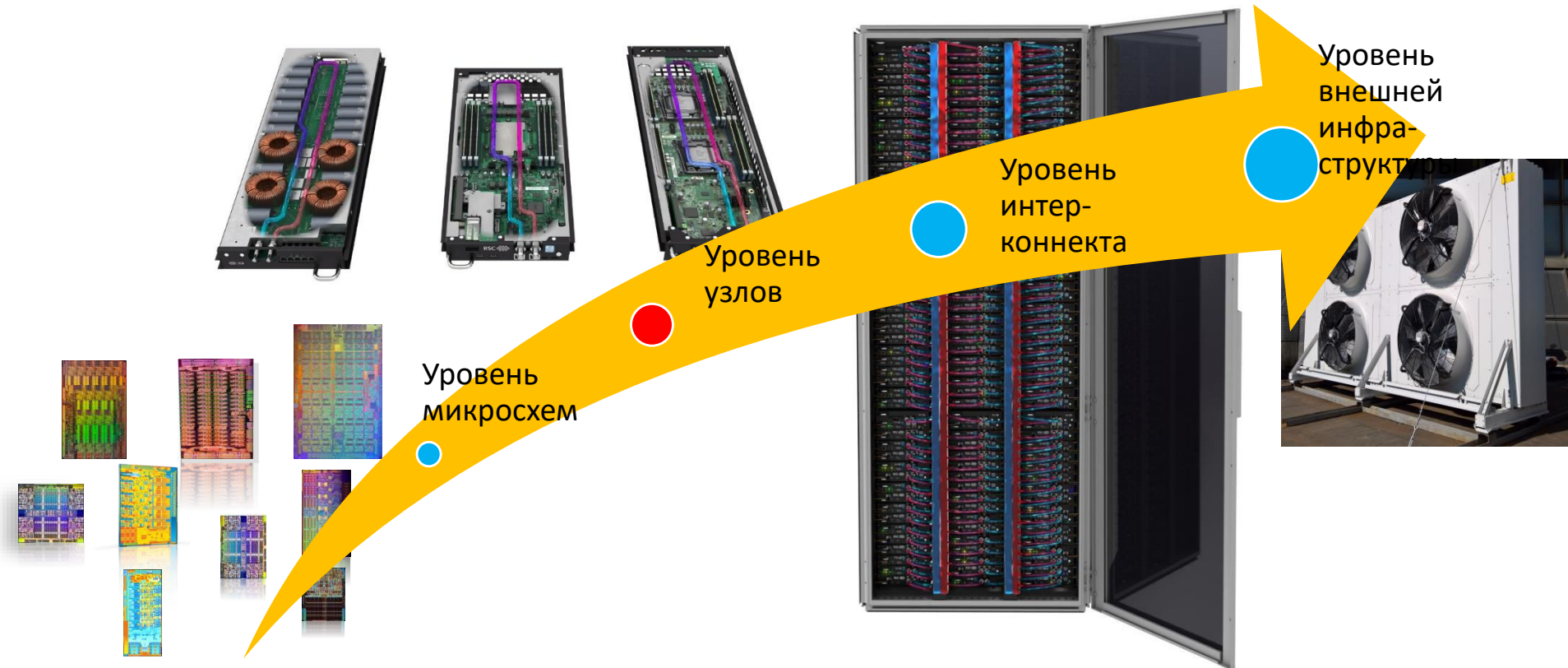
FDR Infiniband / Gigabit Ethernet / Gigabit Ethernet



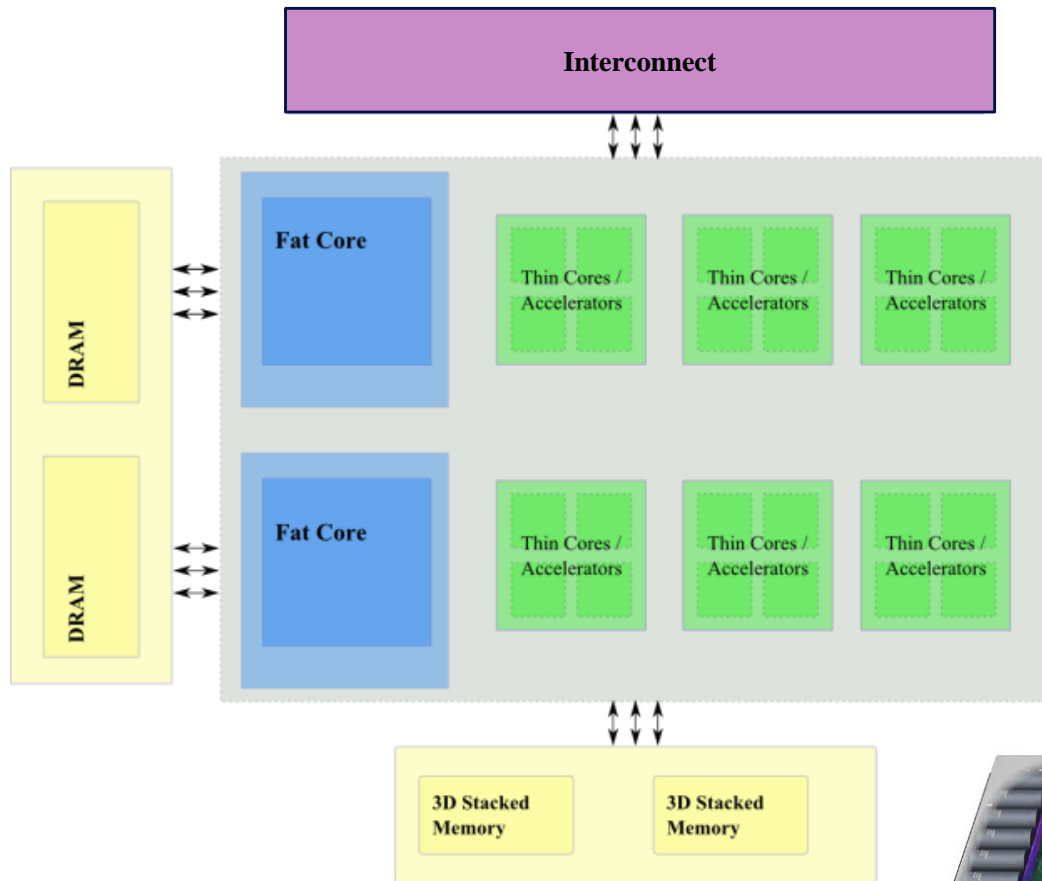
Решение актуальных проблем в:

- механике,
- гидро- и аэродинамике,
- физике твердого тела и плазмы,
- материаловедении,
- электронике,
- вычислительной и квантовой химии,
- биофизике и биотехнологиях.

# Аппаратная архитектура суперкомпьютера

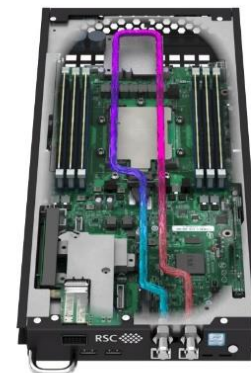
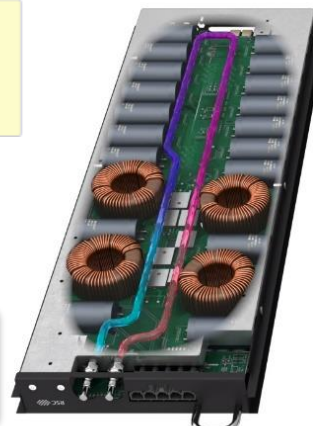


# Современные суперкомпьютерные узлы



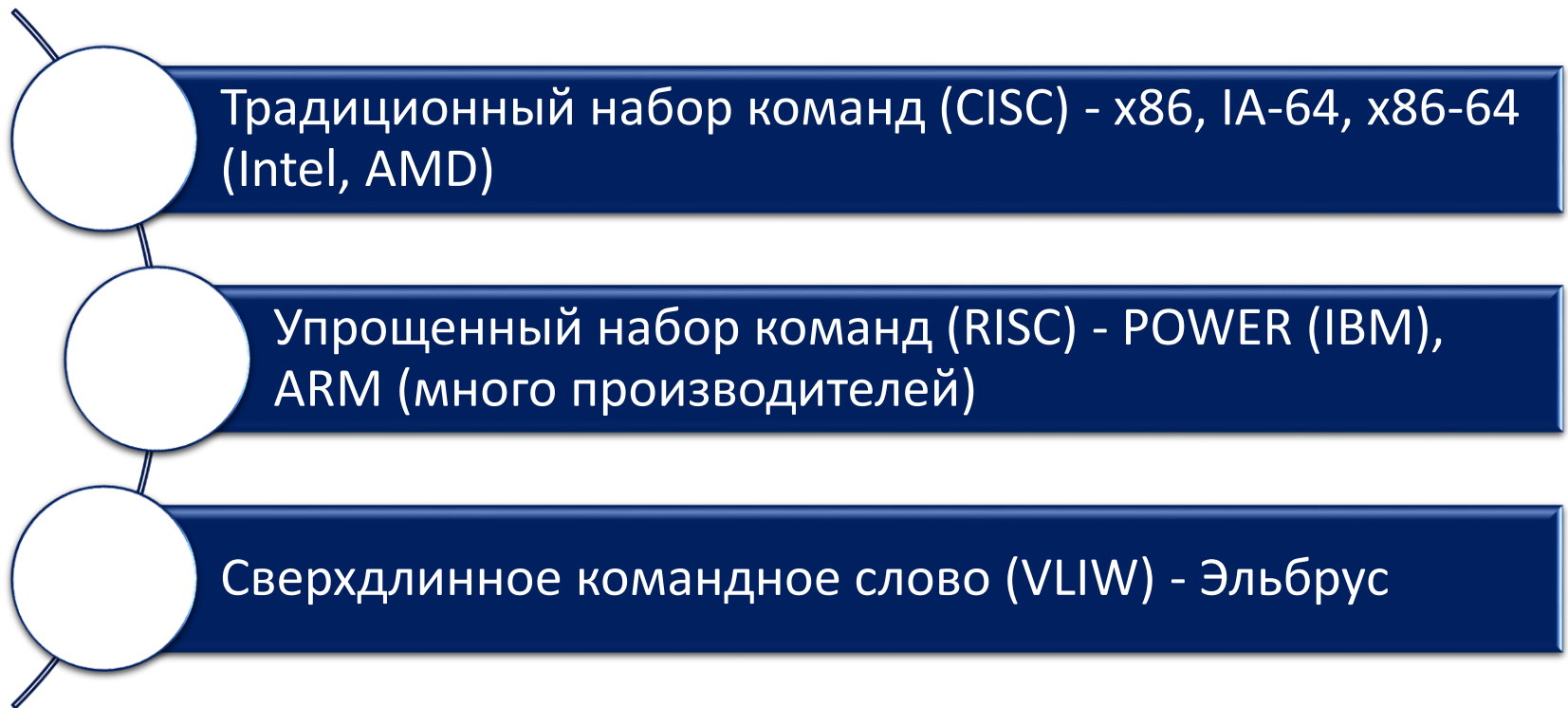
Вычислительная составляющая узла

Инфраструктурная составляющая узла



# Вычислительные универсальные ядра

*(типичная роль – хост, базовые, «классические вычислительные», «толстые»)*



# Специализированные ядра

*(типичная роль – ускоритель, сопроцессор, «тонкие»)*

- Гомогенные, на базе упрощенных универсальных – Xeon Phi (Intel)
- Перепрограммируемые (FPGA)
- Вычислительные графические ускорители (GPGPU) – Tesla (Nvidia)
- Тензорные ускорители (матричное умножение и свертка) – TPU (Google)
- Нейроморфные (самообучающиеся) - Loihi (Intel)
- Ускорители алгоритмов работы машинного зрения - Movidius (Intel)
- Узкоспециализированные (например, оптическое быстрое преобразование Фурье)
- Квантовые ускорители (криптография, молекулярное моделирование, погода)



# Память

## (оперативная и энергонезависимая)

DDR-SDRAM (Double Data Rate Synchronous Dynamic Random Access Memory) — синхронная динамическая память с произвольным доступом и удвоенной скоростью передачи данных

- Самая распространённая технология, поддерживается большинством процессоров. В настоящий момент лидирующий стандарт: DDR4

3D Stacked Memory - технология трехмерного размещения памяти, которая позволяет интегрировать ОЗУ и логические блоки микропроцессора, существенно увеличивая пропускную способность

- Следует отметить, что плотность организации 3D Stacked Memory требует специального подхода к охлаждению – требуется жидкостное охлаждение

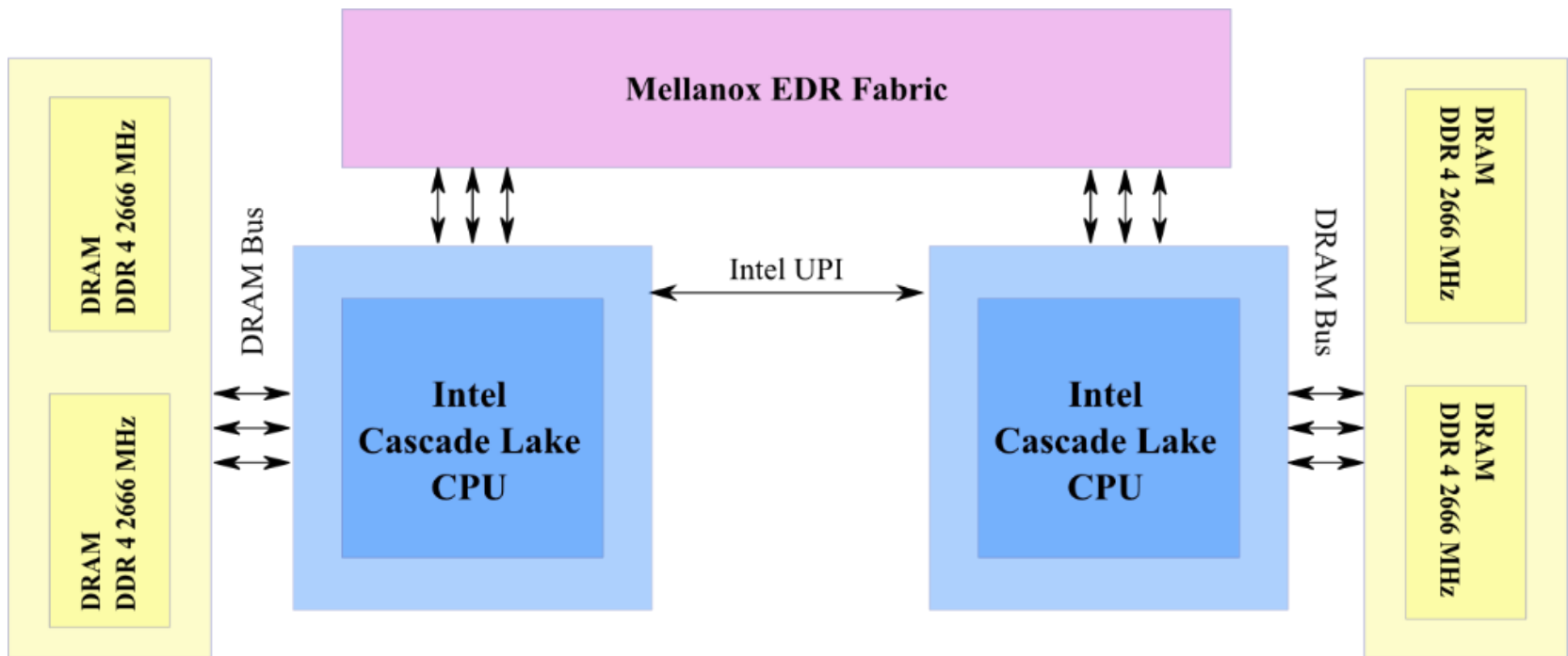
SDM (Software Defined Memory) - программно определяемая память. Программно-аппаратные решения позволяют организовать дополнительный уровень ОЗУ на базе энергонезависимой памяти

- Достоинства - большой объем, недостатки - небольшая пропускная способность по сравнению с DDR. Примером может являться технология IMDT на базе 3DXPoint NVMe накопителей

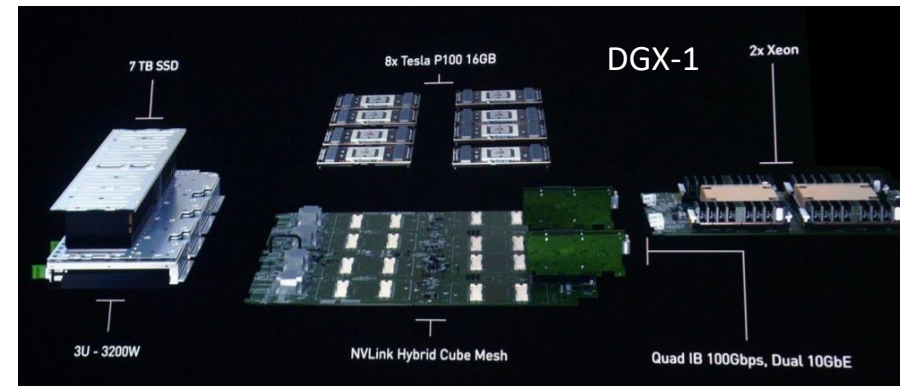
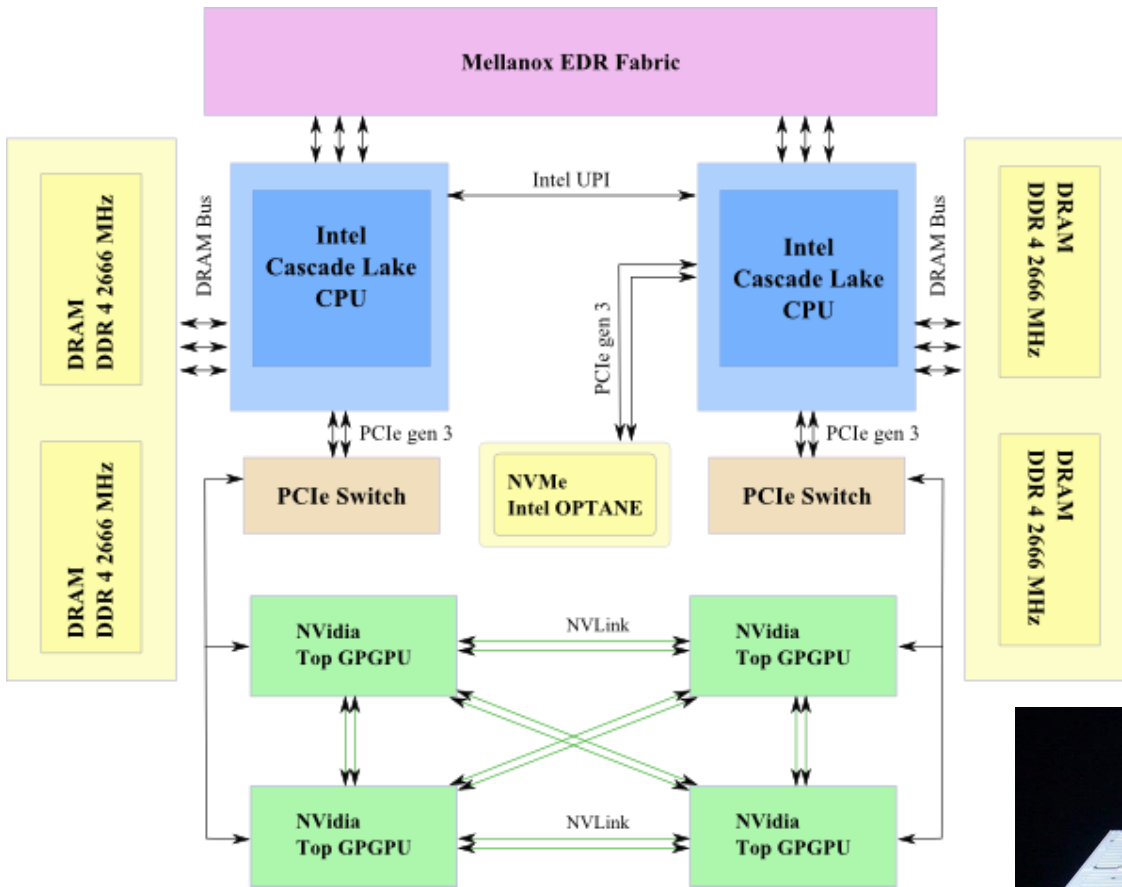
Создание промежуточного буфера (Burst Buffer) для работы с данными подразумевает наличие твердотельных накопителей (в том числе возможно использование устройств SSD) для увеличения производительности ввода-вывода

Полная интеграция уровня хранения с уровнем обработки (гиперконвергентность) - это решение, которое используется в больших, гипер-масштабируемых ЦОД. Многие современные HPC платформы видят будущее именно в создании гиперконвергентных решений

# Пример гомогенного узла

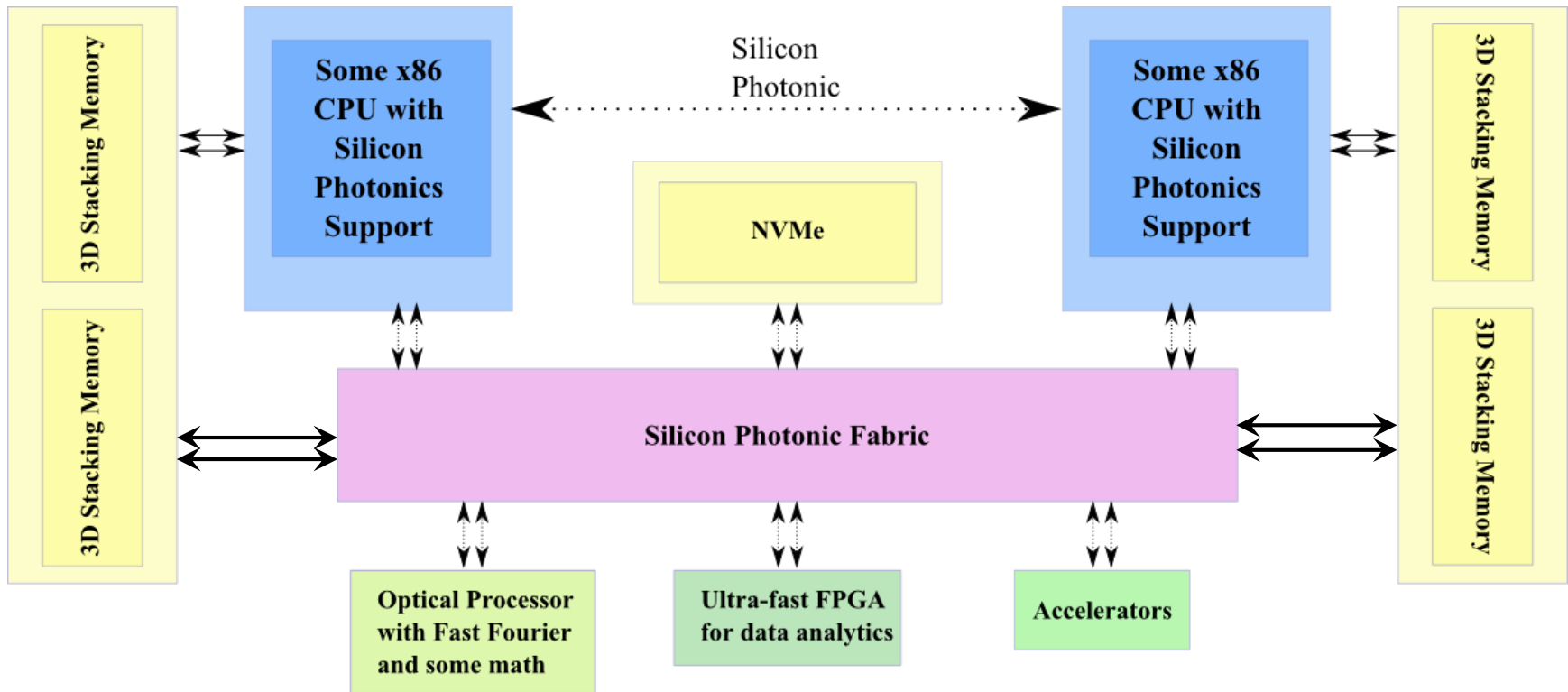


# Пример узла CISC+GPGPU





# Узел на основе новых технологий (через 5-7 лет)

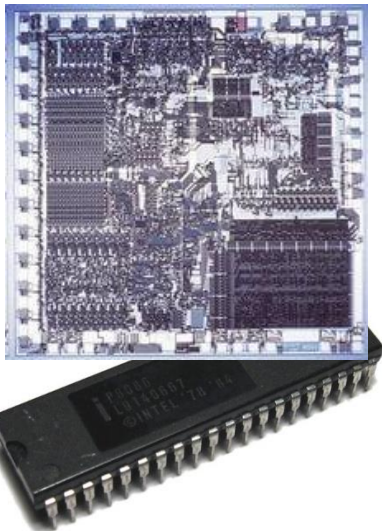


# Что мы ждём от суперкомпьютера?

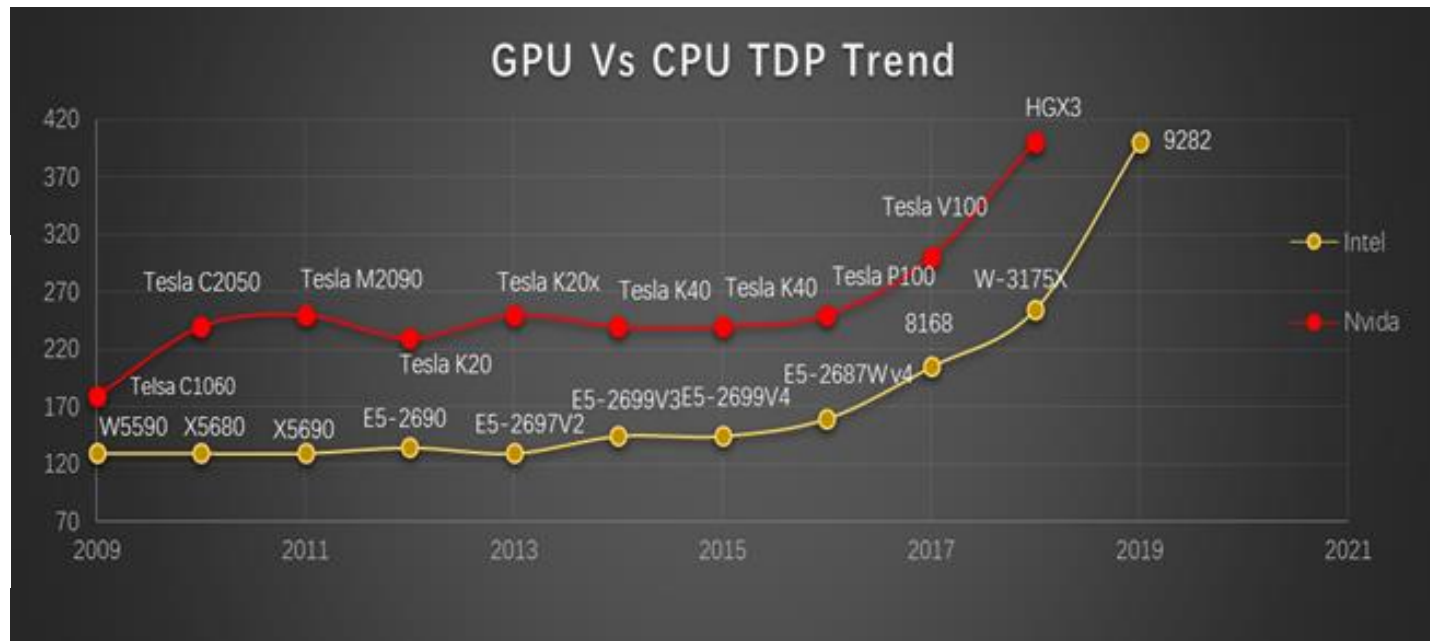


# Об экономии электроэнергии

PUE (Power Usage Effectiveness) определяется как **отношение** общей потребленной суперкомпьютерным центром энергии к энергии, которая пошла непосредственно на нужды суперкомпьютера (например, работу процессора, памяти, ...)



1978г. 8 Вт



## Основная технологии снижения PUE:

- **100% жидкостное охлаждение**
- **Free cooling** – охлаждение за счет холодного воздуха без использования фреоновых машин
- **Горячая вода (hot water cooling)** – охлаждение при температуре хладоносителя более +35°C на подаче (рекордсмены РСК с температурой до +65° С на входе в вычислительные узлы)
- **Эффективное питание**

# Охлаждение вычислителя



Тип охлаждения	Мощность шкафа (кВт)	PUE (min)	Масштабируемость	Free cooling	Горячая вода	Итог
Воздушное	6-8	1.5	Плохая Требуется места	Нет	Нет	
Воздушное (улучшенное)	30	1.3	Нормальная	Ограниченно	Нет	
Комбинированное	50	1.15	Хорошая	Да	Нет	
Иммерсионное	50*	1.06	Плохая Горизонтальная	Да	Да	
100% жидкостное (кулплейт)	>100	1.06	Хорошая	Да	Да	



# От задачи ... к выбору узлов



## Классические мощные вычислительные задачи

Гомогенные узлы с универсальными вычислителями.

При наличии ограничений по энергопотреблению - узлы на базе универсальных процессоров RISC или CISC и дополнительных ускорителей GPGPU.

## Задачи из специфических предметных областей

Узлы на базе универсальных процессоров RISC или CISC и дополнительных ускорителей FPGA, реализующих алгоритмы для данных задач.

## Обработка больших данных

Гиперконвергентные узлы. Высокопроизводительную систему хранения данных, построенную с использованием специальных технологий организации промежуточных буферов между системой хранения и вычислительными узлами.

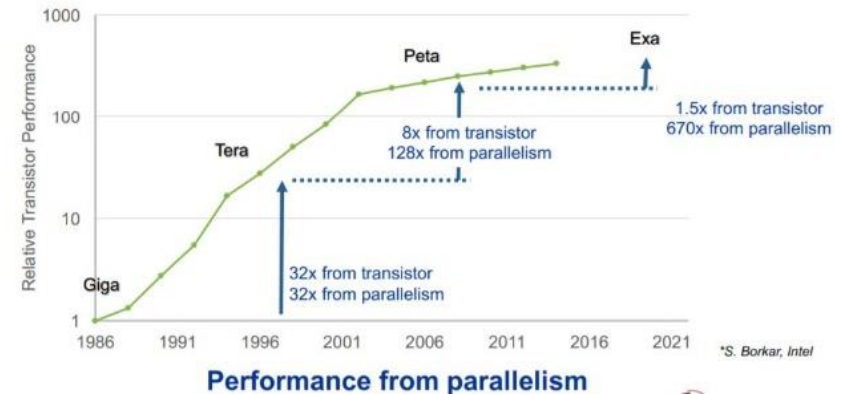
Искусственный интеллект (глубокое обучение)

Гиперконвергентные узлы на базе современных универсальных процессоров CISC с VNNI. CISC или RISC и дополнительных ускорителей GPGPU.

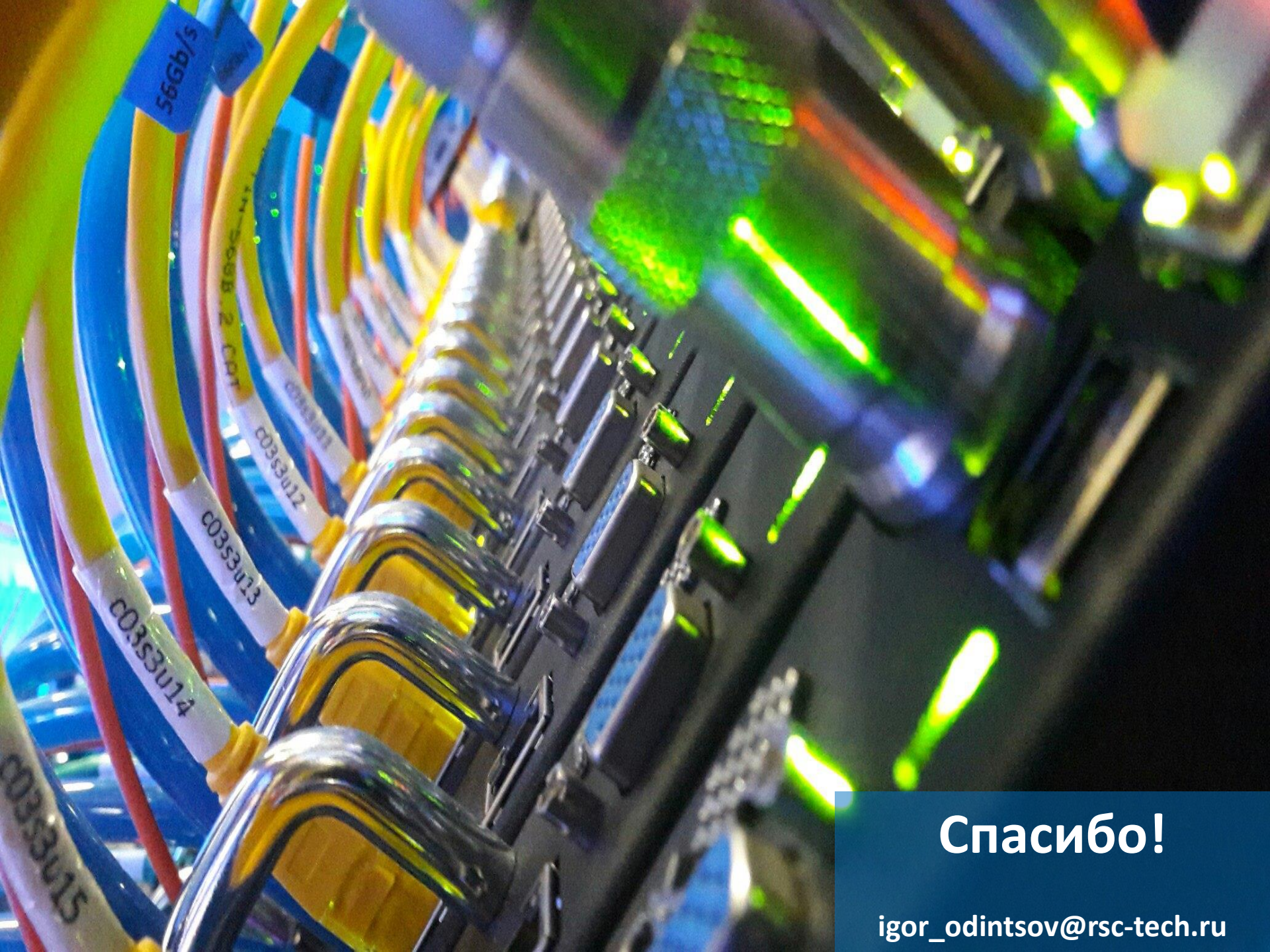
Охлаждение вычислителя: 100% жидкостное (кулплейт)

# Вызовы для программистов

Популярные прикладные пакеты для HPC (с открытым программным кодом)



1. Сложность создания массового параллельного программного обеспечения на современных языках программирования фактически поднимает вопрос о необходимости новой парадигмы программирования.
2. Множество различных высокопроизводительных аппаратных архитектур поднимает задачу не просто переносимости программного обеспечения, а максимально эффективной переносимости и автоматизированного отображения на архитектуру.



**Спасибо!**

[igor\\_odintsov@rsc-tech.ru](mailto:igor_odintsov@rsc-tech.ru)