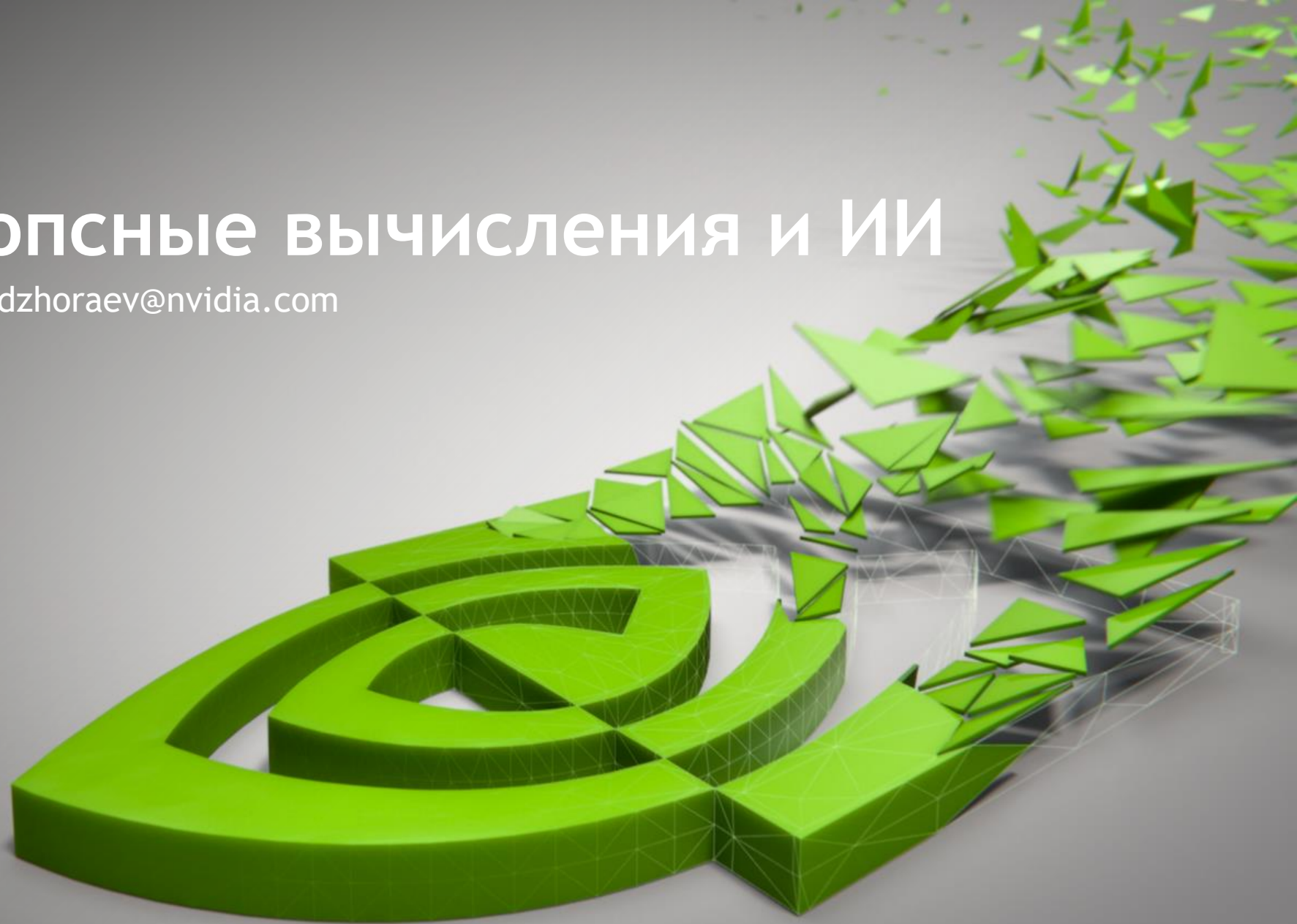


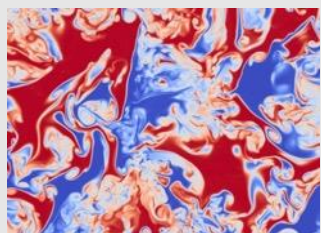
Эксафлопсные вычисления и ИИ

Антон Джораев, adzhoraev@nvidia.com

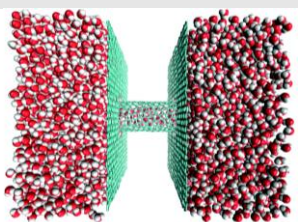


AI SUPERCOMPUTING: НОВЫЙ ЭТАП РАЗВИТИЯ НРС

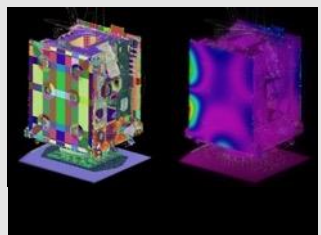
Расширение возможностей НРС с помощью искусственного интеллекта



Turbulent Flow



Molecular Dynamics



Structural Analysis

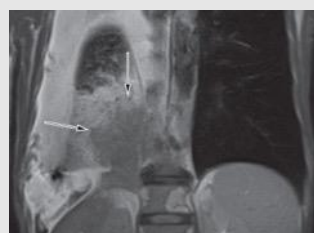


N-body Simulation

COMPUTATIONAL SCIENCE



“What’s happening?”



“Is there cancer?”

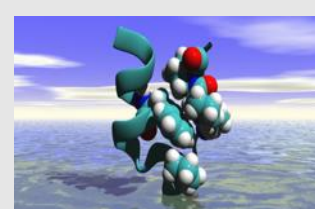


“Next move?”



“What does she mean?”

DATA SCIENCE



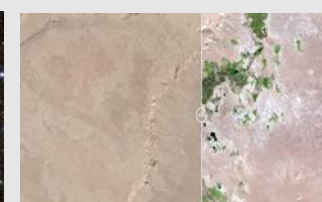
Drug Discovery



Clean Energy



Understanding Universe

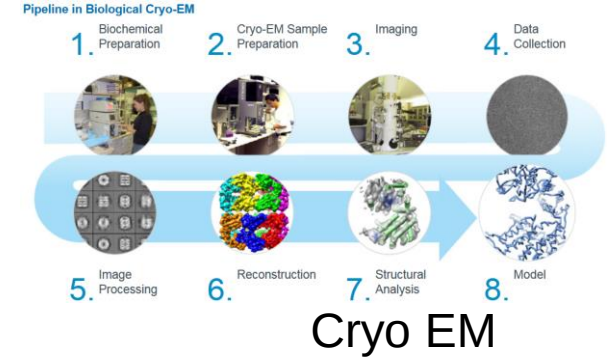
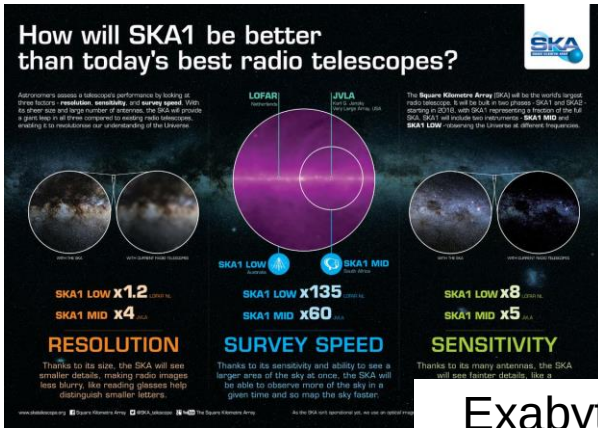


Monitoring Climate Change

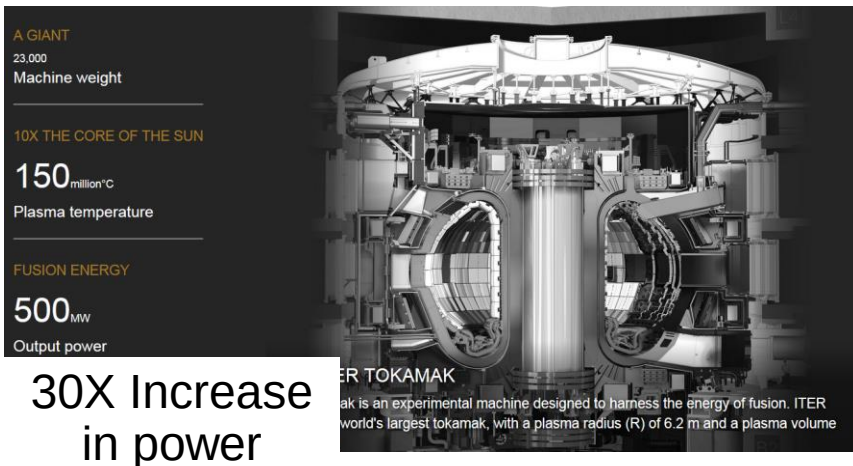
COMPUTATIONAL & DATA SCIENCE

Большие эксперименты = большие данные

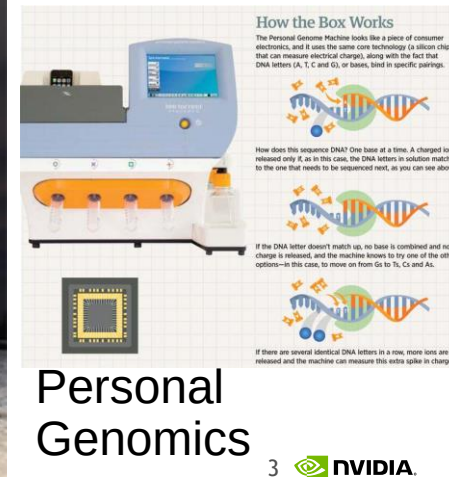
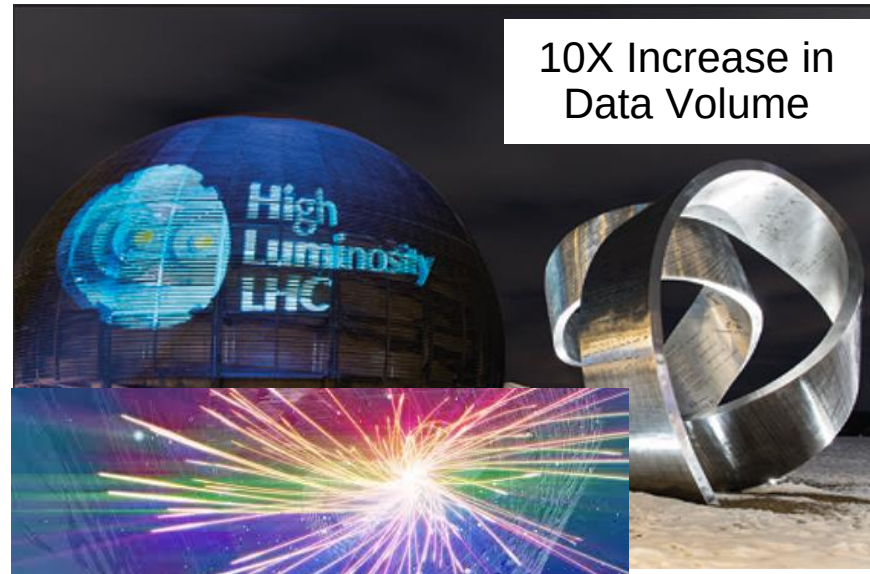
Обработка результатов экспериментов требует применения ИИ



Exabyte/Day



30X Increase in power



Примеры совмещения HPC и ИИ

Deep learning дополняет и расширяет возможности HPC



**Искусственный интеллект - мощное
дополнение к традиционным НРС-методам.**

NVIDIA TESLA - ЛИДЕР ЭФФЕКТИВНОСТИ

- №1 в Green500 - суперкомпьютер TSUBAME 3.0 на базе NVIDIA Tesla P100
- 9 из первых 10 суперкомпьютеров в Green500 - построены на NVIDIA Tesla P100
- Энергоэффективность №1 почти в **три раза выше** чем у первой машины на Xeon Phi (#18 в Green500) - 14.110 против 5.811 ГФлопс/Вт
- Архитектура Volta еще на 50% энергоэффективнее чем Pascal

NVIDIA TESLA V100

Гигантский прорыв для искусственного интеллекта и HPC

7.5 FP64 TFLOPS | 15 FP32 TFLOPS

NEW 120 Tensor TFLOPS

16GB HBM2 @ 900 GB/s

300 GB/s NVLink



NVIDIA TESLA V100

Новые Tensor Core ядра для ИИ

Новые инструкции и форматы данных CUDA
TensorOp для работы с матрицами 4x4

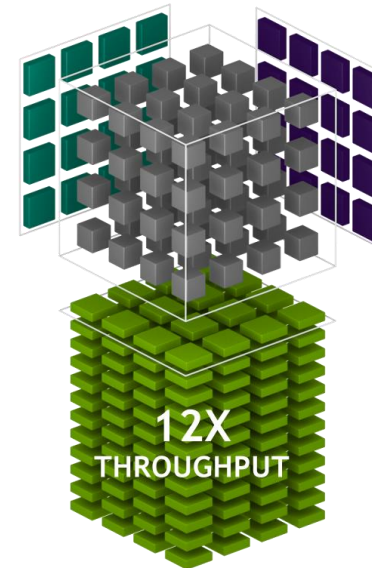
$$D[\text{FP32}] = A[\text{FP16}] * B[\text{FP16}] + C[\text{FP32}]$$

Оптимизированы для обучения и применения
глубоких нейросетей

Pascal



VOLTA Volta CORES



Activation Inputs



Weights Inputs



Output Results

NVIDIA DGX-1 НА БАЗЕ TESLA V100

Мощнейший инструмент для исследования в области ИИ

960 Tensor TFLOPS | 8x Tesla V100 | NVLink Hybrid Cube

Заменяет 400 традиционных серверов на задачах ИИ



NVIDIA DGX-1 НА БАЗЕ TESLA V100

Мощнейший инструмент для исследования в области ИИ

960 Tensor TFLOPS | 8x Tesla V100 | NVLink Hybrid Cube

Заменяет 400 традиционных серверов на задачах ИИ

60 ТФлопс FP64 - достаточно для Топ50

3U в стойке, 3.2 КВт

Интерконнект Dual 10 GbE



NVIDIA DGX-1 НА БАЗЕ TESLA V100

Мощнейший инструмент для исследования в области ИИ

960 Tensor TFLOPS | 8x Tesla V100 | NVLink Hybrid Cube

Заменяет 400 традиционных серверов на задачах ИИ

60 ТФлопс FP64 - достаточно для Топ50

3U в стойке, 3.2 кВт

Интерконнект Dual 10 GbE

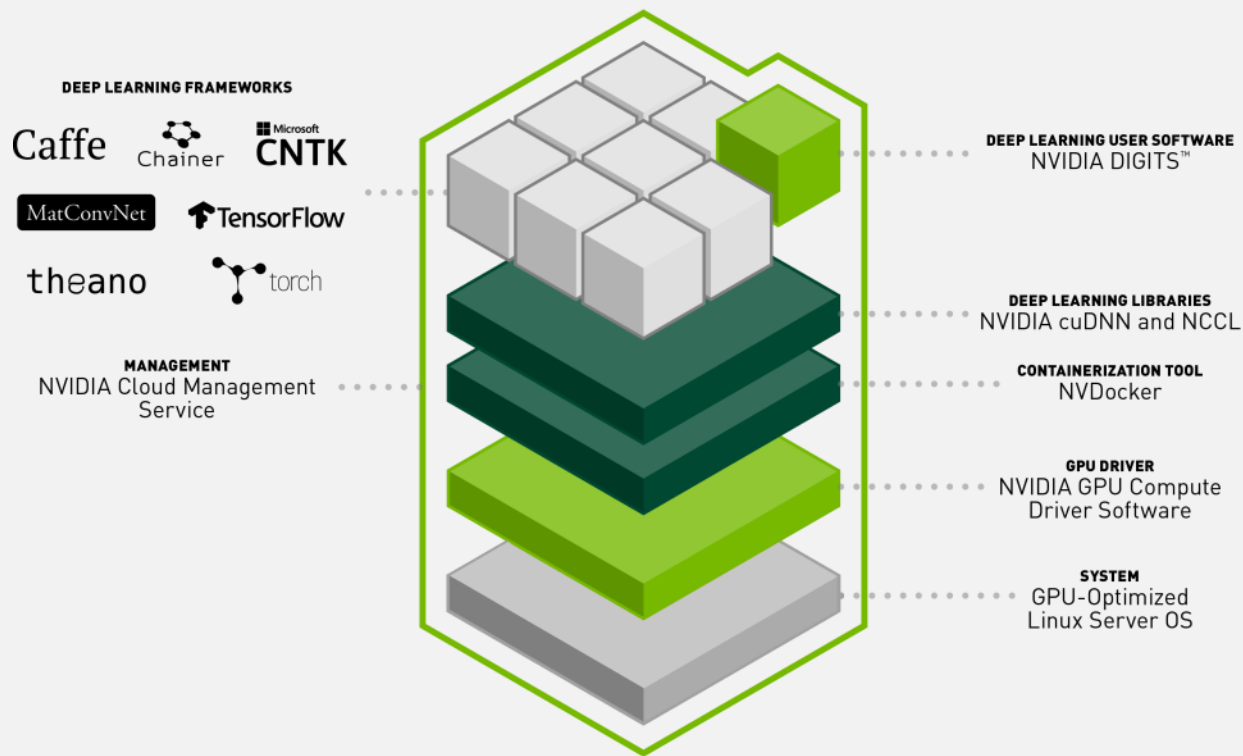
Срок поставки меньше месяца

Пуск в эксплуатацию за 1 день



ПРОГРАММНЫЙ СТЕК DGX

Полностью интегрированная платформа для глубокого обучения



Plug-and-play - производительность здесь и сейчас

Поддержка всех фреймворков глубокого обучения

Оптимизация всех компонентов решения для максимальной производительности

Постоянные обновления из облака

Поддержка нескольких, одновременно работающих фреймворков, благодаря контейнеризации

Командная работа

Простота эксплуатации

NVIDIA DGX STATION

Суперкомпьютер для ИИ на
вашем столе



Персональный суперкомпьютер для исследователей и Data Scientists



Бесшумная настольная рабочая станция



Моментальное начало работы с
программным стеком DGX



Простой перенос задач с рабочей станции
в ЦОД или в облако



Прорывная производительность
архитектуры Volta



DEEP
LEARNING
INSTITUTE

Образовательная инициатива Deep Learning Institute

www.nvidia.ru/dli

Свободный доступ для преподавателей!



- [illegible]

<https://developer.nvidia.com/teaching-kits>

Курсы читаются в:



Лабораторные работы



DEEP
LEARNING
INSTITUTE

Более 20 прикладных лабораторных работ по deep learning

Некоторые темы:

- . Object detection
- . Image classification
- . Image creation using Generative Adversarial Networks
- . Image segmentation
- . Signal processing
- . Medical image analysis
- . Medical image segmentation
- . Deep learning for genomics using DragoNN with Keras and Theano
- . Trading strategy for finance using LSTMs

Приглашение на лекцию



DEEP
LEARNING
INSTITUTE

Вводная лекция по глубокому обучению и нейросетям

14:10 - 15:50, понедельник
Зал «Охотный ряд»

Сегодня каждый вычислительный центр должен иметь инструменты для ИИ:

1. вычислительную систему DGX-1
2. курсы по глубоким нейросетям.

Эксафлопсные вычисления и ИИ

Антон Джораев, adzhoraev@nvidia.com

