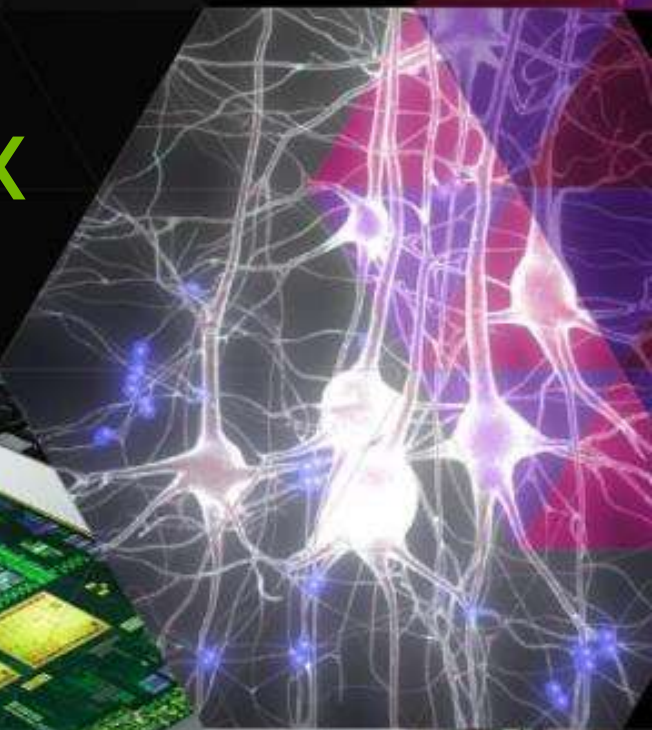




# GPU ДЛЯ ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ ВЫЧИСЛЕНИЙ И ВИЗУАЛИЗАЦИИ

Антон Джораев





GAMING



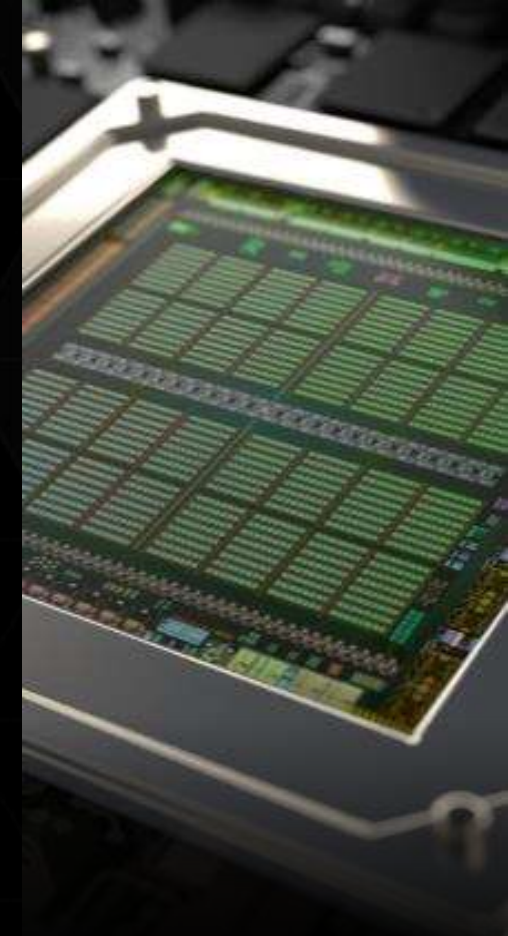
AUTO



ENTERPRISE



HPC & CLOUD



OEM & IP

Лидер в области визуальных вычислений



# Три пути ускорения приложений

Приложения

Библиотеки

Простейший  
способ

Директивы  
OpenACC

Минимум усилий,  
переносимость кода

Языки

Максимальная гибкость  
(unified memory,  
for\_each, lambda)

# Библиотеки с GPU ускорением



Автоматическое ускорение  
до 10 раз

Автоматическое  
масштабирование с multi-GPU  
библиотеками

**75%** разработчиков используют  
GPU библиотеки в своих  
приложениях

# Три пути ускорения приложений

Приложения

Библиотеки

Простейший  
способ

Директивы  
OpenACC

Минимум усилий,  
переносимость кода

Языки

Максимальная гибкость  
(unified memory,  
for\_each, lambda)

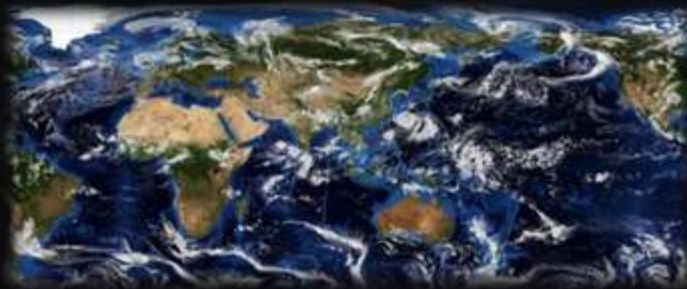
# OpenACC

просто | эффективно | переносимо

Открывает новую эру научных  
открытий в HPC

```
main()
{
  <serial code>
  #pragma acc kernels
  //automatically runs on GPU
  {
    <parallel code>
  }
}
```

RIKEN Japan  
NICAM - моделирование климата



7-8x ускорение  
5% кода модифицировано

University of Illinois  
PowerGrid - реконструкция МРТ



70x ускорение  
2 дня усилий

8000+

Разработчиков  
используют  
OpenACC

# Представляем NVIDIA OpenACC Toolkit

набор инструментов для простого и эффективного ускорения ваших приложений



## PGI Compiler

Бесплатный компилятор OpenACC для академических пользователей



## NVProf Profiler

Легкий поиск мест в коде для эффективного применения директив



## Code Samples

Обучающие примеры на базе реальных алгоритмов



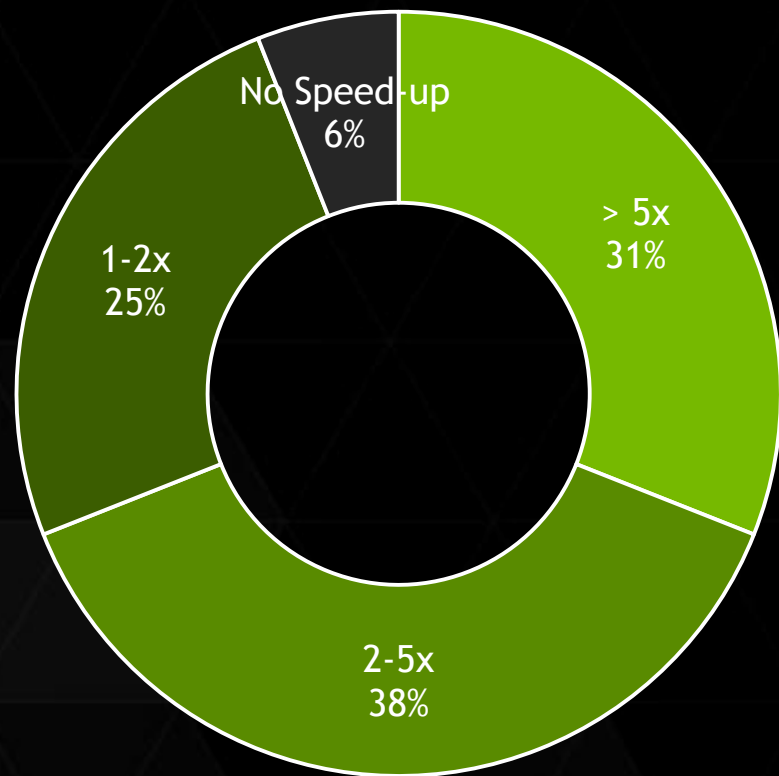
## Documentation

Руководства, документация, вебинары, форумы

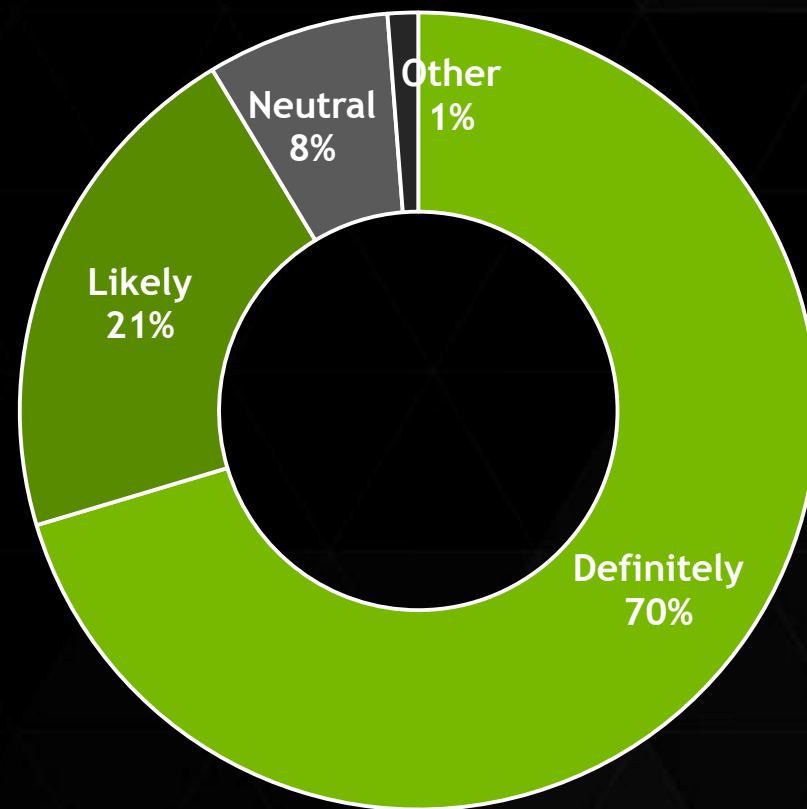
Страница для скачивания: <http://www.nvidia.com/openacc>

# Пользователи OpenACC довольны результатами

**94%** достигли увеличения производительности



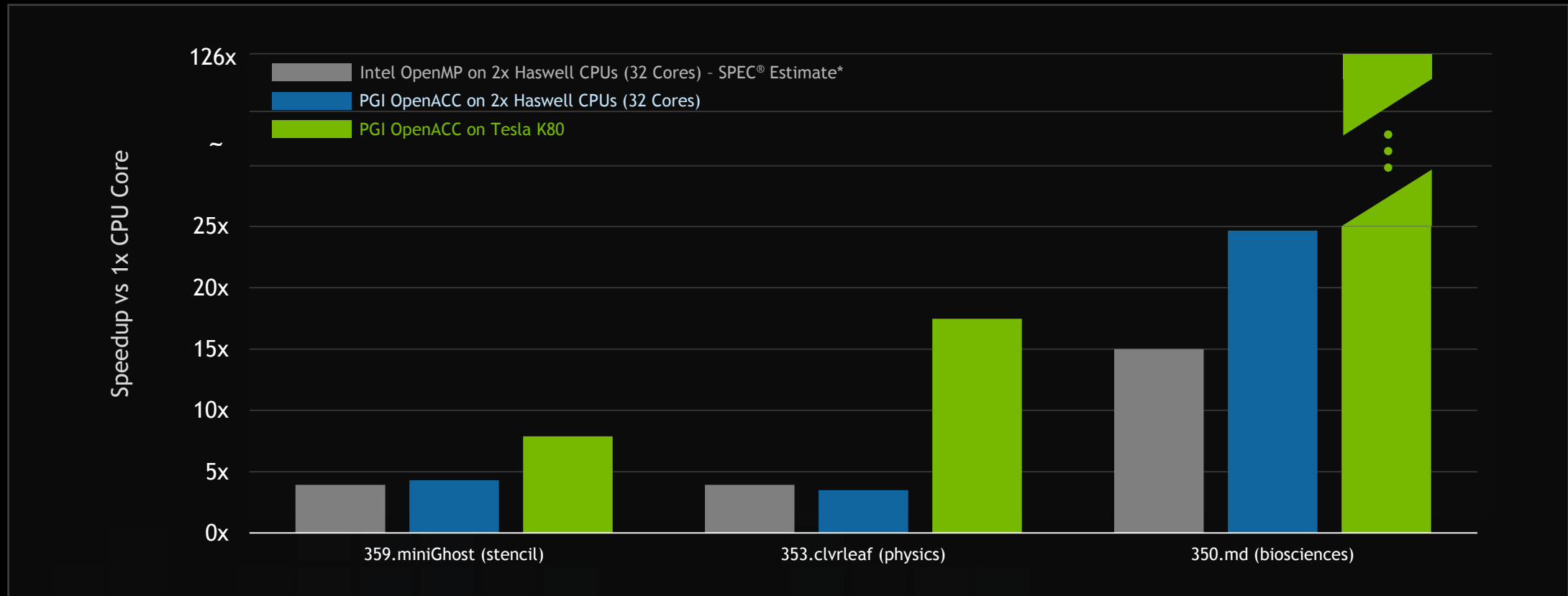
**91%** готовы рекомендовать OpenACC коллегам





# OpenACC – ускорение на любой платформе

Единый код. Исполняется везде. Эффективно.



# Три пути ускорения приложений

Приложения

Библиотеки

Простейший  
способ

Директивы  
OpenACC

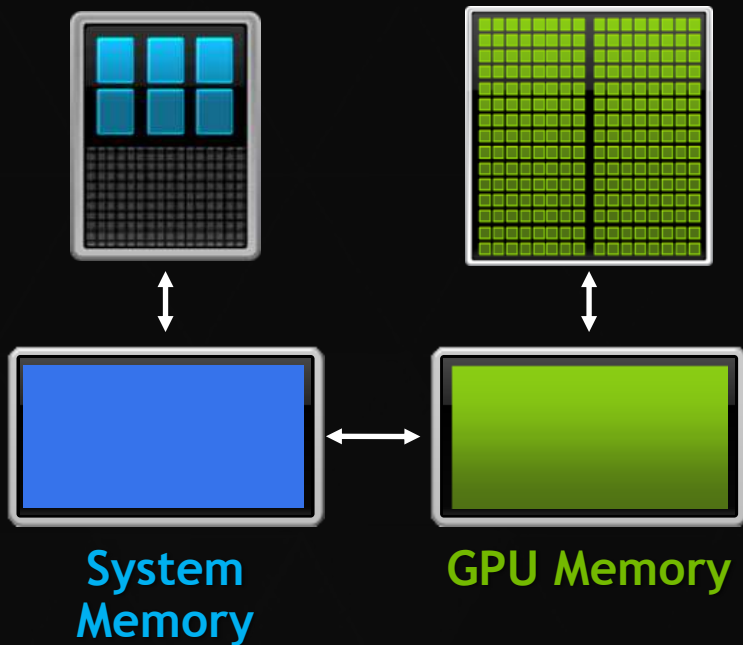
Минимум усилий,  
переносимость кода

Языки

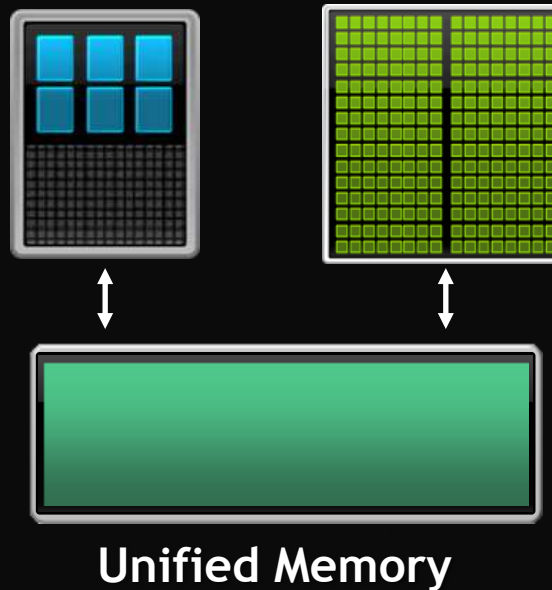
Максимальная гибкость  
(unified memory,  
for\_each, lambda)

# Unified Memory: проще и быстрее с NVLink

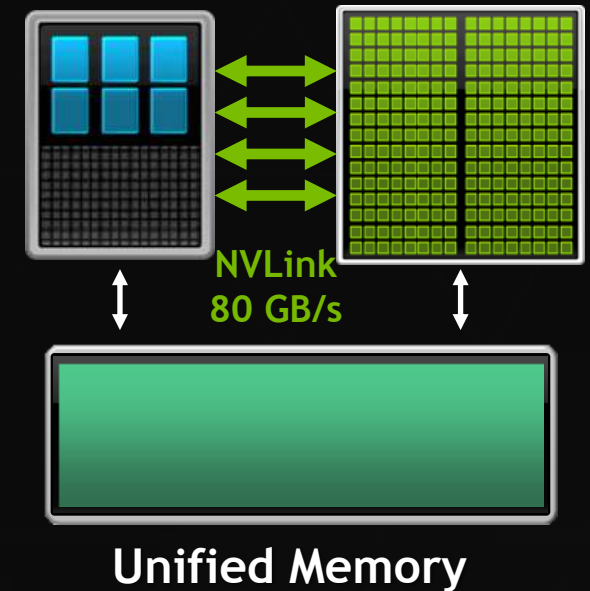
Традиционный взгляд на систему



Вычислительная система с Unified Memory



Архитектура Pascal и интерфейс NVLink



Обмен данными со скоростью памяти хоста, а не шины PCI-E.

# PASCAL

## NVLink

В 5-12 раз быстрее PCIe 3.0

## 3D память

В 2-4 раза выше пропускная  
способность и объем памяти





# TESLA K80

Самый быстрый в мире ускоритель для научных расчетов и аналитики

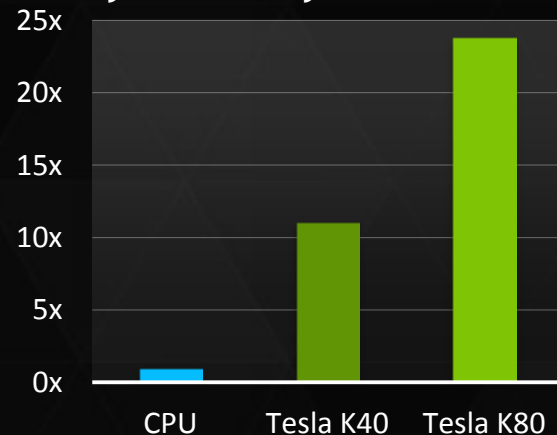


Ускоритель с двумя GPU для максимальной производительности

**2x быстрее**

2.9 TF | 4992 Cores | 480 GB/s

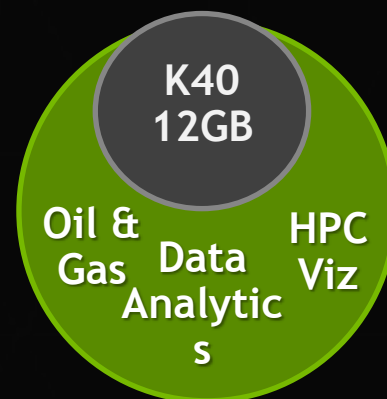
Глубокое обучение: Caffe



**Вдвое больше памяти**

Для приложений Big Data

**24GB**



**Максимум**

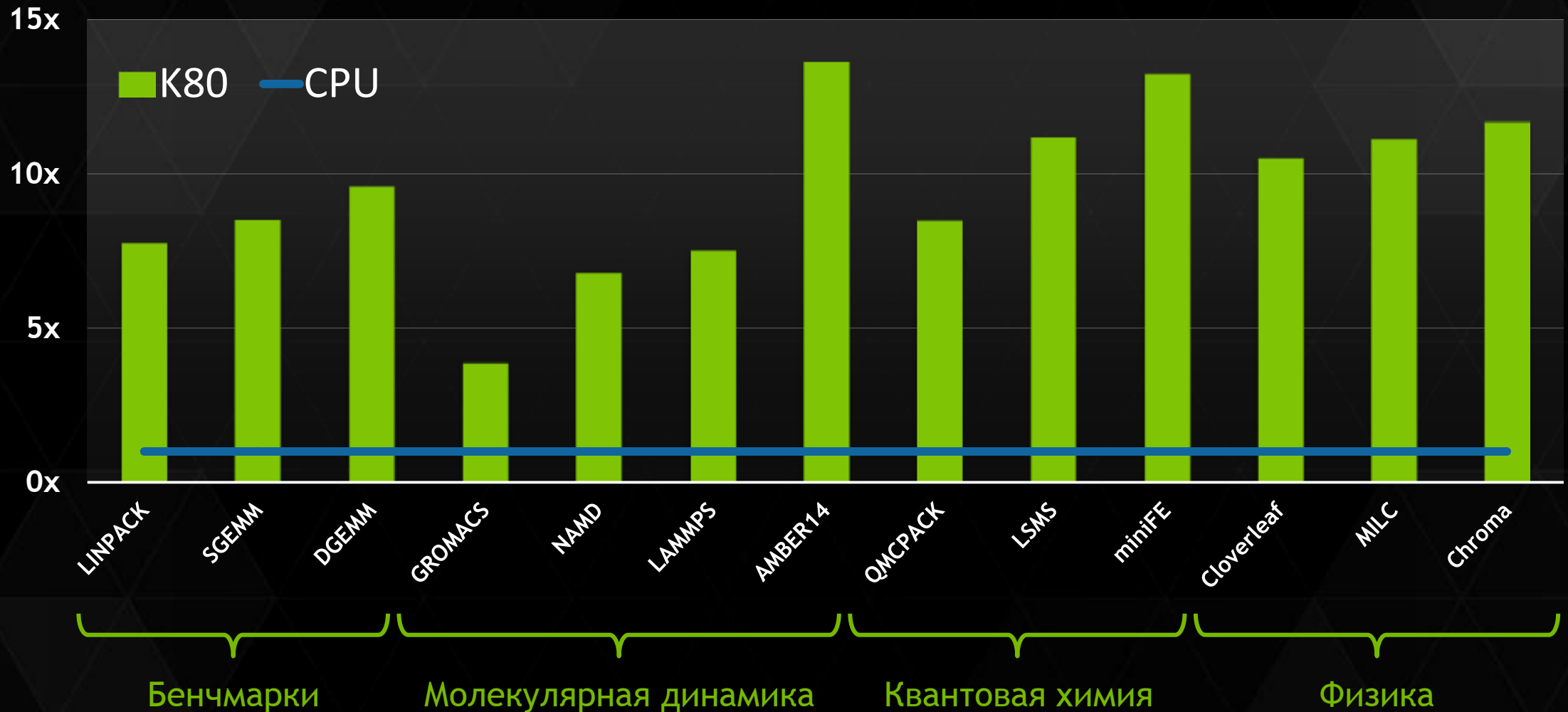
**производительности**

Динамическое повышение производительности для каждого приложения



**GPU Boost**

# Ускорение реальных приложений до 10 раз



CPU: 12 cores, E5-2697v2 @ 2.70GHz. 64GB System Memory, CentOS 6.2  
GPU: Single Tesla K80, Boost enabled



## POPULAR GPU-ACCELERATED APPLICATIONS

- 02 Research: Higher Education and Supercomputing
  - COMPUTATIONAL CHEMISTRY AND BIOLOGY
  - NUMERICAL ANALYTICS
  - PHYSICS
  - WEATHER AND CLIMATE PREDICTION
- 06 Defense and Intelligence
- 07 Computational Finance
- 08 Manufacturing: CAD and CAE
- COMPUTER AIDED DESIGN
- COMPUTATIONAL FLUID DYNAMICS
- COMPUTATIONAL STRUCTURAL MECHANICS
- BLUETOOTH BEAM SIMULATION
- 10 Media and Entertainment
  - ANIMATION, RENDERING AND RENDERING
  - COLOR CORRECTION AND GRAFIC MANAGEMENT
  - COMPOSITING, FINISHING AND EFFECTS
  - EDITING
  - ENCODING AND SOCIAL DISTRIBUTION
  - ON-DEMAND GRAPHICS
  - ON-DEMAND RENDERING AND EFFECTS
  - EMULATOR
  - VIDEO EDITING
- 14 Oil and Gas

### Research: Higher Education and Supercomputing

#### COMPUTATIONAL CHEMISTRY AND BIOLOGY

##### Bioinformatics

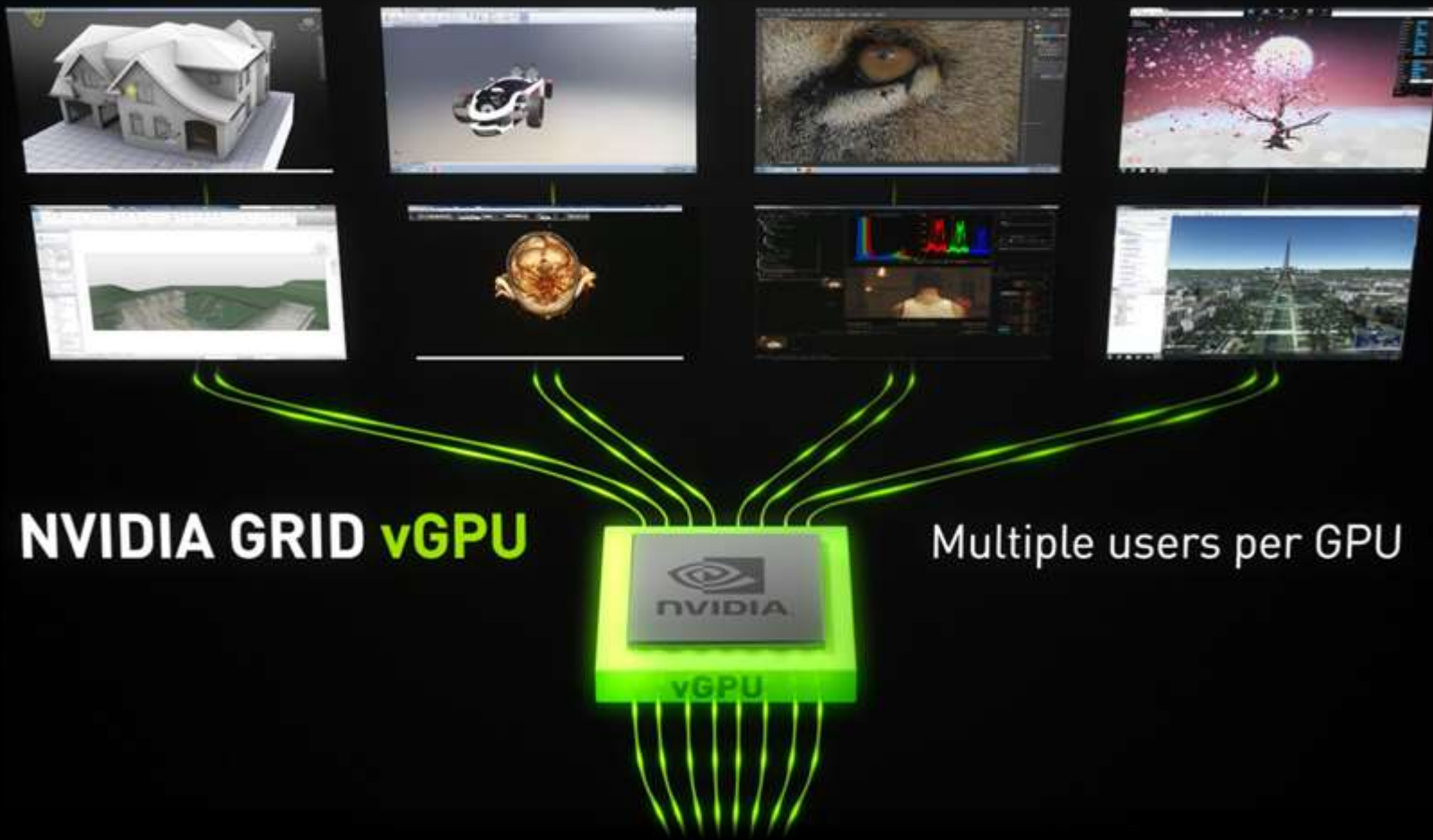
Application	Description	Company/Source	Compute Type	Performance	GPU Support	Availability
BerryCOSA	Sequence mapping software	Alignement of short sequencing reads	8-16x	7.20Tf, 20P, 81G, K20, K20K	Yes	Available now Version 0.3
CUDA-SIM+	Open source software for Smith-Waterman protein database searches on GPUs	Parallel search of Smith-Waterman Database	10-50x	1.20Tf, 20P, 81G, K20, K20K	Yes	Available now Version 2.0.8
COSMAM	Parallelized short read aligner	Parallel, accurate long read aligner - supports alignments to large genomes	10x	7.20Tf, 20P, 81G, K20, K20K	Yes	Available now Version 1.0.0E
GPU-BLAST	Local search with fast k-mers heuristic	Protein alignment according to k-mer, multi-cpu threads	3-4x	7.20Tf, 20P, 81G, K20, K20K	Single only	Available now Version 2.2.34
GPU-HMMER	Parallelized local and global search with profile Hidden Markov models	Parallel local and global search of Hidden Markov Models	40-100x	1.20Tf, 20P, 81G, K20, K20K	Yes	Available now Version 2.3.3
nvCUDA-NEMO	Ultra-fast scalable motif discovery algorithm based on MEME	Scalable motif discovery algorithm based on MEME	4-10x	7.20Tf, 20P, 81G, K20, K20K	Yes	Available now Version 3.0.1E
SeqTMC	A GPU Accelerated Sequence Analysis Toolkit	Reference assembly, blast, gene-annotation, track, de-novo assembly	40x	7.20Tf, 20P, 81G, K20, K20K	Yes	Available now
USBE	Open-source Smith-Waterman for GPU/CUDA, suffix array based repeats filter and output	Fast short read alignment	9-10x	7.20Tf, 20P, 81G, K20, K20K	Yes	Available now Version 1.1E
WdLM	Fits numerous linear models to a fixed design and response	Parallel linear regression on multiple arbitrary-shaped matrices	70x	1.20Tf, 20P, 81G, K20, K20K	Yes	Available now Version 0.1-1

##### Molecular Dynamics

Application	Description	Company/Source	Compute Type	Performance	GPU Support	Availability
Abnema	Models molecular dynamics of biomolecules for simulations of proteins, DNA and ligands	Simulation on 100k GPUs	4-7x	7.20Tf, 20P, 81G, K20, K20K	Single Only	Available now Version 1.8.0E
ACEMD	GPU simulation of molecular dynamics force fields, implicit and explicit solvers	Written for use on GPUs	143 Au/Day (GPU version only)	7.20Tf, 20P, 81G, K20, K20K	Yes	Available now
AMBER	Suite of programs for molecular dynamics simulation					
DL-PCG	Simple and fast iterative memory partitioning					
CHARMM	MD package for molecular dynamics simulation					
SKMACE	Simulation of complex systems with complex force fields					
HOOMD-blue	Particle dynamics simulation package					
LAMMPS	Classical molecular dynamics simulation package					
HOLO	Designed for simulation of polymer chains					
HOLOM	Library and application for simulation of polymer chains					

330+ приложений с ускорением на GPU

[www.nvidia.com/appscatalog](http://www.nvidia.com/appscatalog)



**NVIDIA GRID™ vGPU™** это технология аппаратной виртуализации графического процессора позволяющая множественным виртуальным машинам взаимодействовать с GPU напрямую



# Продукты для виртуализации

GRID K1



Tesla M6



GRID K2



Tesla M60





# GPU ДЛЯ ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ ВЫЧИСЛЕНИЙ И ВИЗУАЛИЗАЦИИ

Антон Джораев

[adzhoraev@nvidia.com](mailto:adzhoraev@nvidia.com)

