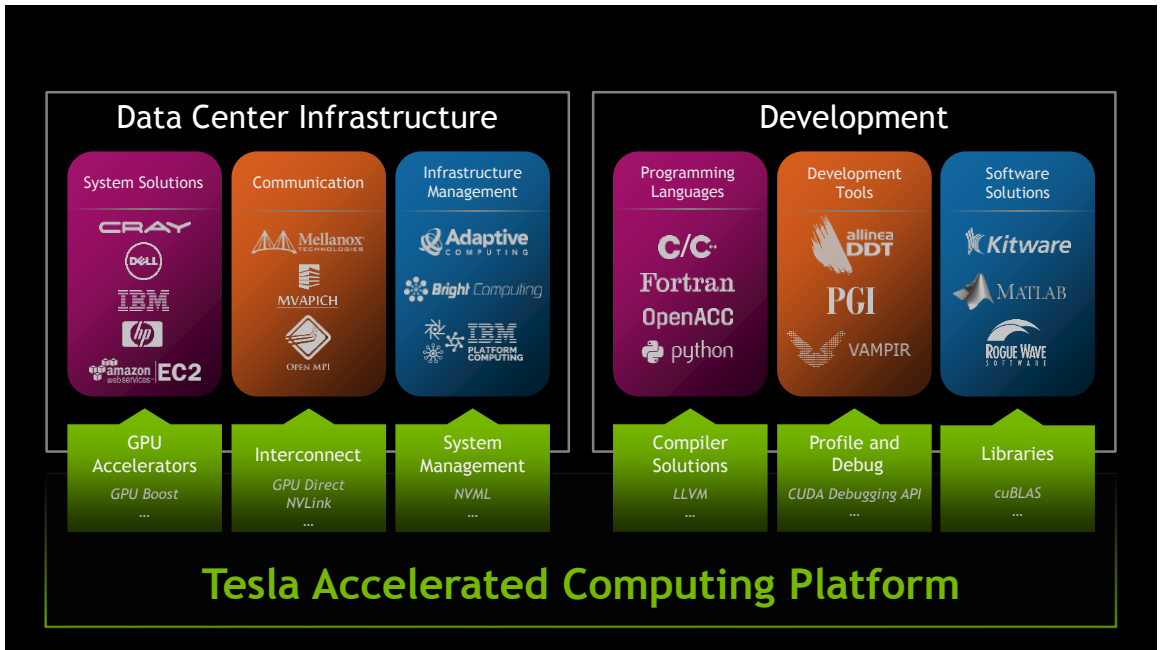
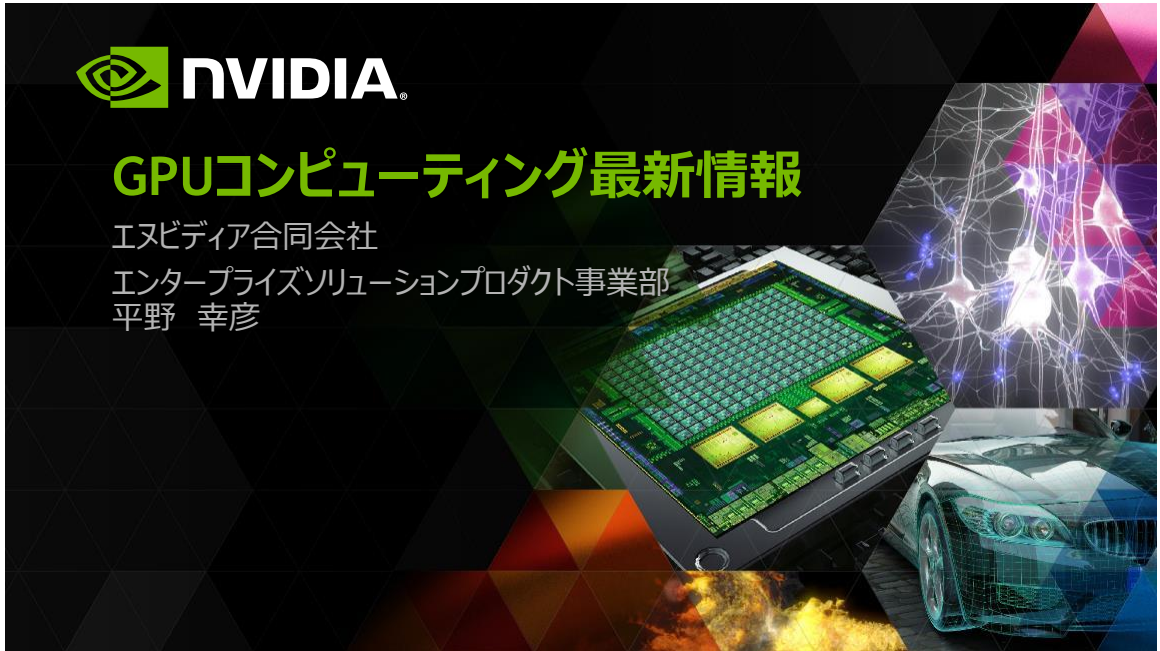




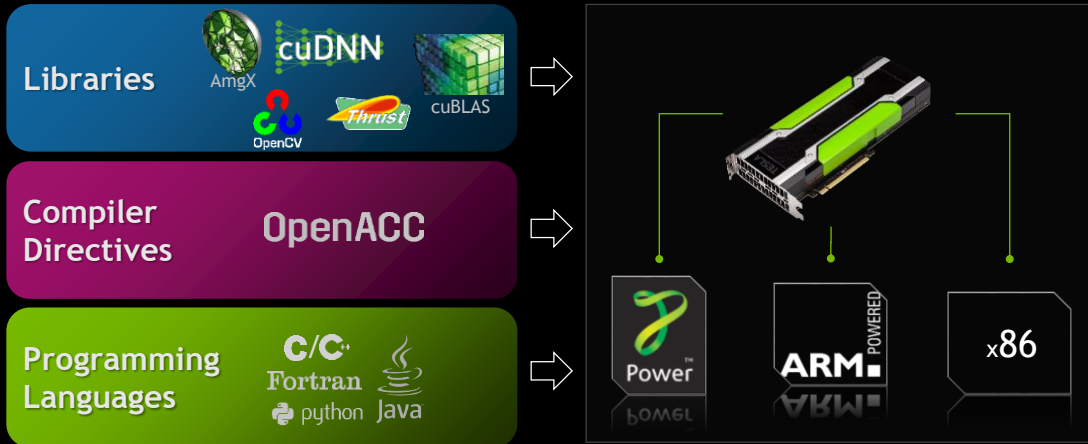
# GPUコンピューティング最新情報

エヌビディア合同会社  
 エンタープライズソリューションプロダクト事業部  
 平野 幸彦



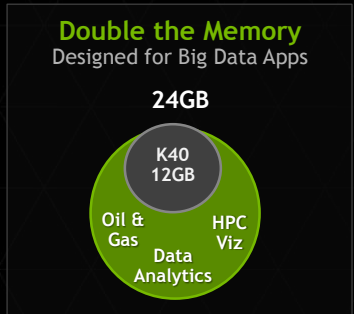
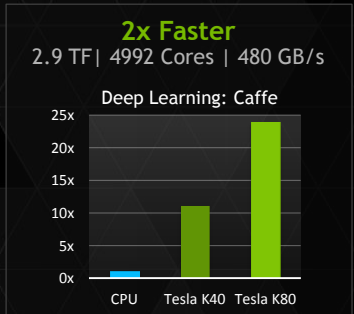
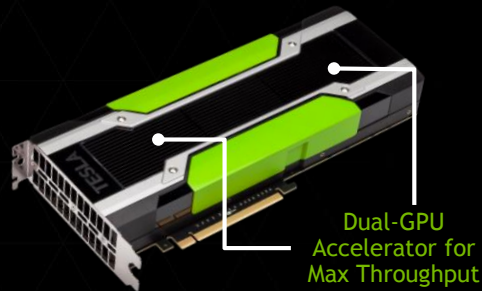
# Tesla: Platform with Open Ecosystem

## Common Programming Models Across Multiple CPUs



## TESLA K80

ビッグデータ解析と科学技術計算のための  
 世界最速のアクセラレータ



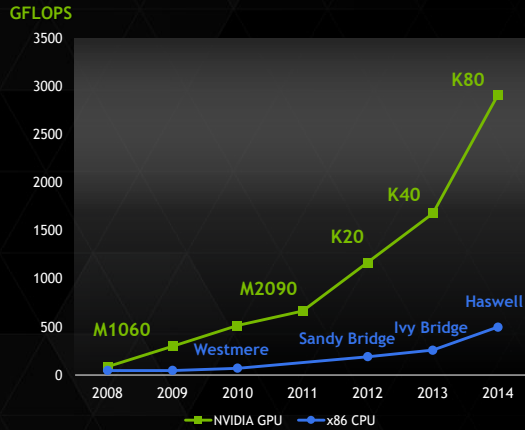
**Maximum Performance**  
 Dynamically Maximize Performance for Every Application

**GPU Boost**

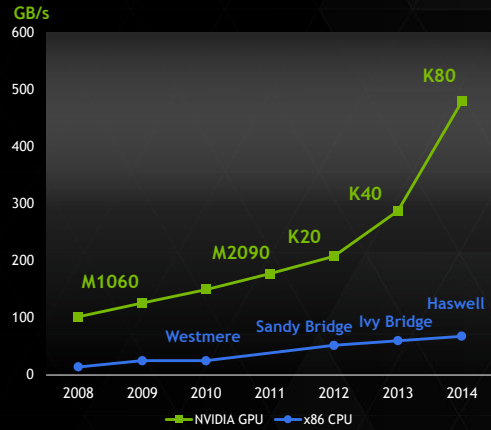
Caffe Benchmark: AlexNet training throughput based on 20 iterations, CPU: E5-2697V2 @ 2.70GHz, 64GB System Memory, CentOS 6.2

# Performance Lead Continues to Grow

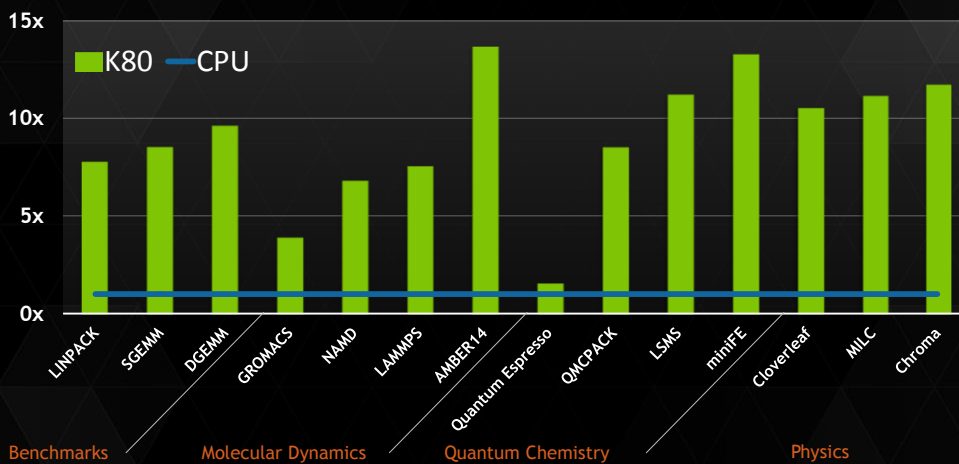
Peak Double Precision FLOPS



Peak Memory Bandwidth

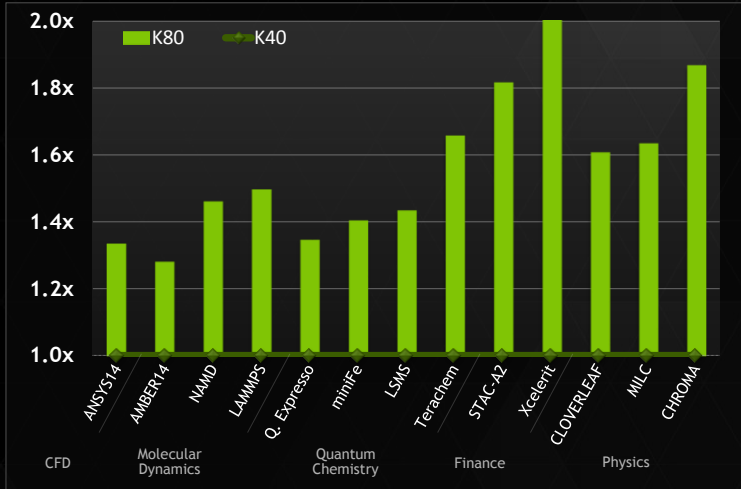
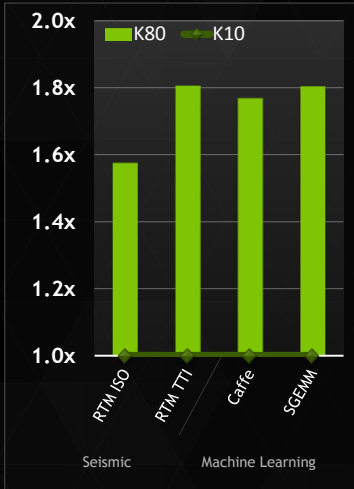


# 10x Faster than CPU on Applications



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

# Tesla K80: Up to 2x Faster than K10 and K40



Single K10 or K80, Boost enabled for K80. Actual boost level dependent on application profile.  
 RTM ISO R4 2x, RTM TTI R8 3 pass

## High GPU Density Servers Now Mainstream



**CRAY**  
THE SUPERCOMPUTER COMPANY

**Cray CS-Storm**  
8 K80s per Node



**Dell C4130**  
4 K80s per Node



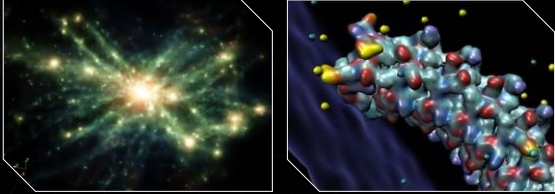
**HP SL270**  
8 K80s per Node




**Quanta S2BV**  
4 K80s per Node

# CSCS Piz Daint

## The World's Largest In-situ Viz Enabled GPU Supercomputer

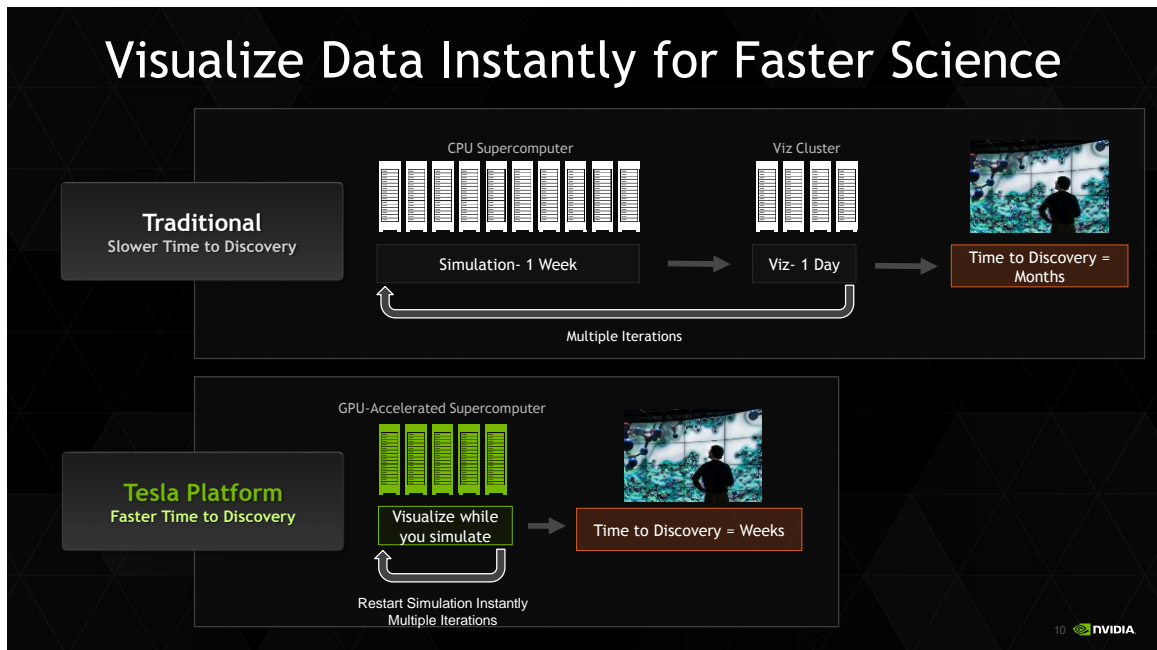


2048 GPU Nodes running Galaxy formation and Molecular Dynamics Simulation + Visualization

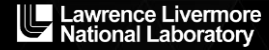


Watch live at NVIDIA SC14 Booth

9 NVIDIA



# US to Build Two Flagship Supercomputers



**SUMMIT**  
150-300 PFLOPS  
Peak Performance

**SIERRA**  
> 100 PFLOPS  
Peak Performance


IBM POWER9 CPU + NVIDIA Volta GPU

NVLink High Speed Interconnect

>40 TFLOPS per Node, >3,400 Nodes


2017

Major Step Forward on the Path to Exascale



**TOP 500<sup>®</sup>**  
SUPERCOMPUTER SITES

Just 4 nodes in Summit would make the Top500 list of supercomputers today

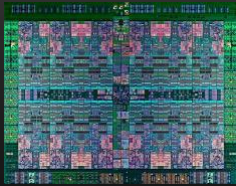


Similar Power as Titan  
5-10x Faster  
1/5th the Size

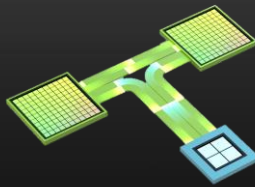


150 PF = 3M Laptops  
One laptop for Every Resident in State of Mississippi

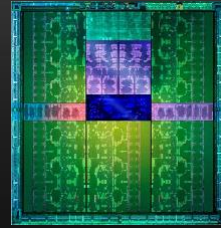
# Accelerated Computing 5x Higher Energy Efficiency



**IBM POWER CPU**  
Most Powerful Serial Processor

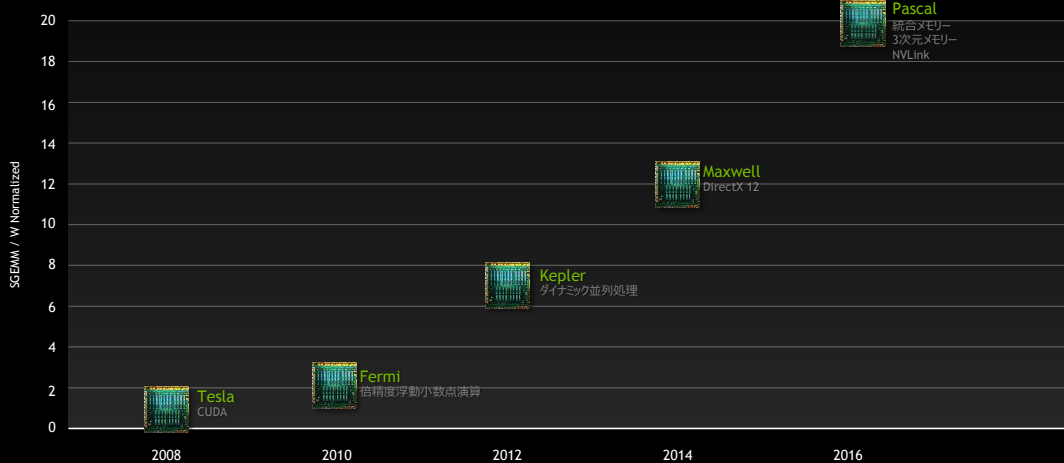


**NVIDIA NVLink**  
Fastest CPU-GPU Interconnect



**NVIDIA Volta GPU**  
Most Powerful Parallel Processor

## GPU ロードマップ



# PASCAL

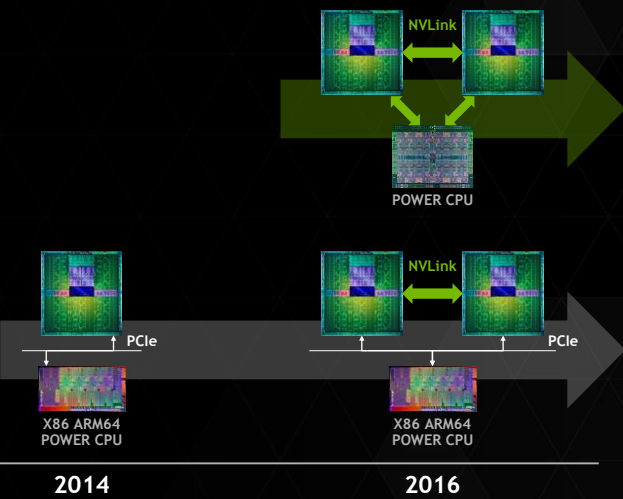


- NVLink PCIe 3.0 の5倍から12倍
- 3次元メモリ 2倍から4倍のメモリバンド幅とサイズ
- モジュール PCIe カードの3分の1のサイズ

## NVLink High-speed GPU Interconnect

### KEPLER GPU

### PASCAL GPU



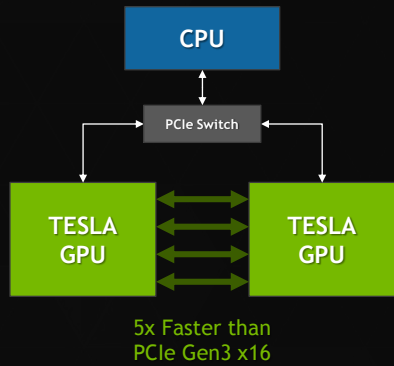
2014

2016



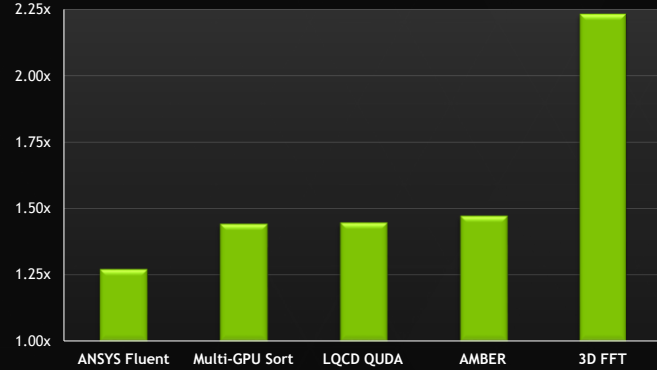
# NVLink Unleashes Multi-GPU Performance

GPUs Interconnected with NVLink



Over 2x Application Performance Speedup  
When Next-Gen GPUs Connect via NVLink Versus PCIe

Speedup vs PCIe based Server



# GPU Technology Theater

Ian Buck with Oak Ridge and Livermore



# US DoE Supercomputers Built on Tesla Platform

Data Center Infrastructure			Development		
<p>System Solutions</p>	<p>Communication</p>	<p>Infrastructure Management</p>	<p>Programming Languages</p>	<p>Development Tools</p>	<p>Software Solutions</p>
<p>GPU Accelerators</p> <p>GPU Boost ...</p>	<p>Interconnect</p> <p>GPU Direct NVLink ...</p>	<p>System Management</p> <p>NVML ...</p>	<p>Compiler Solutions</p> <p>LLVM ...</p>	<p>Profile and Debug</p> <p>CUDA Debugging API ...</p>	<p>Libraries</p> <p>cuBLAS ...</p>

Tesla Accelerated Computing Platform

NVIDIA® TESLA®  
YOUR PLATFORM  
FOR DISCOVERY