

# GPU Computing Ecosystem

## CUDA 5

*Enterprise level GPU Development*

## GPU Development Paths

*Libraries, Directives, Languages*

## GPU Tools

*Tools, libraries and plug-ins for GPU codes*

# Kepler!

## Tesla K10



3x Single Precision

1.8x Memory Bandwidth

Image, Signal, Seismic

Available Now

## Tesla K20



3x Double Precision

Hyper-Q, Dynamic Parallelism

CFD, FEA, Finance, Physics

Available Q4 2012

# Add GPU Processing Your Way

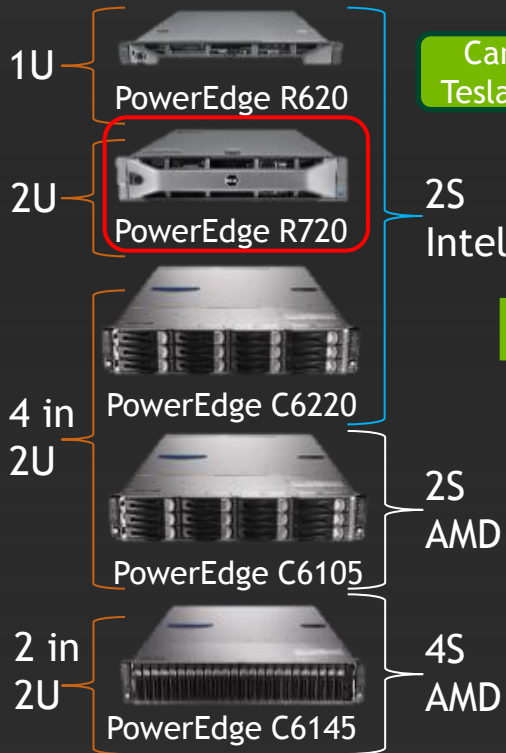
Start with the PowerEdge C410x



- 3U Rack Mount
- **Up to 16 NVIDIA GPUs**
- Up to 8 host server connections
- Programmable GPU:host server ratios
- Individual GPU serviceability



Choose your host server



Can also take 2 Tesla GPU Internal



Choose your GPU:host ratio

- 2:1
  - Two GPUs per server
  - 8 servers per C410x
- 4:1
  - Four GPUs per server
  - 4 servers per C410x

The background features a complex, three-dimensional grid of small, glowing cubes. The cubes are primarily green and cyan, with some transitioning into blue and purple towards the bottom right. The grid is curved and recedes into the distance, creating a sense of depth and perspective. The overall effect is reminiscent of a digital or data landscape.

# Introducing CUDA 5

# CUDA By the Numbers:

---

>375,000,000

CUDA-Capable GPUs

>1,000,000

Toolkit Downloads

>120,000

Active Developers

>500

Universities Teaching CUDA

# Dynamic Parallelism

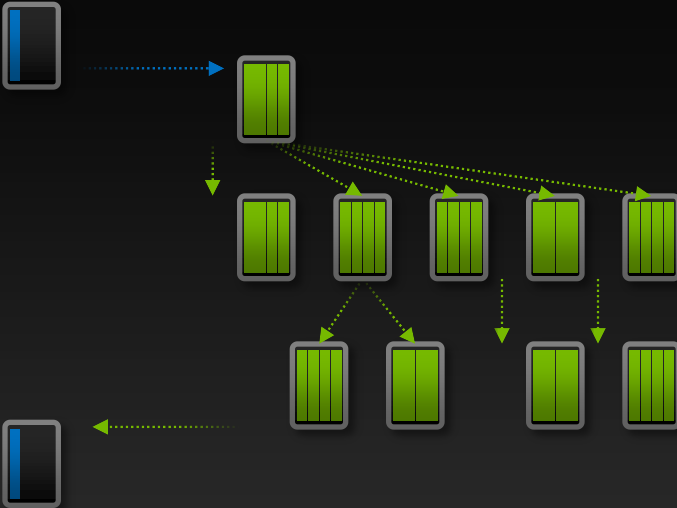
CPU

Fermi GPU



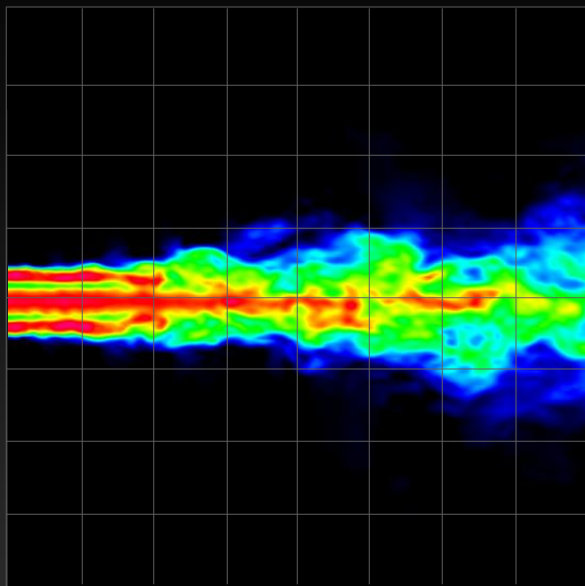
CPU

Kepler GPU



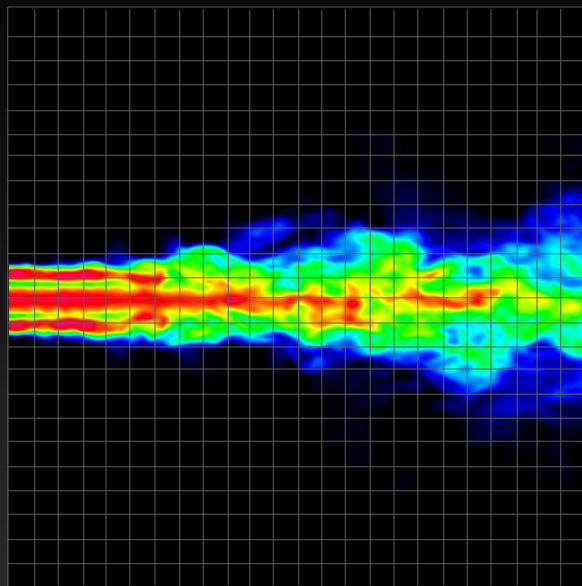
# Dynamic Work Generation

Coarse grid



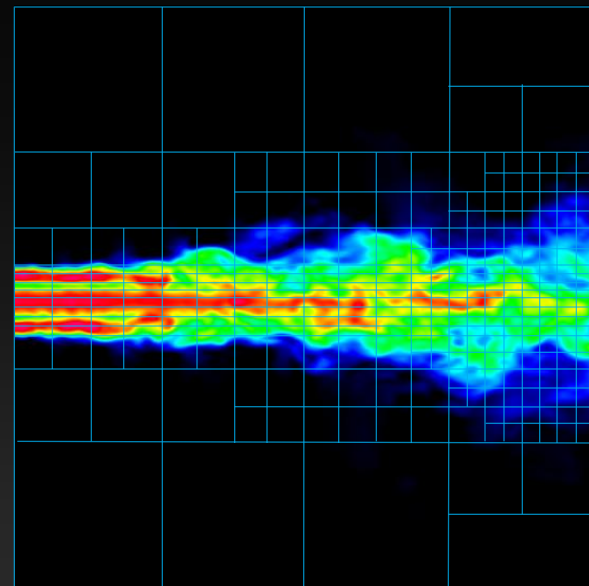
Higher Performance  
Lower Accuracy

Fine grid



Lower Performance  
Higher Accuracy

Dynamic grid



Target performance where  
accuracy is required

# Simpler Code: LU Example

## LU decomposition (Fermi)

```
dgetrf(N, N) {  
  for j=1 to N  
    for i=1 to 64  
      idamax<<<>> → idamax();  
      memcpy ←  
      dswap<<<>> → dswap();  
      memcpy ←  
      dscal<<<>> → dscal();  
      dger<<<>> → dger();  
    next i  
  
    memcpy ←  
    dlaswap<<<>> → dlaswap();  
    dtrsm<<<>> → dtrsm();  
    dgemm<<<>> → dgemm();  
  next j  
}
```

CPU Code

GPU Code

## LU decomposition (Kepler)

```
dgetrf(N, N) {  
  dgetrf<<<>> →  
  
  CPU is Free  
  
  synchronize();  
}
```

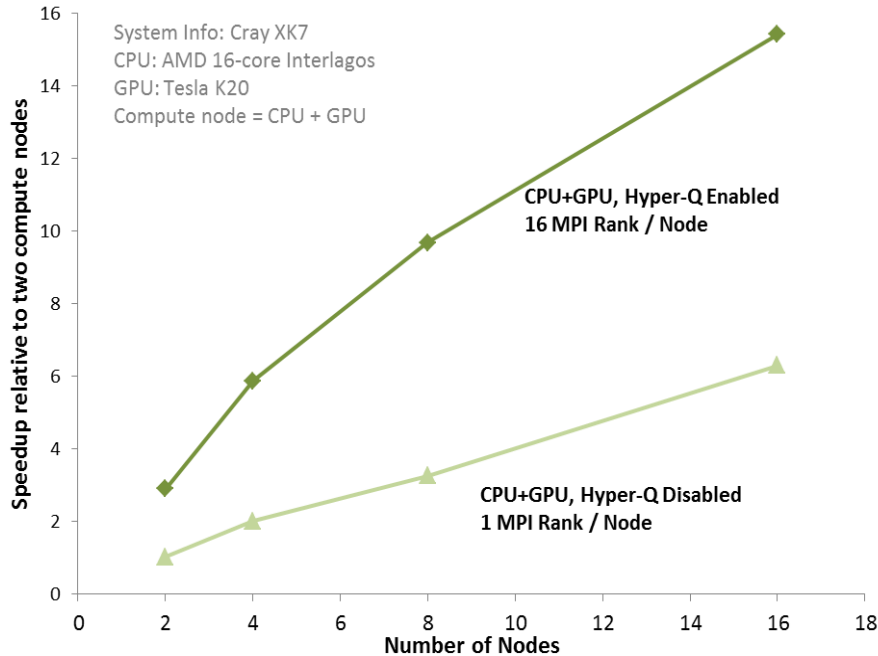
```
dgetrf(N, N) {  
  for j=1 to N  
    for i=1 to 64  
      idamax<<<>>  
      dswap<<<>>  
      dscal<<<>>  
      dger<<<>>  
    next i  
    dlaswap<<<>>  
    dtrsm<<<>>  
    dgemm<<<>>  
  next j  
}
```

CPU Code

GPU Code



# Easy Speed-up for MPI Codes with Hyper-Q



## CP2K Success Story

- M2090 GPU shows no performance improvement over CPUs
- Like most HPC codes, MPI jobs are too small for GPUs
- Hyper-Q provides **2.5x** speed-up without major code rewrite

# GPU Development Paths

“ GPUs have evolved to the point where many real-world applications are easily implemented on them and run significantly faster than on multi-core systems.

Future computing architectures will be hybrid systems with parallel-core GPUs working in tandem with multi-core CPUs

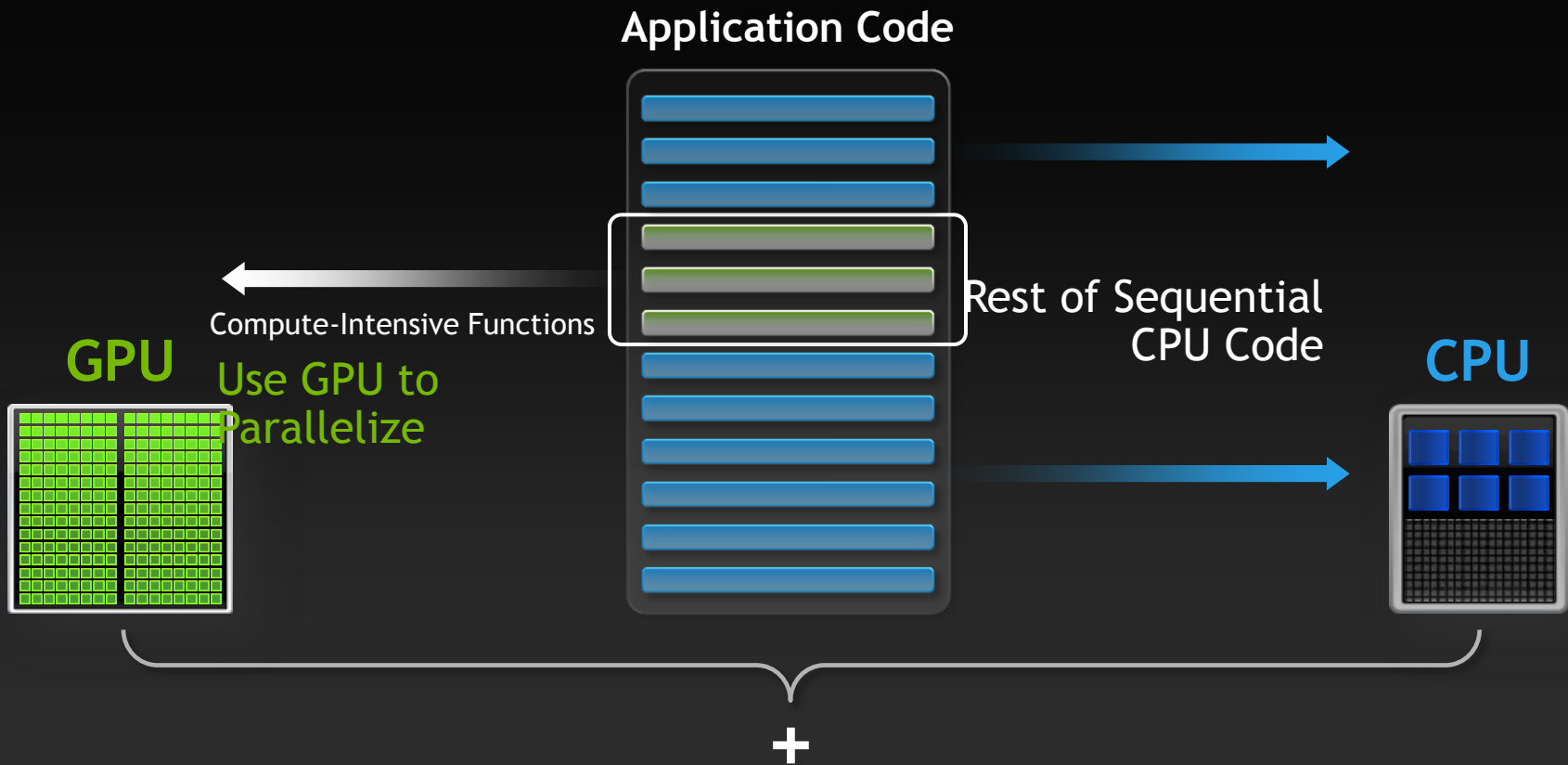
”



**Jack Dongarra**

Professor, University of Tennessee  
Director of the Innovative Computing Laboratory  
Author of LINPACK

# Small Changes, Big Speed-up



# Commercial Apps Accelerated by GPUs

- ▶ **Molecular Dynamics**
  - AMBER • CHARMM • DL\_POLY • GAMESS-US • GROMACS • LAMMPS
  - NAMD
- ▶ **Fluid Dynamics**
  - Altair Acusolve • Autodesk Moldflow • OpenFOAM • Prometech
  - Particlework • Turbostream
- ▶ **Earth Sciences**
  - ASUCA • HOMME • NASA GEOS-5 • NOAA NIM • WRF
- ▶ **Engineering Simulation**
  - Agilent EMPro • ANSYS Mechanical • ANSYS Nexxim • CST Microwave Studio
  - Impetus AFEA • Remcom XFDTD • SIMULIA Abaqus
- ▶ **Others**
  - GADGET2 • MATLAB • Mathematica • NBODY • Paradigm VoxelGeo
  - PARATEC • Schlumberger Petrel

# Rapidly Growing GPU-Accelerated Application Catalog

## POPULAR GPU-ACCELERATED APPLICATIONS

Application	Description	Supported Features	Expected	Multi-GPU	Release Status
<b>Molecular Dynamics</b>					
Abalone	Models molecular dynamics of biopo simulations of proteins, DNA and lig				
ACEMD	Simulation of mechanics force fields, & explicit solvent on CUDA				
AMBER	Suite of programs to simulate molec dynamics on biomolecules				
DL-POLY	Simulate macromolecules, polymers systems, etc on a distributed mem parallel computer				
GROMACS	Simulation of biochemical molecules complicated bond interactions				
HOOMD-Blue	Particle dynamics package written g up for GPUs				
LAMMPS	Classical molecular dynamics packa				
NAMD	Designed for high-performance simu of large molecular systems				
<b>Quantum Chemistry</b>					
GAMESS-US	Computational chemistry suite used to simulate atomic and molecular elect structure				
NWChem	Computational chemistry package de for HPC clusters				
Q-CHEM	Computational chemistry package de for HPC clusters				
TeraChem	Quantum chemistry software designe to run on NVIDIA GPU				
<b>Materials Science</b>					
LSMS	Materials code for investigating the e of temperature on magnetism				
QMCPACK	Solves the many-body Schrodinger eq for electronic structures using a quan Monte Carlo method				
Quantum-Expresso/PWscf	An integrated suite of computer cod electronic structure calculations and modeling at the nanoscale				
VASP	First principles materials code that c electronic structures and quantum-mechanical molecular dynamics				
<b>Visualization &amp; Docking Software</b>					
Amira 5	A multifaceted software platform for visualizing, manipulating, and unders life sciences and bio-medical data				
Core Hopping	Rapid screening of novel cores to imp drug properties				
FastROCS	3D molecular shape comparison				
VMD	Visualizing and analyzing large bio-n systems in 3-D graphics				

## POPULAR GPU-ACCELERATED APPLICATIONS, Continued

Application	Description	Supported Features	Expected	Multi-GPU	Release Status
<b>Weather &amp; Climate Forecasting</b>					
ASUCA	Weather forecasting model fully optimized for GPUs				
CAM/SE	Community Atmospheric Model is a atmosphere model for weather and c research				
GEOS-5	Weather modeling and forecasting ap by NASA				
HIRLAM	Weather forecasting model fully opti for GPUs				
HOMME	Weather modeling tool for atmosph scientists				
HYCOM	Weather forecasting model using ice horizontal grid				
MITgcm	Numerical model designed for study atmosphere, ocean, and climate				
NIM	Weather forecasting model using ice horizontal grid				
WRF	Weather and Ocean modeling applic				
<b>Editing and Effects</b>					
Adobe Premiere Pro	Video editing				
Avi Media Composer	Video editing				
GenArts Sapphire	Effects plug-in for video editing				
Sony Vegas Pro	Video editing				
<b>Animation</b>					
Autodesk 3ds Max	3D modeling, animation, and rendering				
Autodesk Maya	3D modeling, animation, and rendering				
<b>Defense &amp; Intelligence</b>					
DigitalGlobe Advanced Ortho Series	Geospatial Visualization				
Eternix Blaze Terra	Geospatial Visualization				
Exelis (ITT) ENVI	Geospatial Visualization				
GeoEye Analytics Signature Analyst	Geospatial Visualization				
GeoWeb3d Desktop	Geospatial Visualization				
Incognia GIS	Geospatial Visualization				
Intergraph Motion Video Analyst	Video filters and mosaic'ing -- Geo-f PW analytics with intelligence data				
Intuition Panoptes 3.0	Video Analytics				
MetaDef	Video Enhancement				

## POPULAR GPU-ACCELERATED APPLICATIONS, Continued

Application	Description	Supported Features	Expected	Multi-GPU	Release Status
<b>Electronic Design Automation and CEM</b>					
Agilent Technologies ADS	Simulation tool for design of RF, micr and high speed digital circuits				
Agilent Technologies EMPro	Modeling and simulation environm analyzing 3D EM effects of high speed RF/Microwave components				
ANSYS Nexxim	Circuit simulation engine for RF/anal mixed-signal IC design; IBIS-AMI and speedup with GPU computing				
CST Microwave Studio (MWS)	High frequency electromagnetic field simulation				
Gauda OPC, OPV	Collection of several software tools fo computational lithography running of Gauss hardware platform				
Remcom XFDTD	3D EM modeling and simulation				
RocketSim RocketSim	Verilog simulation				
SPEAG SEMCAD-X	3D EM modeling and simulation				
<b>CAD</b>					
CATIA V6 - Live Rendering	Photorealistic rendering				
Bunqspeed Pro Suite	Easy to use photorealistic rendering s				
RTT DeltaGen 10.x	Photorealistic rendering used for des				
RTT DeltaPix	Photorealistic rendering with integrat TeamCenter and RTT formats				
<b>Numerical Analysis</b>					
Jacket Acceler Eyes	GPU acceleration for MATLAB				
Mathematica Wolfram	Symbolic math analysis				
MATLAB Mathworks	Technical computing language and integrated development environment (MATLAB PCT, MDSC)				

## POPULAR GPU-ACCELERATED APPLICATIONS, Continued

Application	Description	Supported Features	Expected	Multi-GPU	Release Status
<b>Oil &amp; Gas</b>					
Accelerate RTM	Seismic Processing				
CGG/Veritas RTM	Seismic Processing				
FFA SVI Pro	Seismic Interpretation				
Headwave Suite	Seismic Imaging				
Geoteric	Seismic Processing/Interpretation				
Paradigm EarthStudy260	Reservoir Modeling				
Paradigm Echos RTM	Seismic Processing				
Paradigm SKUA	Reservoir Modeling				
Paradigm VoxelGeo	Seismic Interpretation				
Schlumberger WesterGeo Omega2 RTM	Seismic Processing				
Seismic City Prostack Interpretation	Seismic Processing				
SpectraSeis	Seismic Processing / Imaging				
Stoneridge Reservoir Simulation	Reservoir Simulation				
Tsunami RTM	Seismic Processing				
<b>Computational Finance</b>					
Howweek Associates	Real-time options analytical engine				
MATLAB Mathworks	Data parallel mathematics (MATLAB PCT, MDSC)				
Murx	Risk analytics (MACS)				
Numerical Algorithms Group	Random Number Generators				
SciComp, Inc	Derivative pricing (SciFinance)				
Wolfram Mathematica	Mathematical Development Environm				

\*GPU performance compared against multi-core x86 CPU socket. kernel to kernel performance comparison. Performance results

## POPULAR GPU-ACCELERATED APPLICATIONS, Continued

Application	Description	Supported Features	Expected Speed Up*	Multi-GPU Support	Release Status
<b>Physics</b>					
Chroma	General purpose LOCD application	Wilson-clover fermions, Krylov solvers, Domain-decomposition	5-6x	Yes	Available now
MILC	General purpose LOCD application	Staggered fermions, Krylov solvers, Gauge-link fattening	5-6x	Yes	Available now
<b>Computational Fluid Dynamics</b>					
Altair AcuSolve	General purpose CFD flow solver	Linear equation solver	2x	Yes	Available now
Autodesk Moldflow	Optimize design of plastic parts and injection molds	Linear equation solver	2x	Single Only	Available now
FEFLD (GMU-Lohner)	Navier-Stokes flow solver based on unstructured grids for modeling both compressible and incompressible flows	Explicit solver	10x	Yes	In Development
Fluidyna L. Bultra	Computing physical flows in and around solid bodies	IBM, particle CFD	20x	Yes	Available now
Fluidyna Cullis-OpenFOAM	Computing physical flows with Cullis - a software library with special algorithms for solving systems of equations	Linear equation solvers	3x Solver	Single Only	Available now
Promotech Particleworks	Fluid simulation for free surface flow like Tsunami, material processing and liquids	MPS, Particle CFD	4x-9x	Yes	Available now
S3D (Sandia NL S3D)	Massively parallel direct numerical solver (DNS) for the full compressible Navier-Stokes	Chemistry kernel	8x SP, 5x DP kernel	Yes	In Development
Turbostream	Ultrafast CFD solver for turbomachines	Explicit solver	19x	Yes	Available now
Vratis SpeedIT-OpenFOAM Solver	Set of accelerated solvers for sparse linear systems of equations	Linear equation solvers	3x Solver	Yes	Available now
<b>Computational Structural Mechanics</b>					
Abaqus/Standard	Simulation and analysis tool for structural mechanics	Linear equation solver	1.5-2.5x	Single Only	Available now
ANSYS Mechanical	Simulation and analysis tool for structural mechanics	Linear equation solver	2x	Single Only	Available now
Impetus Afea	Predicts large deformations of structures and components exposed to extreme loading conditions	Linear equation solver, SPH	10x SPH, 2x Total	Yes	Available now
LS-DYNA Implicit	Multiphysics simulation package used	Linear equation solver	3x	Yes	In Development
MSC Nastran	Simulation and analysis tool for structural mechanics	Linear equation solver	1.4-2x	Yes	Available now
Marc	Simulation and analysis tool for structural mechanics	Linear equation solver	1.5x	Yes	In Development
RADIOSS Implicit	Used to maximize durability, NVH, crash, safety, manufacturability and fluid-structure interaction performance	Linear equation solver	2x	Single Only	In Development

# 3 Ways to Accelerate Applications

Applications

Libraries

“Drop-in”  
Acceleration

OpenACC  
Directives

Easily Accelerate  
Applications

Programming  
Languages

Maximum  
Flexibility

# GPU Accelerated Libraries

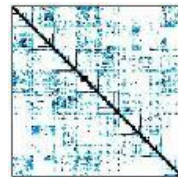
## “Drop-in” Acceleration for your Applications



NVIDIA cuFFT



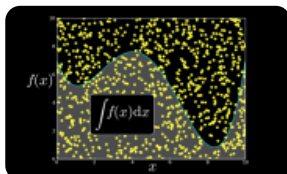
NVIDIA cuBLAS



NVIDIA cuSPARSE



NVIDIA NPP



NVIDIA cuRAND



Matrix Algebra on  
GPU and Multicore



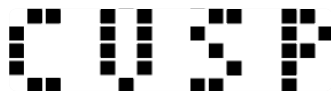
GPU Accelerated  
Linear Algebra



Vector Signal  
Image Processing



IMSL Library



Sparse Linear Algebra

ArrayFire



Building-block  
Algorithms



C++ Templated  
Parallel Algorithms



# GPU-Aware MPI Libraries

## Integrated Support for GPU Computing



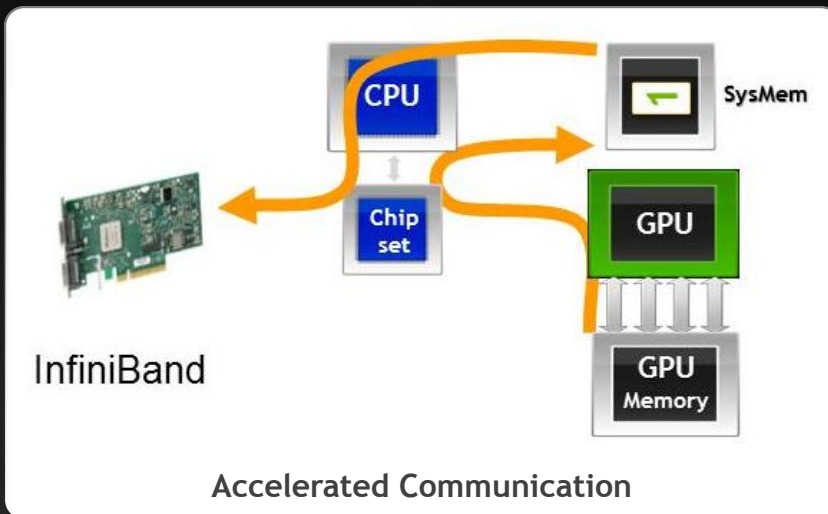
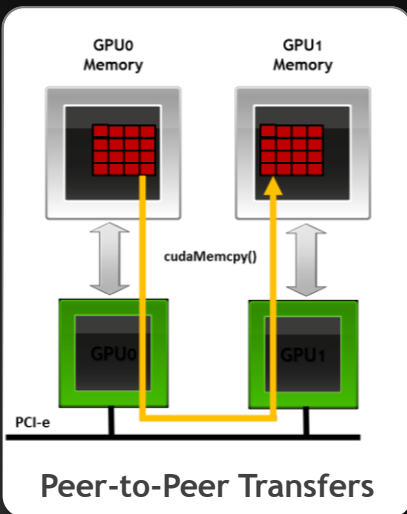
OpenMPI



MVA PICH



Platform MPI



# 3 Ways to Accelerate Applications

Applications

Libraries

“Drop-in”  
Acceleration

OpenACC  
Directives

Easily Accelerate  
Applications

Programming  
Languages

Maximum  
Flexibility

# OpenACC

## Open Programming Standard for Parallel Computing

“OpenACC will enable programmers to easily develop portable applications that maximize the performance and power efficiency benefits of the hybrid CPU/GPU architecture of Titan.”

*--Buddy Bland, Titan Project Director, Oak Ridge National Lab*



“OpenACC is a technically impressive initiative brought together by members of the OpenMP Working Group on Accelerators, as well as many others. We look forward to releasing a version of this proposal in the next release of OpenMP.”

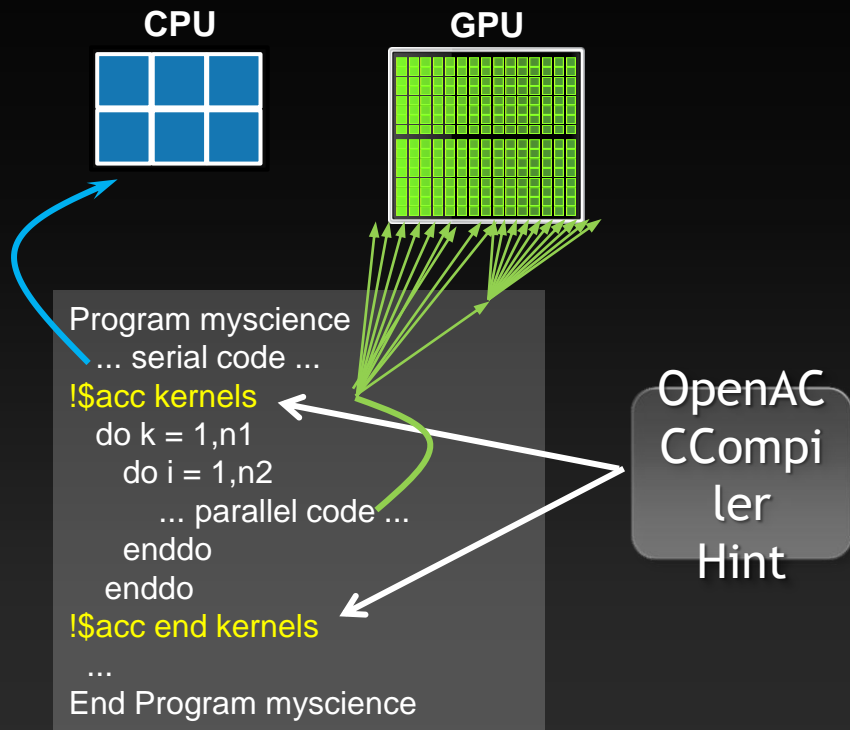
*--Michael Wong, CEO OpenMP Directives Board*



## OpenACC Standard



# OpenACC Directives



Your original  
Fortran or C code

Simple Compiler hints

Compiler Parallelizes code

Works on many-core GPUs &  
multicore CPUs



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

# Directives: Easy & Powerful

## Real-Time Object Detection

Global Manufacturer of Navigation Systems



**5x** in 40 Hours

## Valuation of Stock Portfolios using Monte Carlo

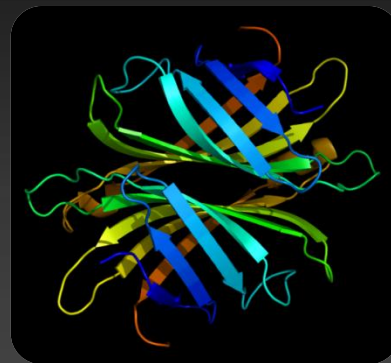
Global Technology Consulting Company



**2x** in 4 Hours

## Interaction of Solvents and Biomolecules

University of Texas at San Antonio



**5x** in 8 Hours

“Optimizing code with directives is quite easy, especially compared to CPU threads or writing CUDA kernels. The most important thing is avoiding restructuring of existing code for production applications.”

# 3 Ways to Accelerate Applications

Applications

Libraries

“Drop-in”  
Acceleration

OpenACC  
Directives

Easily Accelerate  
Applications

Programming  
Languages

Maximum  
Flexibility

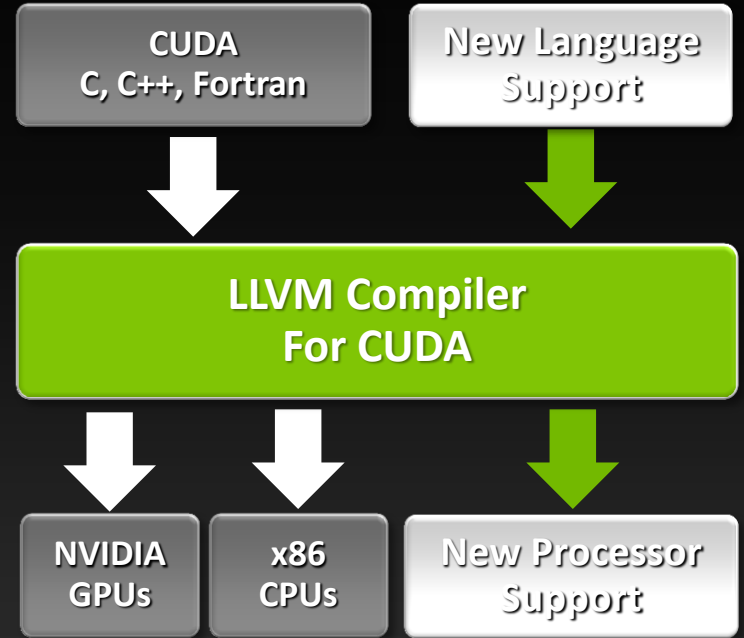
# Opening the CUDA Platform with LLVM

CUDA compiler source contributed to open source LLVM compiler project

SDK includes specification documentation, examples, and verifier

**Provides ability for anyone to add CUDA to new languages and processors**

Learn more at  
[developer.nvidia.com/cuda-source](http://developer.nvidia.com/cuda-source)



# GPU Programming Languages

---

Numerical analytics ▶

MATLAB, Mathematica, LabVIEW

Fortran ▶

OpenACC, CUDA Fortran

C ▶

OpenACC, CUDA C

C++ ▶

Thrust, CUDA C++

Python ▶

PyCUDA, Copperhead

C# ▶

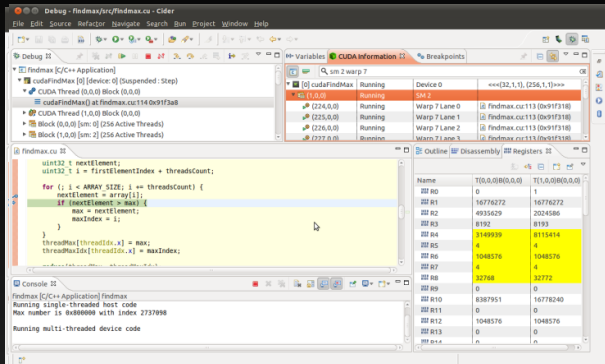
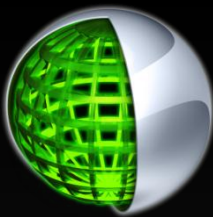
GPU.NET





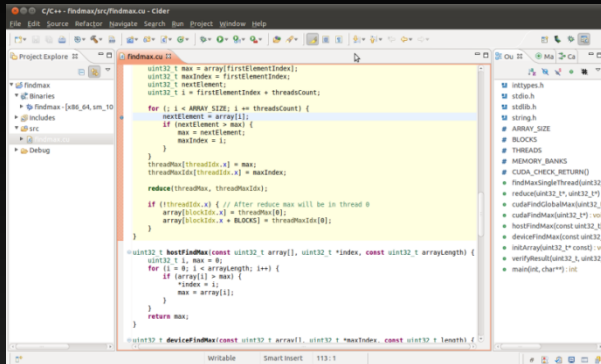
# Development Tools

# NVIDIA® Nsight™ Eclipse Edition for Linux and MacOS



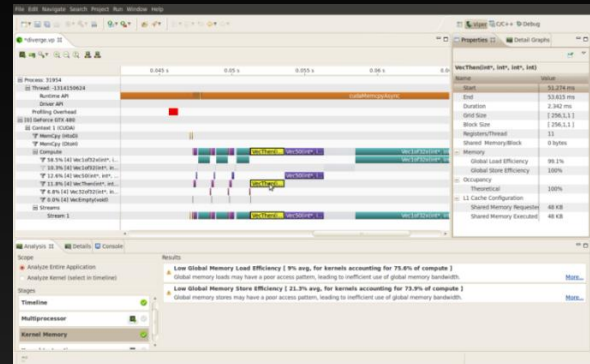
## CUDA-Aware Editor

- Automated CPU to GPU code refactoring
- Semantic highlighting of CUDA code
- Integrated code samples & docs



## Nsight Debugger

- Simultaneously debug CPU and GPU threads
- Inspect variables across CUDA
- Use breakpoints & single-step debugging



## Nsight Profiler

- Quickly identifies performance issues
- Integrated expert system
- Source line correlation

# Debugging Solutions

Command Line to Cluster-Wide



**NVIDIA Nsight**  
Eclipse & Visual Studio Editions



**NVIDIA CUDA-GDB**  
for Linux & Mac



**NVIDIA CUDA-MEMCHECK**  
for Linux & Mac



**Allinea DDT with CUDA**  
Distributed Debugging Tool



**TotalView for CUDA**  
for Linux Clusters

# Performance Analysis Tools

Single Node to Hybrid Cluster Solutions



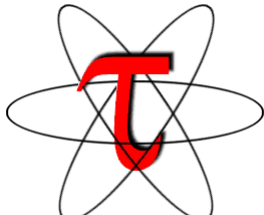
**NVIDIA Nsight**  
Eclipse & Visual Studio Editions



**NVIDIA Visual Profiler**



**Vampir Trace Collector**



Tuning and Analysis Utilities

**TAU Performance System**



**PAPI CUDA Component**



**Under Development**

# Job Scheduling & Cluster Management



LSF, HPC, Cluster Manager



Bright Cluster Manager



ROCKS+MOAB



PBS Professional



NVML Plugin for GPUs



Univa Grid Engine



# GPU Test Drive

## Experience the Acceleration

**What** ▶

Program that provides free access to a remote/cloud GPU cluster

**Who** ▶

For academic researchers

**Why** ▶

To experience how applications accelerate with GPUs

### Benefit to Researchers

FREE & EASY way to start with GPUs

No GPU programming expertise needed

Access to a remote, pre-configured GPU cluster for evaluation

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

# GPU Technology Conference 2013

## March 18-21 | San Jose, CA

### Reasons to attend GTC

- Learn about leading-edge advances in GPU computing
- Explore the research as well as the commercial applications
- Discover advances in computational visualization
- Take a deep dive into parallel programming

### Ways to participate

- Submit a Research Poster - share your work and gain exposure as a thought leader
- Register - learn from the experts and network with your peers
- Exhibit/Sponsor - promote your organization as a key player in the GPU ecosystem

Visit [www.gputechconf.com](http://www.gputechconf.com)



**Thank you**