Maximize automotive simulation productivity with ANSYS HPC and NVIDIA GPUs

Bhushan Desam, Ph.D.
NVIDIA Corporation
NVIDIA Enterprise Group
Visualization, Accelerated Computing & Virtualization

**QUADRO**
Revolutionizing Design & Visualization

**TESLA**
Accelerating Momentum in HPC and Big Data Analytics

**GRID**
Enabling End-to-End Enterprise Virtualization
GPU enable tremendous breakthroughs by simply enabling us to do more, faster.

ANSYS and other major ISVs leverage GPUs to accelerate engineering simulation for better design.

Many of the world’s most powerful supercomputers run on Tesla GPUs, including TITAN at Oak Ridge National Labs.
NVIDIA’s Use of ANSYS in Product Development

**ANSYS Icepak**
- Active and passive cooling of IC packages

**ANSYS Mechanical**
- Large deflection bending of PCBs

**ANSYS Mechanical**
- Comfort and fit of 3D emitter glasses

**ANSYS Mechanical**
- Shock and vibration of solder ball assemblies
Changing Role of Simulation in Product Development

From insight to product innovation

Building insight

Experience envelope

Design variable 1

Design variable 2

Final design concept

E – Experiment

S - Simulation

Innovation envelope

Design variable 1

Design variable 2

Final design concept

Simulation-driven product innovation
Computing Capacity is Still a Major Challenge

Frequency of limiting size/detail in simulation models due to compute infrastructure or turnaround time limitations

- For some models: 57%
- Nearly every model: 34%
- Almost never: 9%

Source: Survey by ANSYS with over 1,800 respondents
An ongoing effort designed to remove computing limitations from engineers who use computer aided engineering in all phases of design, analysis, and testing.

It is a hardware and software initiative!
Hardware trends in HPC

Recent advancements have revolutionized the computational speed available on the desktop

- Multi-core processors
- GPGPU (general-purpose computing on graphics processing units)
- Large amounts of RAM
- Solid state derives (SSD)

NVIDIA GPU

6X more flops compared to the latest CPU
Basics of GPU Computing

- GPU is an accelerator attached to an x86 CPU
- GPU acceleration is user-transparent
  - Jobs launch and complete without additional user steps
- Schematic of a CPU with an attached GPU accelerator
  - CPU begins/ends job, GPU manages heavy computations
What do GPUs bring to ANSYS simulations?

By handling parallel tasks in matrix solutions efficiently, GPUs accelerate ANSYS simulation by 2-3 X.
Benefits of GPU-accelerated simulations

More simulations in the same amount of time or same number of simulations in less amount of time

- Improve product-quality: More design points can be analyzed for better-quality products without slipping project schedules
- Faster time-to-market: Simulation times can be cut into half, thus shorter product development times
- Complex simulations: Mesh sizes can be doubled or advanced models can be used without increasing simulation times
ANSYS 15.0 HPC licenses (new)

- Treats each GPU socket as a CPU core, which significantly increases simulation productivity from HPC licenses

- All ANSYS HPC products unlock GPUs in 15.0, including HPC, HPC Pack, HPC Workgroup, and HPC Enterprise products.

GPU support in ANSYS – Licensing comparison

<table>
<thead>
<tr>
<th>License type</th>
<th>ANSYS 14.5</th>
<th>ANSYS 15.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPC per core licenses</td>
<td>None</td>
<td>1 HPC license – 1 GPU</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 HPC licenses – 2 GPUs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‘N’ HPC licenses – ‘N’ GPUs</td>
</tr>
<tr>
<td>HPC pack</td>
<td>1 HPC pack – 1 GPU</td>
<td>1 HPC pack – 4 GPUs</td>
</tr>
<tr>
<td></td>
<td>2 HPC packs – 4 GPUs</td>
<td>2 HPC packs – 16 GPUs</td>
</tr>
</tbody>
</table>
ANSYS® Fluent on NVIDIA GPUs

ANSYS® Fluent 15.0
GPU Acceleration in Fluent 15.0

In flow problems, pressure-based coupled solver typically spends 60-70% time in AMG whereas segregated solver only spends 30-40% of the time in AMG.

Higher AMG times are ideal for GPU acceleration, thus coupled-problems benefit from GPUs.
Fluent Speed-up from GPU acceleration

![Graph showing speed-up factors for Fluent with different solver types and linear solver fractions.]

- **Coupled solver**
- **Segregated solver**
- **AMG speed-up on GPU**

- For coupled solver, the speed-up factor increases as the linear solver fraction increases.
- For segregated solver, the speed-up factor is constant at 2.0 across different linear solver fractions.
- AMG speed-up on GPU shows a substantial increase in speed-up with a linear solver fraction of 0.7, reaching a speed-up factor of 3.5.

**Fluent Speed-up from GPU acceleration**
ANSYS Fluent convergence

Coupled vs segregated solver

Coupling Momentum and Continuity Increases CFD Robustness

FLUENT technology introduces a pressure-based coupled solver to reduce computation time for low-speed compressible and incompressible flow applications.

By Franklin J. Kreys, Applications Specialist, ANSYS, Inc.


TRUCK BODY MODEL
(14 million cells)

Coupled
Stable convergence for drag coefficient at ~550 iterations

Segregated
Oscillating behavior for drag even after ~6000 iterations
GPU Acceleration of Sedan Model

ANSYS Fluent 15.0 performance on pressure-based coupled Solver

1.9x

12
Jobs/day

15
Jobs/day

CPU only

Coupled solver

27
Jobs/day

CPU only

CPU + GPU

Higher is Better

Sedan Model

• Sedan geometry
• 3.6M mixed cells
• Steady, turbulent
• External aerodynamics
• Coupled PBNS, DP
• CPU: Intel Xeon E5-2680; 8 cores
• GPU: 2 X Tesla K40

NOTE: Times for total solution until convergence

Convergence criteria: 10e-03 for all variables; No of iterations until convergence: segregated CPU-2798 iterations (7070 secs); coupled CPU-967 iterations (5900 secs); coupled 985 iterations (3150 secs)
GPU Acceleration of Water Jacket Analysis

ANSYS Fluent 15.0 performance on pressure-based coupled Solver

**Water jacket model**
- Unsteady RANS model
- Fluid: water
- Internal flow
- CPU: Intel Xeon E5-2680; 8 cores
- GPU: 2 X Tesla K40

**NOTE:** Times for 20 time steps
GPU value proposition for Fluent 15.0

Formula 1 aerodynamic study (144 million cells)

25 secs/iter
2.1x
12 secs/iter

Lower is Better

2X

CPU only
160 cores

CPU + GPU
32 X K40

Additional cost of adding GPUs and HPC licenses
CPU-only solution cost: 100%

Simulation productivity from CPU-only system

Cost

Benefit

CPU
GPU

Additional productivity from HPC licenses and GPUs

Additional productivity from GPUs

55%
110%

100%
100%

CPU-only solution cost is approximated and includes both hardware and paid-up software license costs. Benefit/productivity is based on the number of completed Fluent jobs/day.

All results are based on turbulent flow over an F1 case (144 million cells) over 1000 iterations; steady-state, pressure-based coupled solver with single-precision; CPU: 8 Ivy Bridge nodes with 20 cores each, F-cycle, size 8; GPU: 32 X Tesla K40, V-cycle, size 2 and max AMG iterations 5.
GPU Scaling in Fluent 15.0 (F1 case)

All results are based on turbulent flow over an F1 case (144 million cells) over 1000 iterations; steady-state, pressure-based coupled solver with single-precision; CPU: Ivy Bridge with 10 cores per socket, F-cycle, size 8; GPU: Tesla K40, V-cycle, size 2 and max AMG iterations 5.
Shorter Time to Solution with GPUs at PSI Inc.

A customer success story

Objective
Meeting engineering services schedule & budget, and technical excellence are imperative for success.

HPC Solution
- PSI evaluates and implements the new technology in software (ANSYS 15.0) and hardware (NVIDIA GPU) as soon as possible.
- GPU produces a 43% reduction in Fluent solution time on an Intel Xeon E5-2687 (8 core, 64GB) workstation equipped with an NVIDIA K40 GPU

Design Impact
Increased simulation throughput allows meeting delivery-time requirements for engineering services.
Performance Tips for ANSYS Fluent on GPUs

• GPUs accelerate the AMG solver portion of the CFD analysis, thus benefit problems with relatively high %AMG
  – Coupled solvers have high %AMG in the range of 60-70%
  – Fine meshes and low-dissipation problems have high %AMG

• In some cases, pressure-based coupled solvers offer faster convergence compared to segregated solvers (problem-dependent)

• The whole problem must fit on GPUs for the calculations to proceed
  – In pressure-based coupled solver, each million cells need approx. 4 GB of GPU memory
  – High-memory cards such as Tesla K40 or Quadro K6000 are ideal

• Moving scalar equations such as turbulence may not benefit much because of low workloads (using ‘scalar yes’ option in ‘amg-options’)

• Better performance on lower CPU core counts
  – A ratio of 4 or 5 CPU cores to 1 GPU is recommended
ANSYS® Mechanical 15.0
GPU Acceleration in Mechanical 15.0

ANSYS Mechanical 15.0 on Tesla K40

**Simulation productivity (with a HPC core license)**
- 2 CPU cores: 108 jobs/day
- 2 CPU cores + Tesla K40: 410 jobs/day
- 3.8X Increase

**Simulation productivity (with a HPC pack)**
- 8 CPU cores: 307 jobs/day
- 7 CPU cores + Tesla K40: 757 jobs/day
- 2.5X Increase

Distributed ANSYS Mechanical 15.0 with Sandy Bridge (Xeon E5-2687W 3.1 GHz) 8-core CPU and a Tesla K40 GPU with boost clocks; V145sp-5 model, Turbine geometry, 2.5 M DOF, and direct sparse solver.

V14sp-5 Model
- Turbine geometry
- 2,100,000 DOF
- SOLID187 FEs
- Static, nonlinear
- Distributed ANSYS 15.0
- Direct sparse solver
GPU value proposition for Mechanical 15.0

*With one HPC license + Tesla K20*

### Simulation productivity

V14sp-6 benchmark, 4.9 M DOF, static non-linear analysis; direct sparse solver, distributed ANSYS Mechanical 15.0 with Intel Xeon E5-2697 v2 2.7 GHz CPU; Tesla K20 GPU.

**ANSYS Mechanical jobs/day**

- 2 CPU cores: 59
- 2 CPU cores + Tesla K20: 165

**Higher is Better**

**2.8X**

**7X**

**Additional productivity from a GPU and one HPC license**

**Cost**

- CPU: 100%
- GPU: 25%

**Benefit**

- Additional cost of adding a GPU + HPC license: 180%
- Simulation productivity from GPU: 100%
- Additional productivity from CPU-only system: 100%

Additional cost of adding a GPU + HPC license

CPU-only solution cost is approximated and includes both hardware and software license costs. Benefit is based on the number of completed Mechanical jobs/day.
GPU value proposition for Mechanical 15.0

*With an HPC pack + Tesla K20*

**Simulation productivity**

V14sp-6 benchmark, 4.9 M DOF, static non-linear analysis; direct sparse solver, distributed ANSYS Mechanical 15.0 with Intel Xeon E5-2697 v2 2.7 GHz CPU; Tesla K20 GPU.

**Additional productivity from a GPU and HPC pack**

**Cost**

- CPU: 100%
- GPU: 50%

**Benefit**

- Additional productivity from GPU: 50%
- Additional cost of adding a GPU: 12%

CPU-only solution cost is approximated and includes both hardware and software license costs. Benefit is based on the number of completed Mechanical jobs/day.
Performance Tips for ANSYS Mechanical on GPUs

- GPUs accelerate the solver part of analysis, consequently problems with high solver workloads benefit the most from GPUs
  - Characterized by both high DOF and high factorization requirements
  - Models with solid elements (such as castings) and have >500K DOF experience good speedups
- Better performance when run on DMP mode over SMP mode
- GPU and system memories both play important roles in performance
  - Sparse solver:
    - Bulkier and/or higher-order FE models are good and will be accelerated
    - If the model exceeds 5M DOF, then either add another GPU with 5-6 GB of memory (Tesla K20 or K20X) or use a single GPU with 12 GB memory (Tesla K40 or Quadro K6000).
  - PCG/JCG solver:
    - Memory saving (MSAVE) option should be turned off for enabling GPUs
    - Models with lower Level of Difficulty value (Lev_Diff) are better suited for GPUs
NVIDIA Kepler family GPUs for ANSYS simulations

K20 (5 GB)
K20X (6 GB)
K40 (12 GB)
K6000 (12 GB)
Thank you

Bhushan Desam
bdesam@nvidia.com