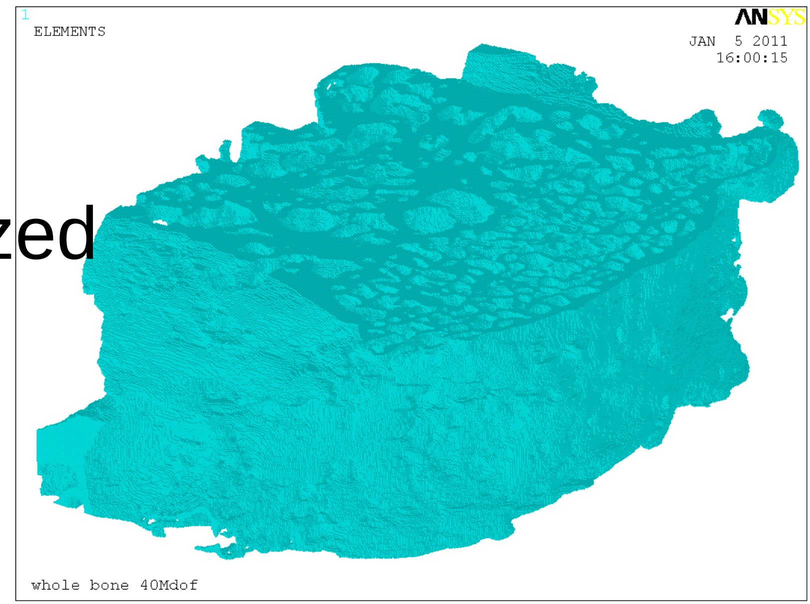
HPC Solutions for Efficiently Solving Large Structural Models

Pierre LOUAT
ANSYS FRANCE

- Agenda : Structural models : Needs & Trends
  - Introduction / Short ANSYS HPC history
  - Shared Memory vs Distributed Memory
  - ANSYS Structural solver scalability
  - Examples
  - ANSYS using GPU technology
    - Introduction
    - Benefits
  - Conclusion

Need for Speed

- Implicit structural FEA codes
  - Runtimes can be hours, days or even weeks
    - **Lots of computations!!**
- Mesh fidelity continues to increase
  - More equations to solve
    - **More computations!!**
- More complex physics being analyzed
  - More nonlinear solutions
    - **More computations!!**



Long History of HPC!

ANSYS Inc. has a strong commitment to High Performance Computing.

2009
- Ideal scaling to 2048 cores (fluids)
- Teraflop performance at 512 core (structures)
- Parallel I/O (fluids)
- Domain Decomposition introduced (HFSS 12)

2005 - 2006
- Parallel meshing (fluids)
- Support for clusters using Windows HPC

2001 - 2003
- Parallel dynamic moving/deforming mesh
- Distributed memory particle tracking

1998-1999
- Integration with load management systems
- Support for Linux clusters, low latency interconnects
- 10M cell fluids simulations, 128 processors

1994 - 1995
- Parallel dynamic mesh refinement and coarsening
- Dynamic load balancing

1993
- 1st general-purpose parallel CFD with interactive client-server user environment

1980 — 1990 — 2000 — 2010

GPU

2007 - 2008
- Optimized performance on multicore processors
- 1st One Billion cell fluids simulation

2005 -2007
- Distributed sparse solver
- Distributed PCG solver
- Variational Technology
- DANSYS released
- Distributed Solve (DSO) HFSS 10

2004
- 1st company to solve 100M structural DOF

1999 - 2000
- 64bit large memory addressing
- Shared memory multiprocessing (HFSS 7)

1994
- Iterative PCG Solver Introduced for large structural analysis

1990
- Shared Memory Multiprocessing for structural simulations

1980's
- Vector Processing on Mainframes

Parallel Processing - Hardware

- 2 Types of memory systems
  - Shared memory (SMP) ← single box, workstation/server
  - Distributed memory (DMP) ← multiple boxes, cluster

Workstation

Cluster

Parallel Processing – Hardware + Software

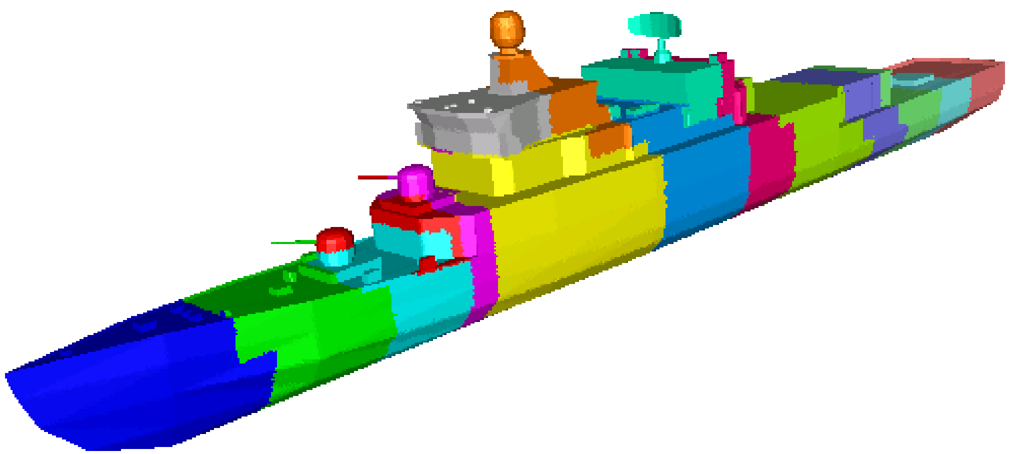
	Laptop/Desktop or Workstation/Server	Cluster
ANSYS	YES	--
Distributed ANSYS	YES	YES

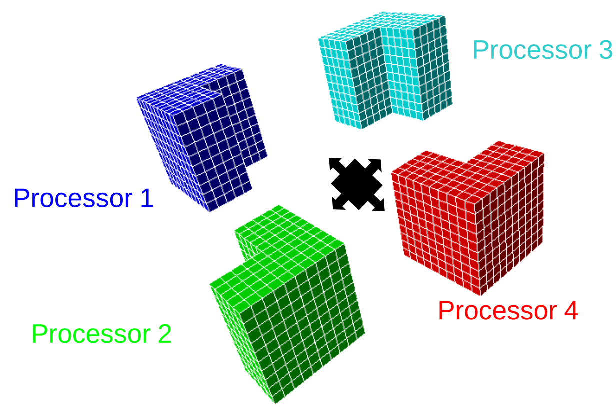
- R9.0 First release of Distributed ANSYS
- R11.0 Support for 1st distributed eigensolver (LANPCG)
  - Support for full harmonic analyses
- R12.0 Support for unsymmetric matrices
  - Achieved over 1 Tflops using 512 cores
- R13.0 ***Achieved over 3 Tflops using over 1024 cores***

Distributed ANSYS Design Requirements

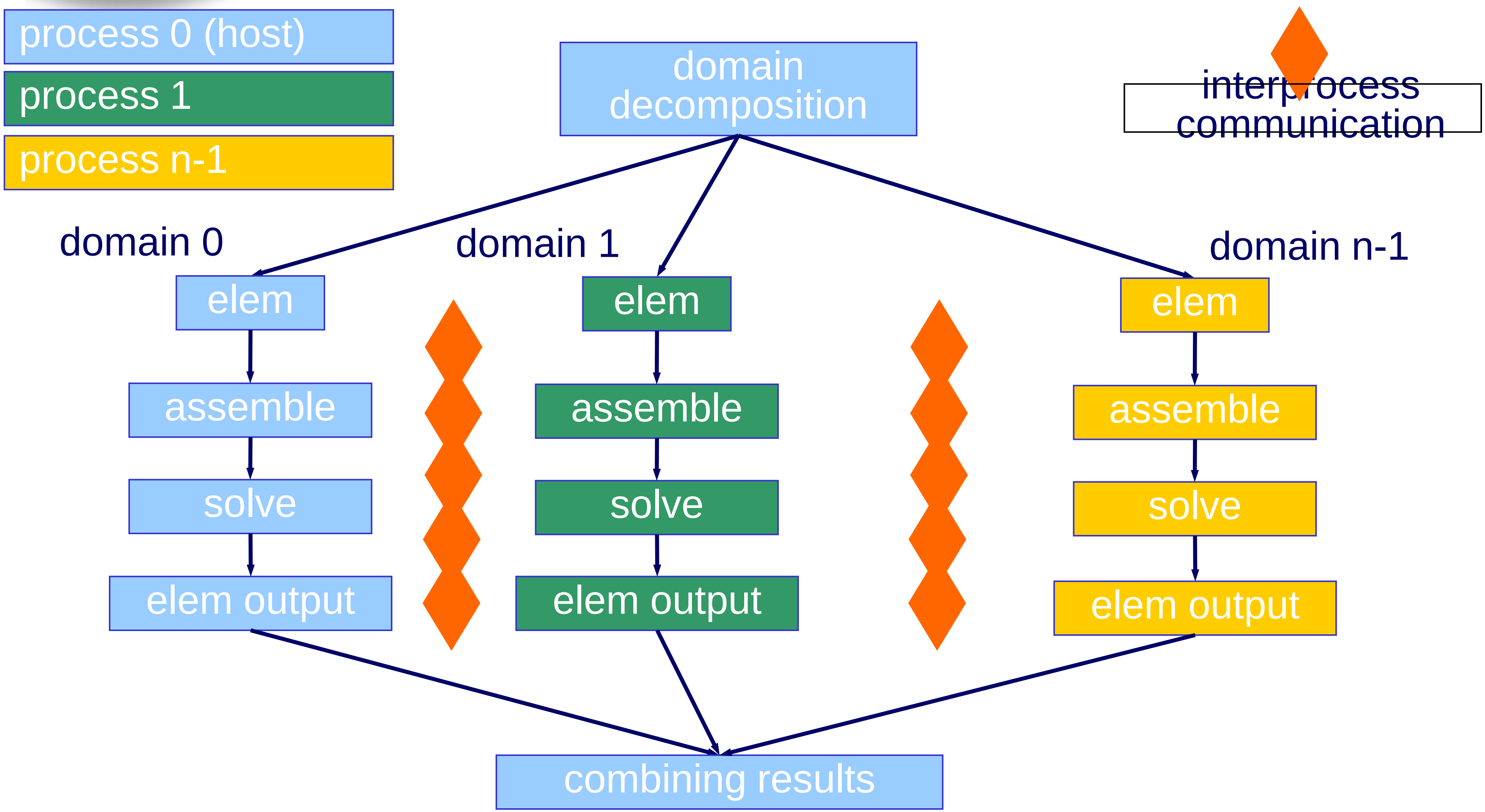
- No limitation in simulation capability
- Reproducible and consistent results
- Support all major platforms

- Domain decomposition approach
  - Break problem into *N* pieces (domains)
  - "Solve" the global problem independently within each domain
  - Communicate information across the boundaries as necessary





Distributed ANSYS Architecture



Distributed ANSYS Solvers

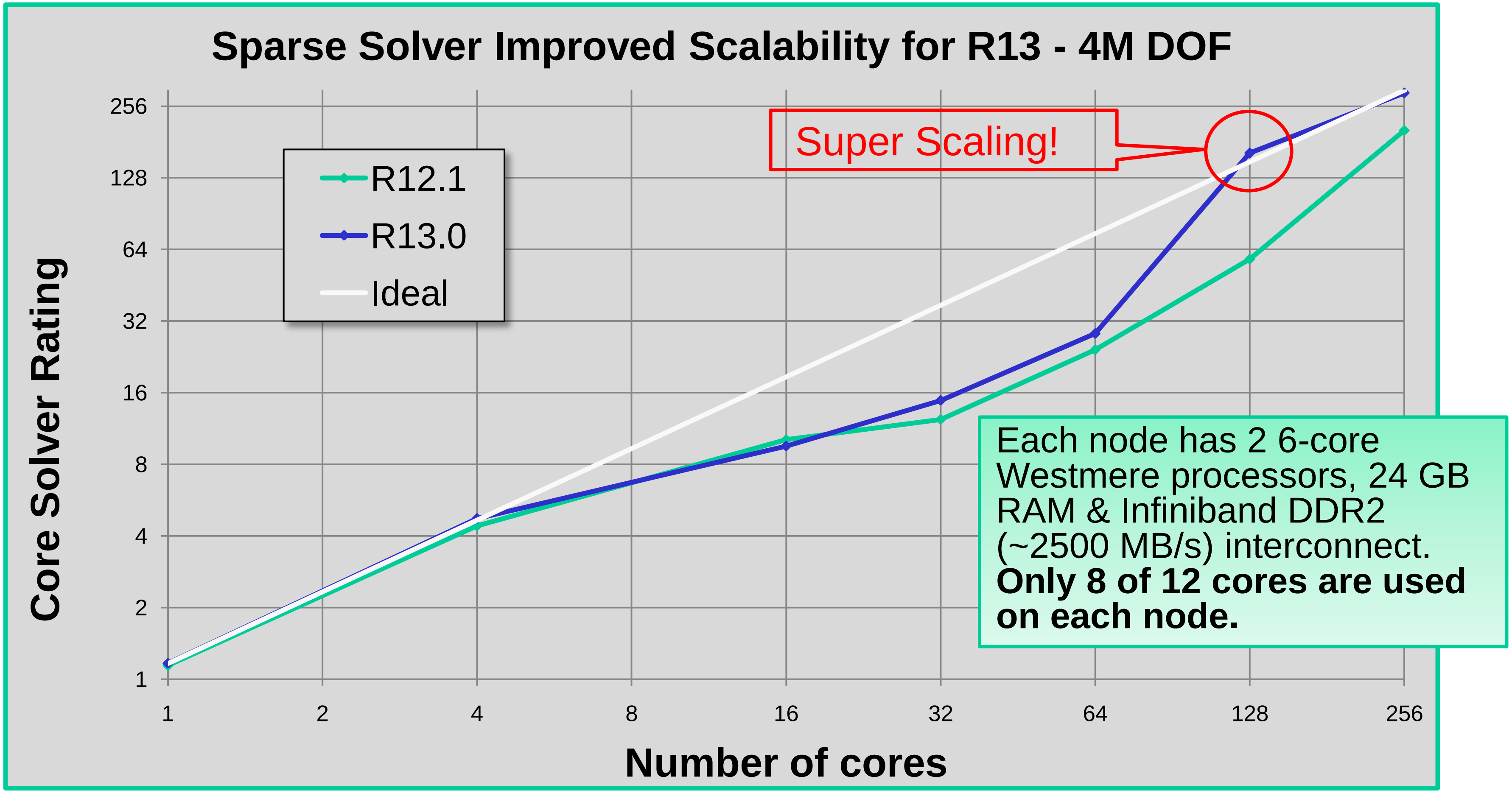
- Distributed sparse (default)
  - Supports all analyses supported with DANSYS (Linear, Non Linear, Static , Transient)
- Distributed PCG
  - For static and full transient analyses
- Distributed LANPCG (eigensolver)
  - For modal analyses

Benefits of Distributed ANSYS

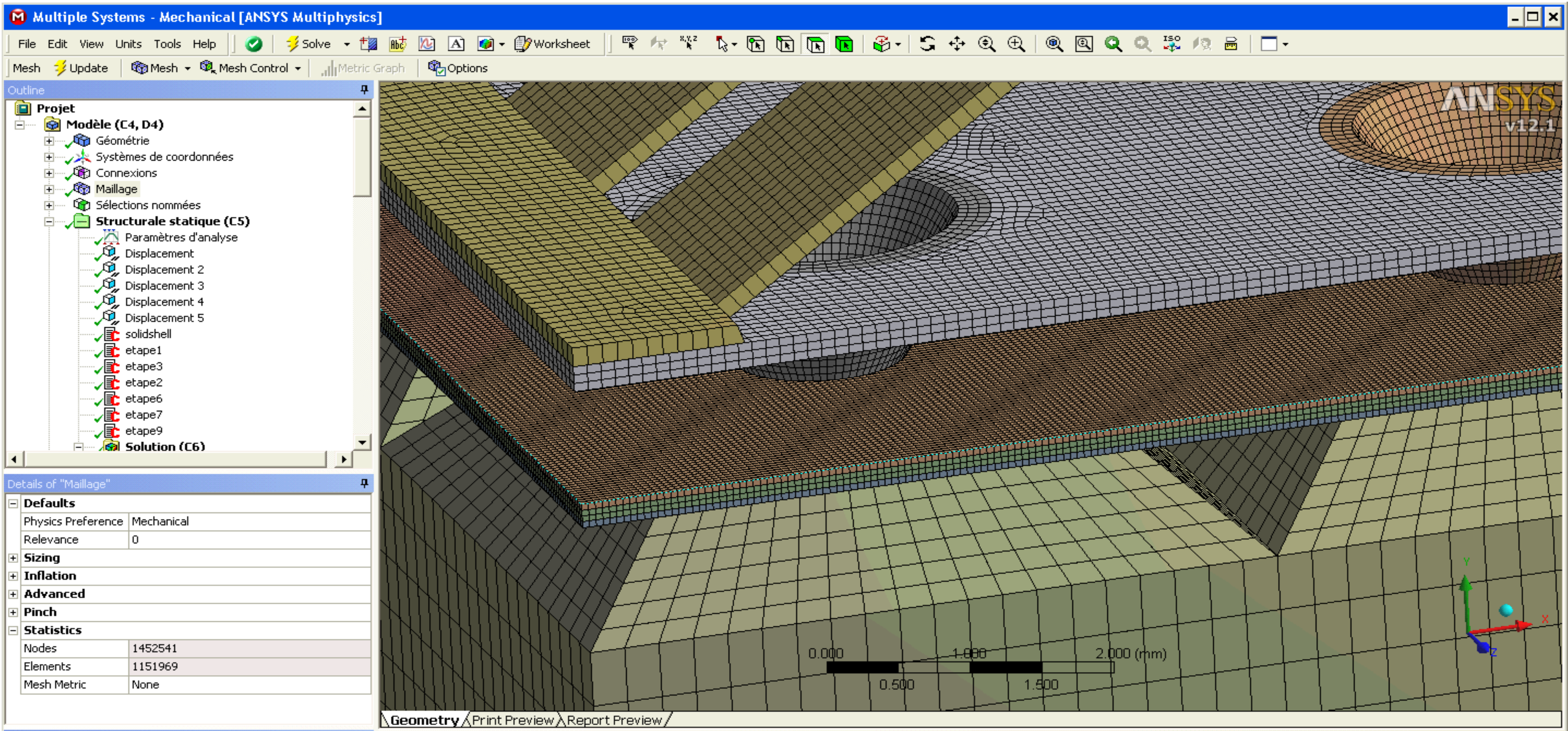
- The entire SOLVE phase is parallel
  - More computations performed in parallel → faster solution time
- Better speedups than SMP
  - Can achieve > 4x on 8 cores (**Try getting that with SMP!!!!**)
  - Can be used for jobs running on up to 1024 cores
- Can take advantage of resources on multiple machines
  - ***Whole new class of problems can be solved!***
  - Memory usage and bandwidth scales
  - Disk (I/O) usage scales (**i.e. parallel I/O**)

Distributed ANSYS

- Enhanced scalability
  - Parallel equation ordering scheme is now default for sparse solver



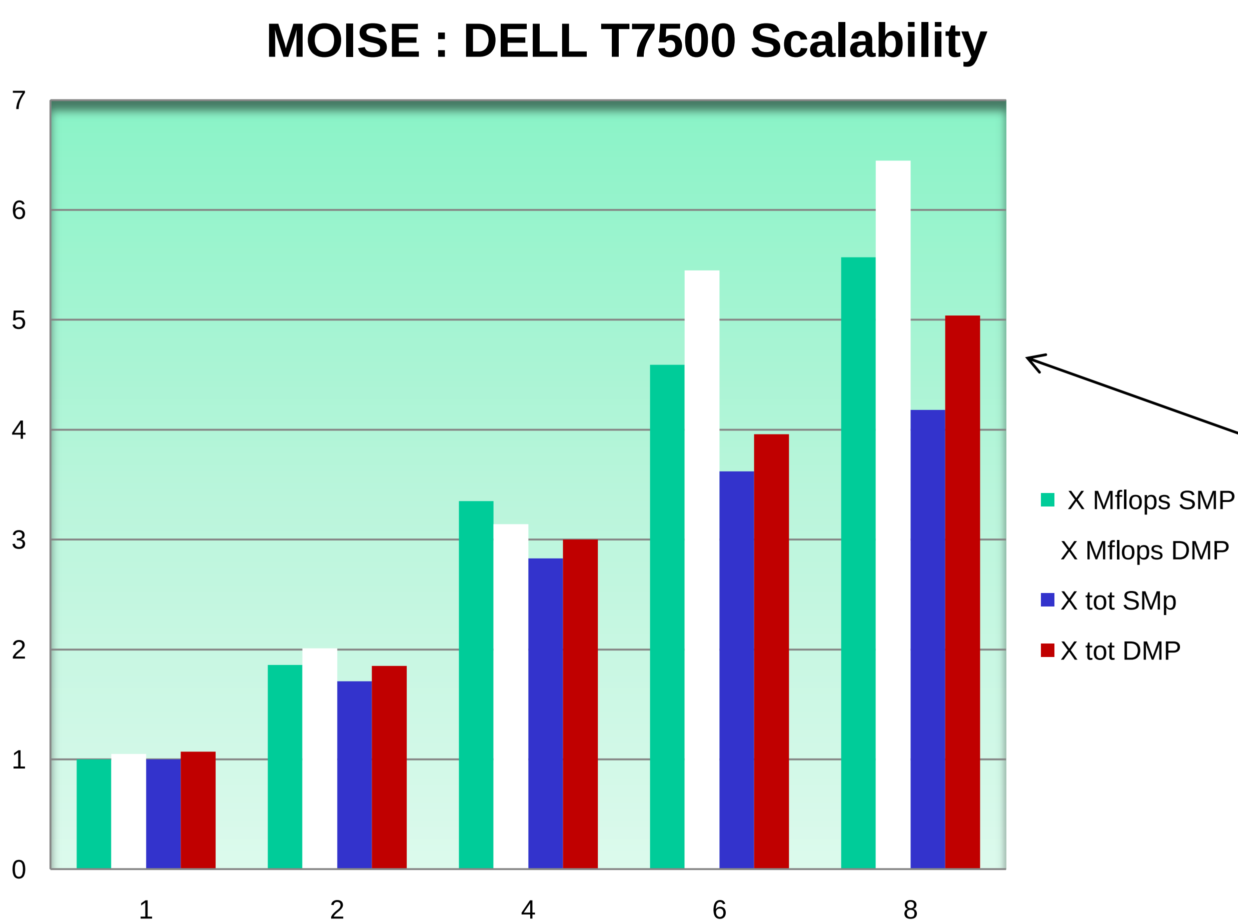
HPC start on Desktop : ANR Project MOISE



4.3 Mdof , Large displacement, plasticity, creep, (> 100 000 frictional contact element) 116 iter cum en 48 heures Dell T7500 8 cœurs w5580

ANSYS speedup on Desktop
(01/2010 ANR Project MOISE)

- Gain de temps en multi cœur sur un PC Desktop :



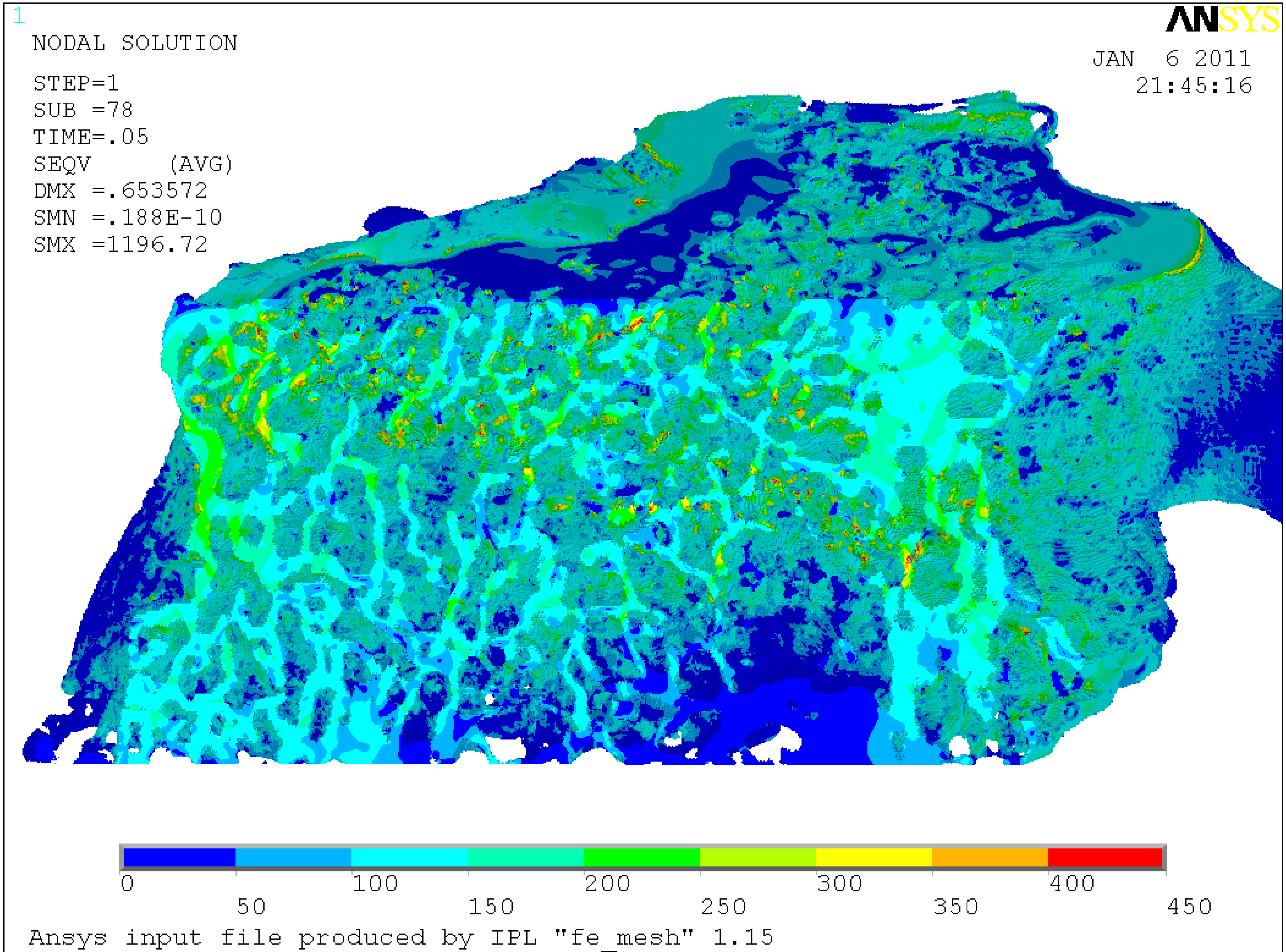
1.36 Mdof ,
Grand déplacement ,
plasticité,
fluage,
contact avec frottement

46 iter cum :

<u>accélération 5 x avec 8 cœurs</u>

(calculs réalisés en 1/2010)

VPHOP Project : Bone & Osteoporosis

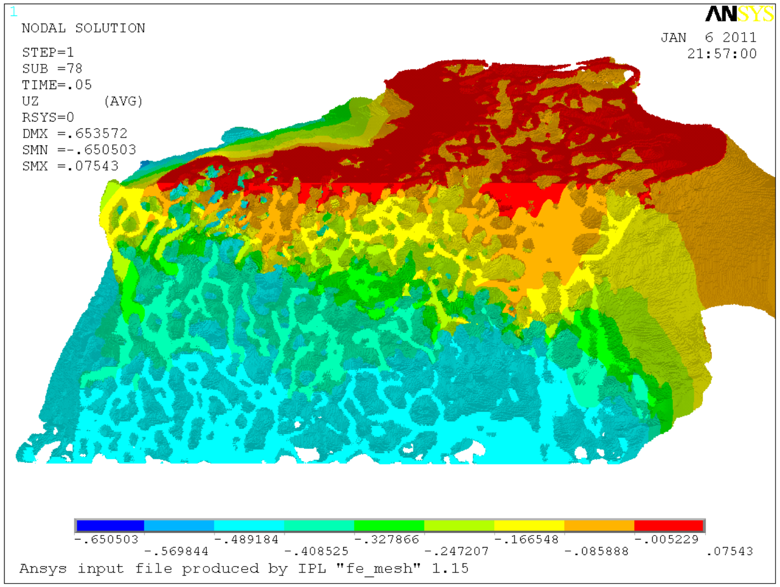


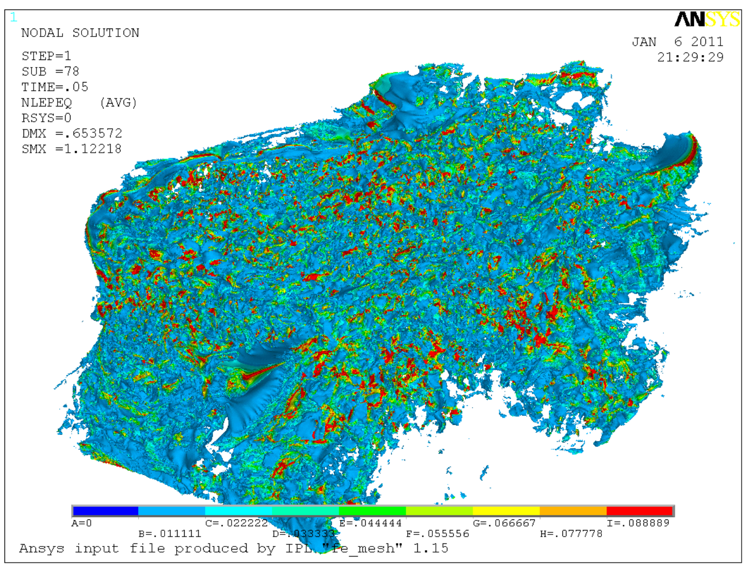
40 Mdof,
Plasticity,
Large
Deformation,

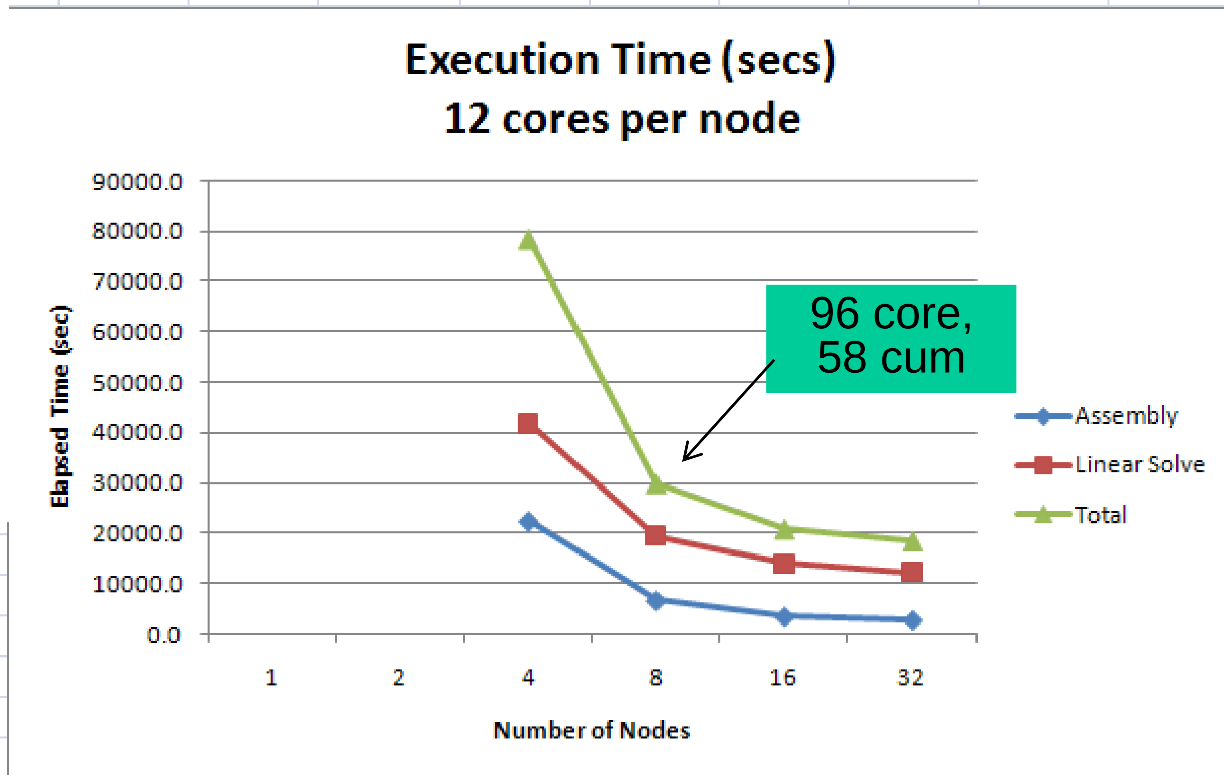
Cumulated plastic
Deformation 112%

Cumulative
Iteration number
340

VPHOP Project : Bone & Osteoporosis





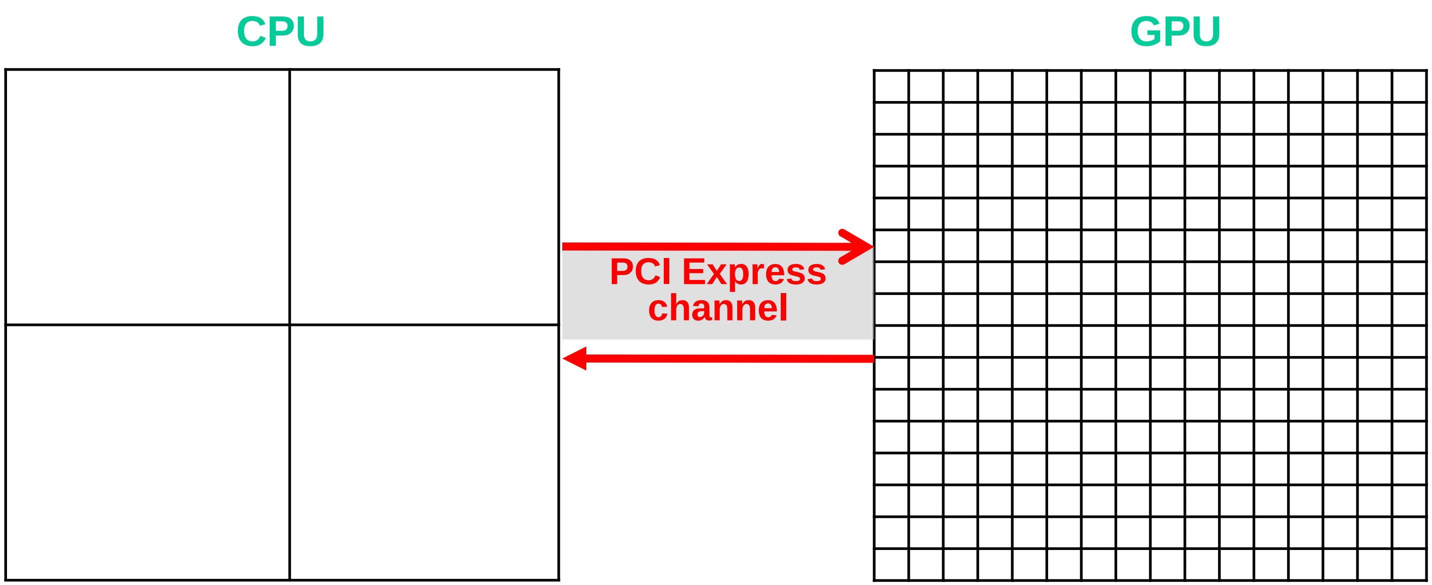


GPU : Introduction

- Graphics processing units (GPUs)
  - Widely used for gaming, graphics rendering
  - Recently been made available as general-purpose "accelerators"
    - Support for double precision arithmetic
    - Performance exceeding the latest multicore CPUs

  - **So how can ANSYS Mechanical make use of this new technology to reduce the overall time to solution??**

- CPUs and GPUs used in a ***collaborative*** fashion



- Typically 4-12 cores
- Powerful, general purpose

- Typically hundreds of cores
- Great for highly parallel code

GPU Accelerator capability

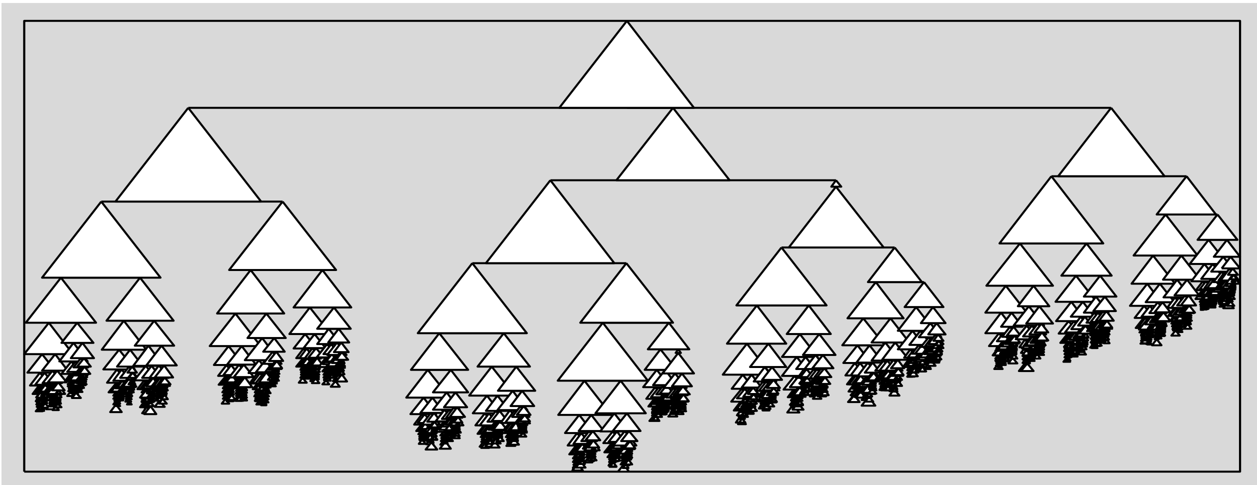
- Motivation
  - Equation solver dominates solution time
    - Logical place to add GPU acceleration

5%-30% time	Element Formation
5%-10% time	Solution Procedures
5%-10% time	Global Assembly
60%-90% time	Equation Solver (e.g., [A]{x} = {b})
1%-10% time	Element Stress Recovery

"solver"

GPU Accelerator capability

- "Accelerate" sparse direct solver (Boeing/DSP)
  - GPU is only used to factor a dense frontal matrix
  - Decision is made based on frontal matrix size on when to send data to GPU or not:
    - Too small, too much overhead, stays on CPU
    - Too large, exceeds GPU memory, stays on CPU



GPU Accelerator capability

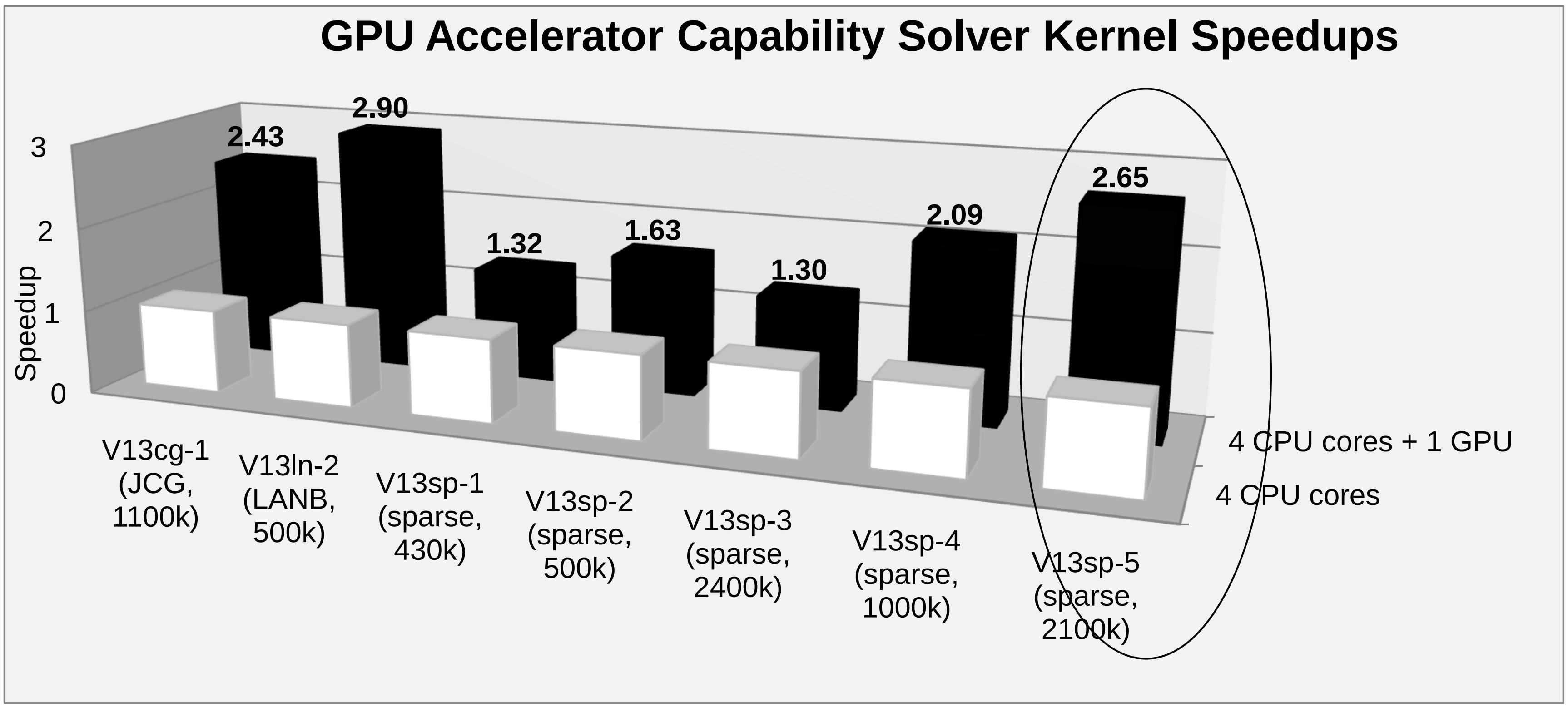
- Supported hardware
  - Currently recommending NVIDIA Tesla 20-series cards
  - Recently added support for Quadro 6000
  - Requires the following items
    - Larger power supply (1 card needs about 225W)
    - Open 2x form factor PCIe x16 Gen2 slot
  - Supported on Windows/Linux 64-bit

	NVIDIA Tesla C2050	NVIDIA Tesla C2070	NVIDIA Quadro 6000
Power	225 Watts	225 Watts	225 Watts
CUDA cores	448	448	448
Memory	3 GB	6 GB	6 GB
Memory Bandwidth	144 GB/s	144 GB/s	144 GB/s
Peak Speed (SP/DP)	1030/515 Gflops	1030/515 Gflops	1030/515 Gflops

Performance Results

- R13.0 benchmark set
- Solver kernel performance

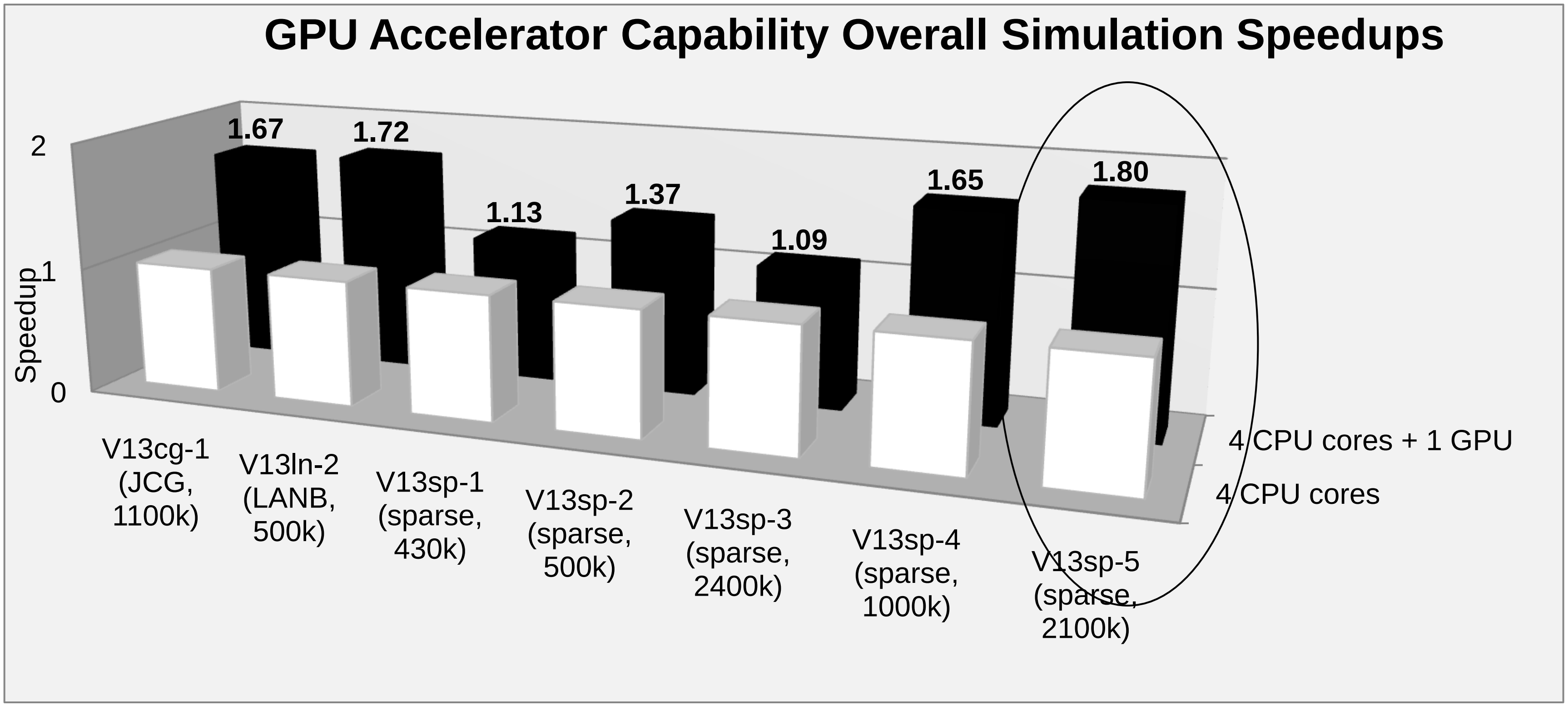
- Intel Xeon 5560 processors (2.8 GHz, 8 cores total)
- 32 GB of RAM
- Windows XP SP2 (64-bit)
- Tesla C2050 (ECC,ON; WDDM driver)



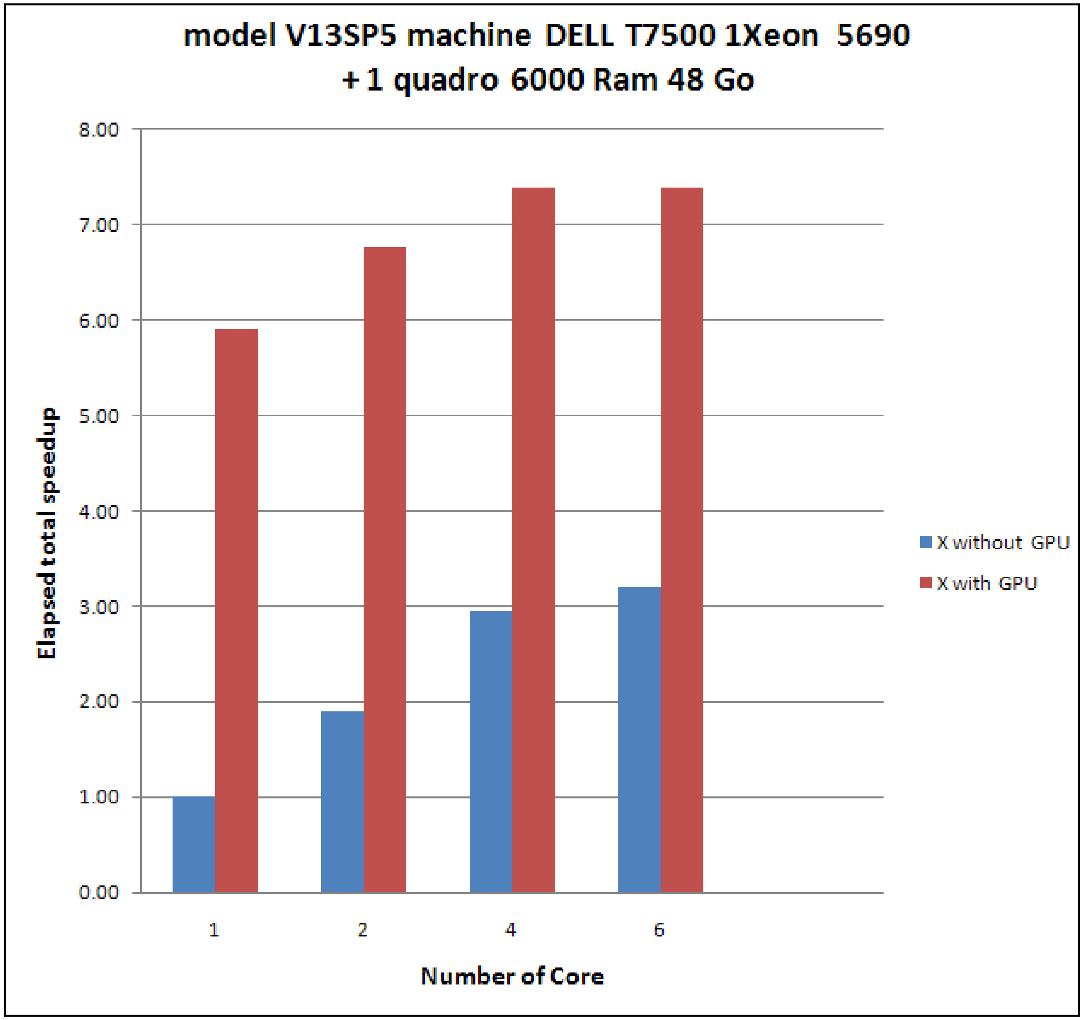
Performance Results

- R13.0 benchmark set
- Total simulation performance

- Intel Xeon 5560 processors (2.8 GHz, 8 cores total)
- 32 GB of RAM
- Windows XP SP2 (64-bit)
- Tesla C2050 (ECC,ON; WDDM driver)

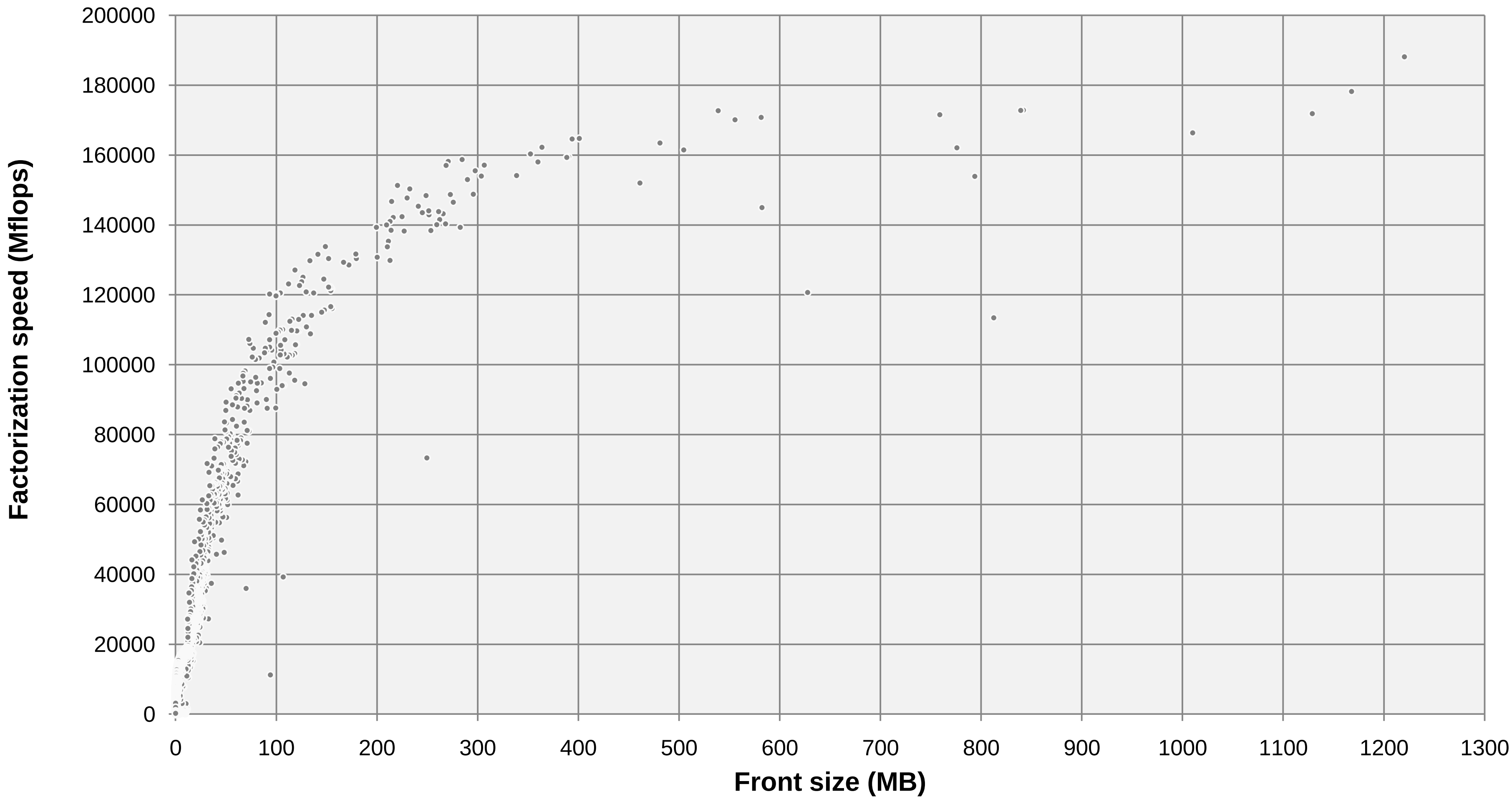


Latest Results for V13SP5 2.1Mdof Benchmark



GPU Accelerator capability

- V13sp-5 benchmark (turbine model)



Which hardware for HPC ?

lyohpc_SMP.txt - Notepad

File Edit Format View Help

SERVER / CLUSTER

ANSYS Mechanical APDL Version 13.0 Benchmark Runs Summary

```
                 Company: DELL
          Computer Model: R910
         Processor Model: X7560
         Total Num Nodes: 1
         Total Num Cores: 32
               GPU Model: none
Total system memory (GB): 256 Go
      Disk Configuration: sas 6 x 10000 kpm raid 0
        MPI Interconnect: 10 gb
        Operating System: windows2008r2
        MPI Version Used: native
```

Job Name	Machlist N X core	Num nodes	Num cores	Total (sec)	Prep Solve	Core Solver	Post Solve	TotMem (GB)	Master (GB)	SlvMax (GB)	Solver Mflops	I/O Mb/sec	Comm Mb/sec	Latency (usec)
v13sp-5	1X1	1	1	3888	13	3850	0	35.3	-	-	5493	3858	-	-
v13sp-5	1X2	1	2	2610	10	2576	0	35.3	-	-	8316	6628	-	-
v13sp-5	1X4	1	4	1752	10	1717	0	35.3	-	-	12879	9008	-	-
v13sp-5	1X8	1	8	1213	10	1180	0	35.3	-	-	19759	7329	-	-
v13sp-5	1X16	1	16	879	9	848	0	35.3	-	-	29474	6164	-	-
v13sp-5	1X32	1	32	578	9	543	0	35.3	-	-	54347	6507	-	-

T7500_SMP.txt - Notepad

File Edit Format View Help

DESKTOP

ANSYS Mechanical APDL Version 13.0 Benchmark Runs Summary

```
                 Company: DELL
          Computer Model: T7500
         Processor Model: X5690
         Total Num Nodes: 1
         Total Num Cores: 6
               GPU Model: Quadro 6000
Total system memory (GB): 48 Go
      Disk Configuration: sas 2 x 15 0000 kpm raid 0
        MPI Interconnect: 1 gb
        Operating System: windows 7
        MPI Version Used: native
```

Job Name	Machlist N X core	Num nodes	Num cores	Total (sec)	Prep Solve	Core Solver	Post Solve	TotMem (GB)	Master (GB)	SlvMax (GB)	Solver Mflops	I/O Mb/sec	Comm Mb/sec	Latency (usec)
v13sp-5	1X1	1	1	419	4	405	0	35.3	-	-	78330	8676	-	-
v13sp-5	1X2	1	2	367	4	353	0	35.3	-	-	85061	11663	-	-
v13sp-5	1X4	1	4	335	3	321	0	35.3	-	-	90438	12382	-	-
v13sp-5	1X6	1	6	332	4	318	0	35.3	-	-	90367	11653	-	-

Guidelines : Which hardware for HPC ?

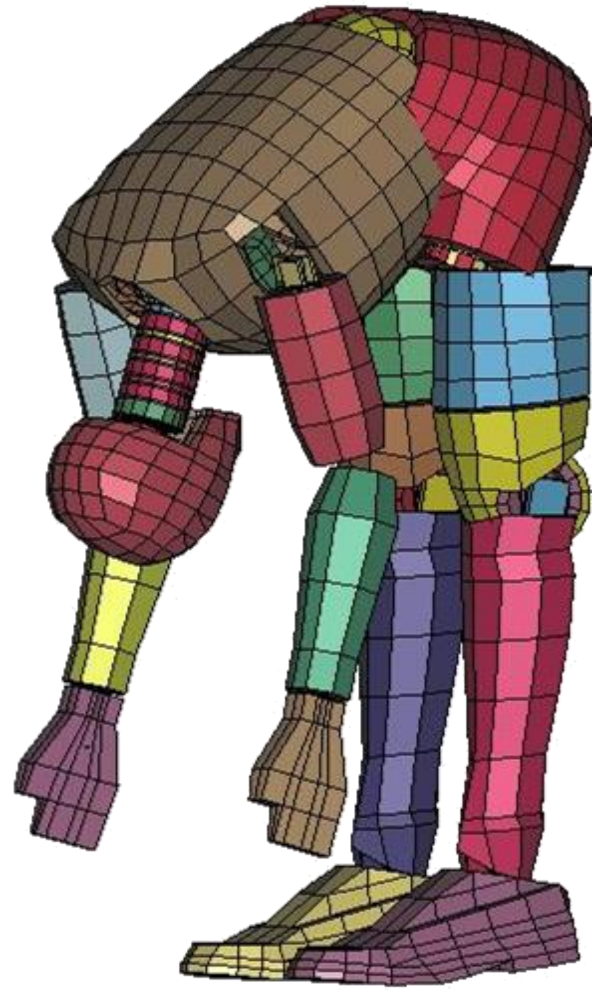
- For 1 or 2 Users :
  - Multicore Desktop (6 to 12 core) + GPU
  - (RAM 48 go , 2 or more Hard disk SAS 15 kpm RAID 0)
- For 4 users :
  - Multisocket server (4 socket , 32 core to 40 core)
  - (RAM 256 Go , 8 or more hard disk SAS 15 kpm RAID 0)
- For More users :
  - Cluster with multiple nodes having Desktop configuration
  - Cluster with some nodes having Server configuration

Summary

- HPC & ANSYS Structural simulation is a reality Today
  - Multicore Provide speedup (Up to 5X on Desktop)
  - GPU if present give additionnal power & decrease core load
  - Server & Cluster can achieve better on larger Models
  - Usage of 6 or 8 core and a GPU is highly recommended for most of structurals problems

Questions ?

Thank You!



Pierre LOUAT

Pierre.Louat@ansys.com