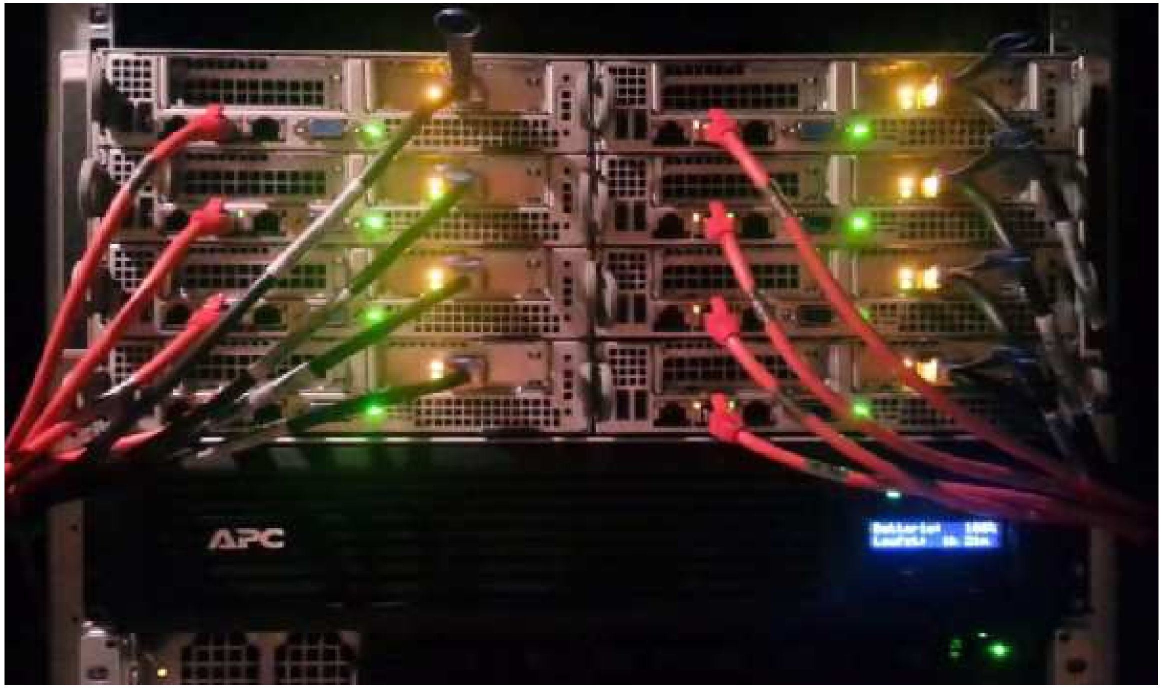


TFLOP Performance for ANSYS Mechanical

Dr. Herbert Güttler

MicroConsult Engineering GmbH
Holunderweg 8
89182 Bernstadt
www.microconsult-engineering.de

MicroConsult Engineering

ANSYS CONFERENCE &
31. CADFEM USERS' MEETING
19. - 21. Juni 2013, Mannheim

May 2009, Ansys12, 512 cores, 1 TFLOP per secoond

ANSYS 12.0 High-Performance Computing Capabilities Mean More Productive Use of Engineering Simulation

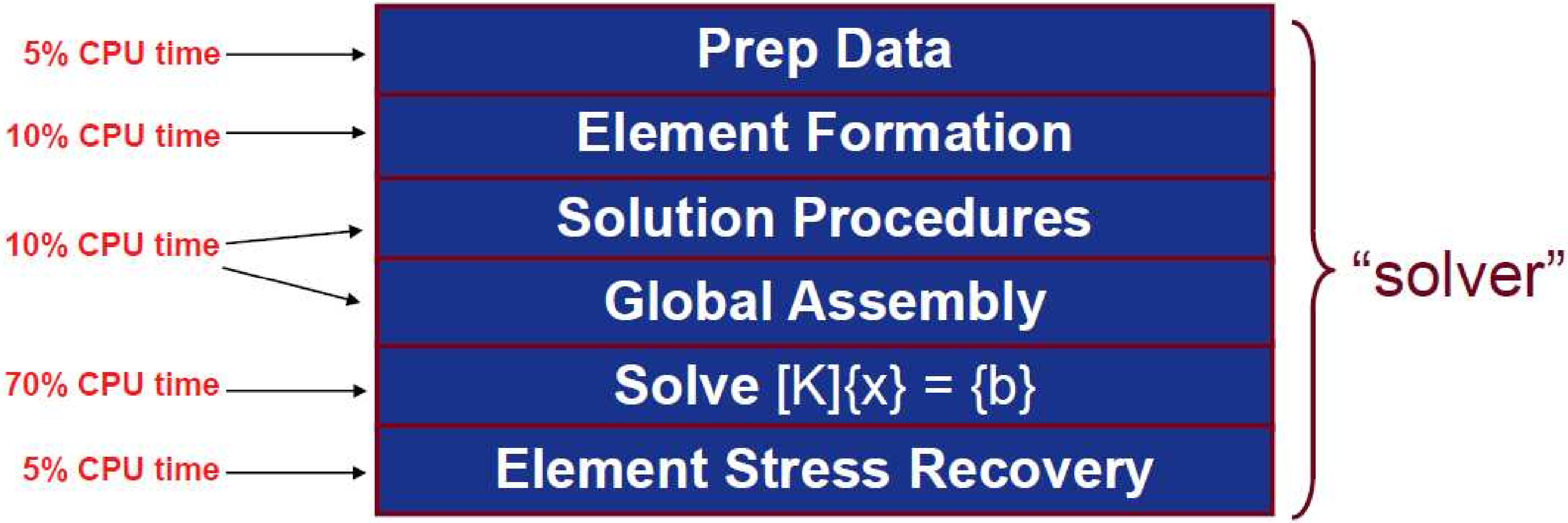
Major Advances in New Release Deliver Teraflop Performance

SOUTHPOINTE, Pa.--(BUSINESS WIRE)--May. 11, 2009-- ANSYS, Inc. (NASDAQ: ANSS), a global innovator of simulation software and technologies designed to optimize product development processes, today announced significant high-performance computing (HPC) milestones achieved with ANSYS® 12.0. The new release delivers impressive performance gains that enable product development teams to increase the value of simulation by considering large, high-fidelity models in shorter turnaround times. Key HPC achievements in the recent release include optimized parallel computing performance on multi-core processors, expanded support for large simulations, scaling breakthroughs, and support for parallel file systems. These product enhancements deliver best-in-class HPC capability for multiphysics simulations, engineered to scale from multi-core desktop workstations to departmental clusters and large enterprise supercomputers.

"Engineers today are including more geometric details in their analyses and looking for a more-realistic treatment of physical phenomena. Both of these factors drive the need for outstanding performance on the latest computing platforms," said Jim Cashman, president and CEO of ANSYS, Inc. "ANSYS 12.0 is our most HPC-capable release to date, and it delivers tremendous value to customers who need faster turnaround and the ability to consider high-fidelity multiphysics simulations. HPC is an important part of Simulation Driven Product Development™, since it delivers efficiency and productivity improvements to engineering organizations seeking to gain a competitive advantage in today's business climate."

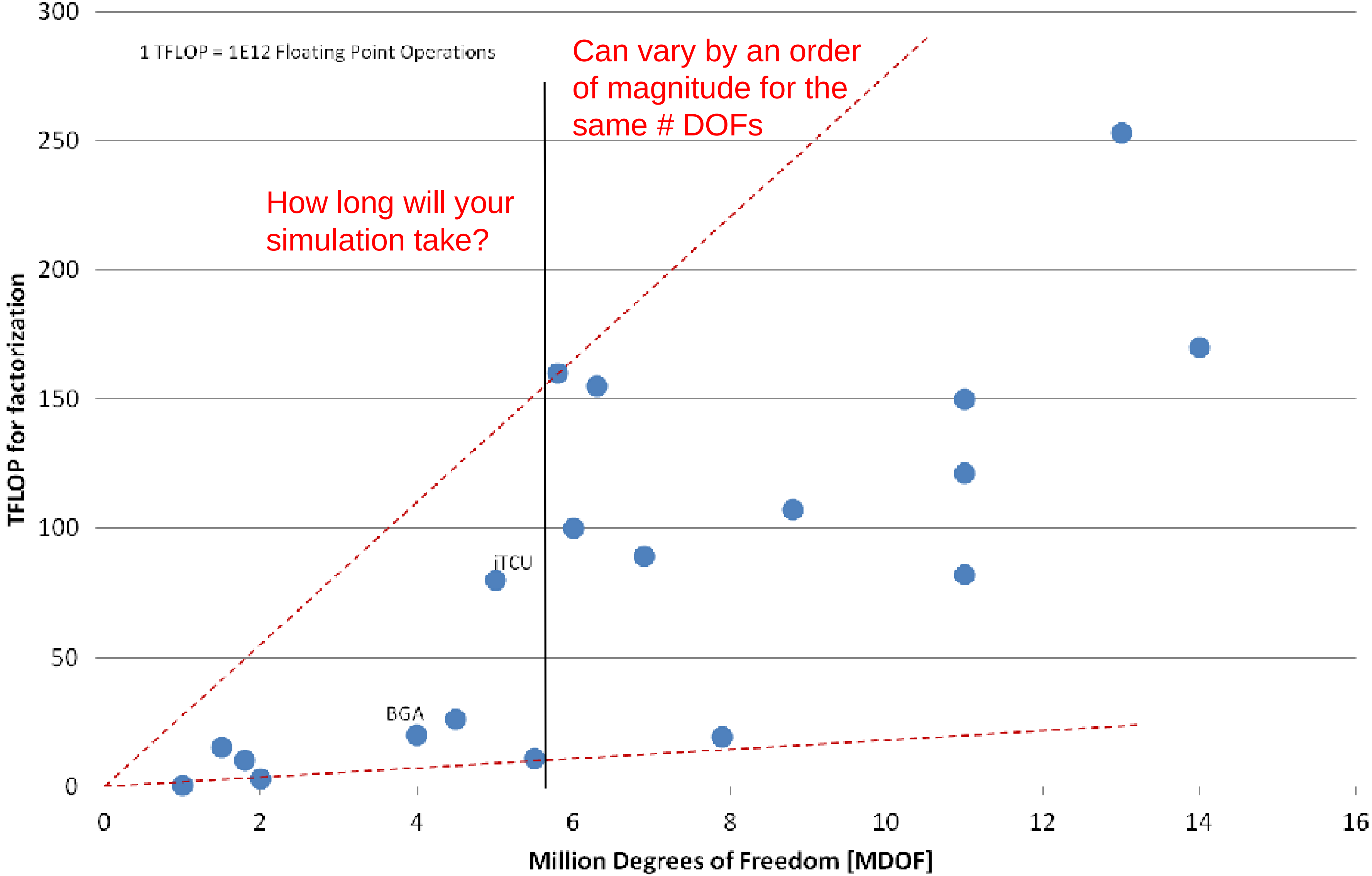
ANSYS 12.0 technology incorporates optimization for the latest multi-core processors and benefits greatly from recent improvements in processor architecture, resulting in highly efficient use of parallel processing to reduce the turnaround time for simulation. Improved algorithms for model partitioning, combined with optimized communications and load balancing between processors, have yielded additional parallel scaling breakthroughs. For structural mechanics analyses, dramatically improved scaling is observed on desktop parallel systems, and a major performance milestone of over 1 teraflop has been achieved via cluster computing on 512 cores. Teraflop performance is well over 100 times faster than the fastest single-core performance currently observed, reducing run times from days to minutes on the most challenging simulations. For fluid dynamics simulations, nearly ideal linear speedup has been demonstrated out to 1,024 cores, roughly doubling the core count for ideal scaling compared to previous releases. At 2,048 cores, scaling remained at approximately 80 percent of ideal linear performance. ANSYS FLUENT® 12.0 software introduces parallel input and output (I/O) of files, dramatically reducing turnaround time for large simulations that involve extensive I/O and removing file handling as a bottleneck for scaling on large clusters. In addition, ANSYS 12.0 technology includes important enhancements that enable larger simulations than ever before, setting the stage for customers to consider highly detailed physical phenomena and full-assembly models in their fluids or structural simulations.

ANSYS Solver Basics: Solution Overview



- Equation solver dominates solution CPU time! Need to pay attention to equation solver
- Equation solver also consumes the most system resources (memory and I/O)

Numerical Effort for a random selection of MCE Projects ANSYS MAPDL, sparse solver

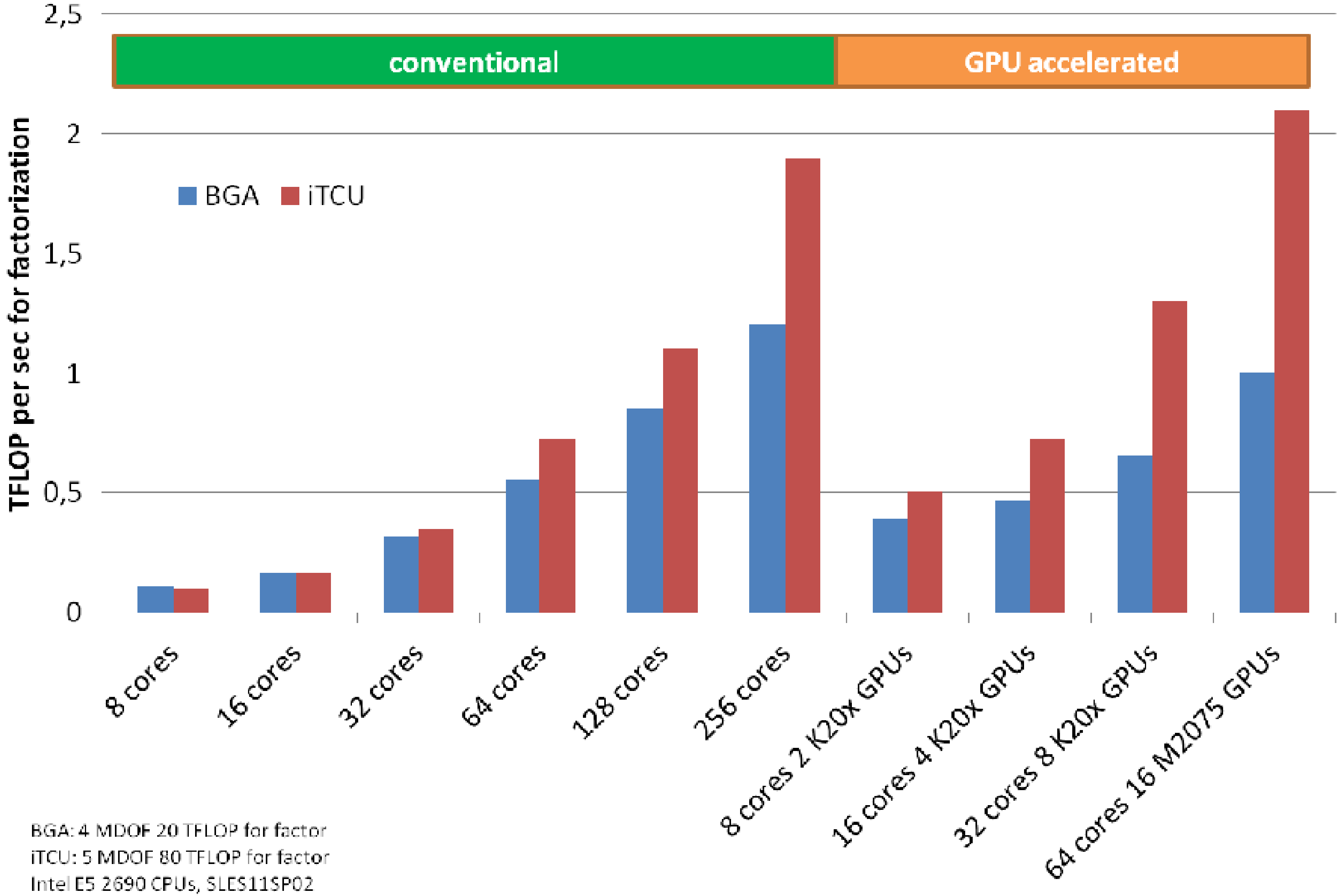


Stats data can be found here

p.ex. file.DSP

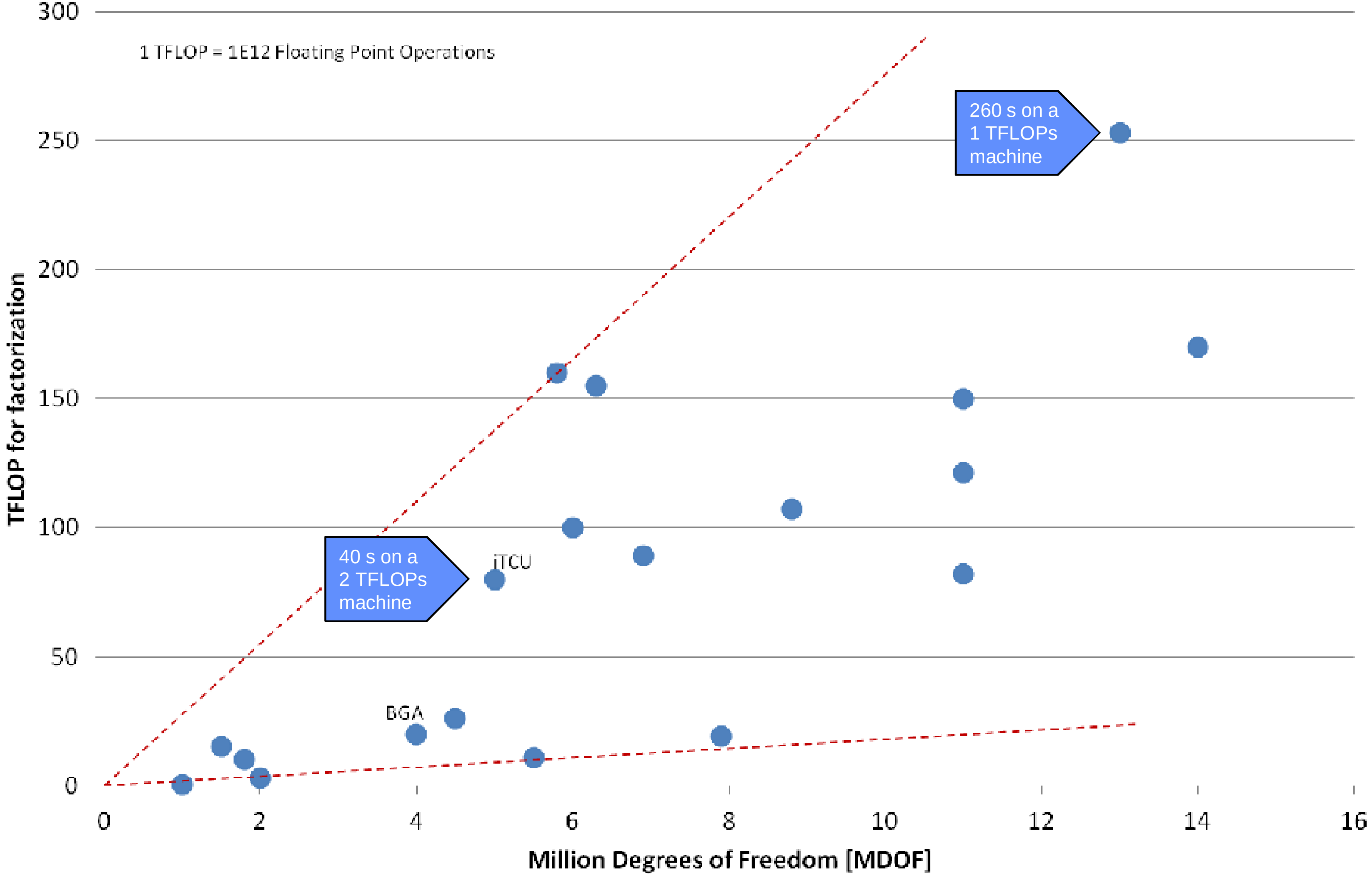
```
==========================
 = multifrontal statistics =
 ==========================

    number of equations                       =        5162979
    no. of nonzeroes in lower triangle of a =      406162561
    no. of nonzeroes in the factor l          =    11877671289
    ratio of nonzeroes in factor (min/max)  =         0.1692
    number of super nodes                     =          64238
    maximum order of a front matrix           =          13914
    maximum size of a front matrix            =       96806655
    maximum size of a front trapezoid         =       59723199
    no. of floating point ops for factor      =     7.8384D+13
    no. of floating point ops for solve       =     4.6973D+10
    ratio of flops for factor (min/max)       =         0.4117
    near zero pivot monitoring activated
    number of pivots adjusted                 =              0
    negative pivot monitoring activated
    number of negative pivots encountered     =              0
    factorization panel size                  =            128
    number of cores used                      =             64
    GPU acceleration activated
    percentage of GPU accelerated flops       =        99.7144
    time (cpu & wall) for structure input     =       1.060000        1.058888
    time (cpu & wall) for ordering            =      14.095459       14.095459
    time (cpu & wall) for other matrix prep =       6.354541        7.541591
    time (cpu & wall) for value input         =       0.950000        0.944036
    time (cpu & wall) for matrix distrib.     =       3.340000        3.326019
    time (cpu & wall) for numeric factor      =      35.900000       37.974798
    computational rate (mflops) for factor  =  2183395.338670  2064102.972591
    time (cpu & wall) for numeric solve       =       0.730000        0.723194
    computational rate (mflops) for solve     =   64346.638293    64952.225467
    effective I/O rate (MB/sec) for solve     =  245160.688215   247467.975314

 Memory allocated on core    0        =    2420.815 MB
 Memory allocated on core    1        =    2283.246 MB
…
 Memory allocated on core   62        =    2758.796 MB
 Memory allocated on core   63        =    3805.647 MB
 Total Memory allocated by all cores  =  156488.815 MB


 DSP Matrix Solver         CPU Time (sec) =         62.870
 DSP Matrix Solver     ELAPSED Time (sec) =         66.165
 DSP Matrix Solver      Memory Used ( MB) =       2420.815
```

Performance Results

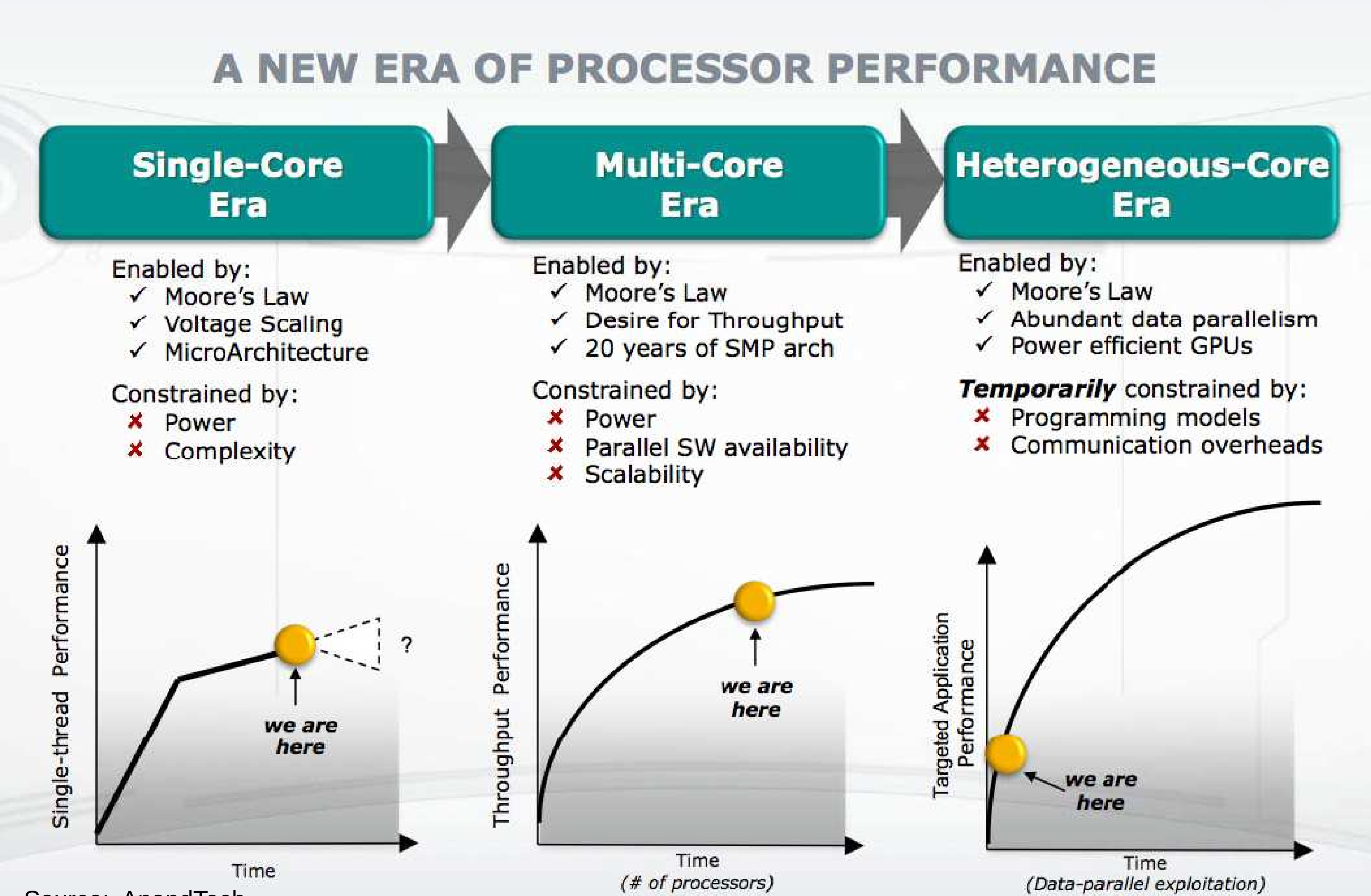


BGA: 4 MDOF 20 TFLOP for factor
iTCU: 5 MDOF 80 TFLOP for factor
Intel E5 2690 CPUs, SLES11SP02

Numerical Effort for a random selection of MCE Projects ANSYS MAPDL, sparse solver

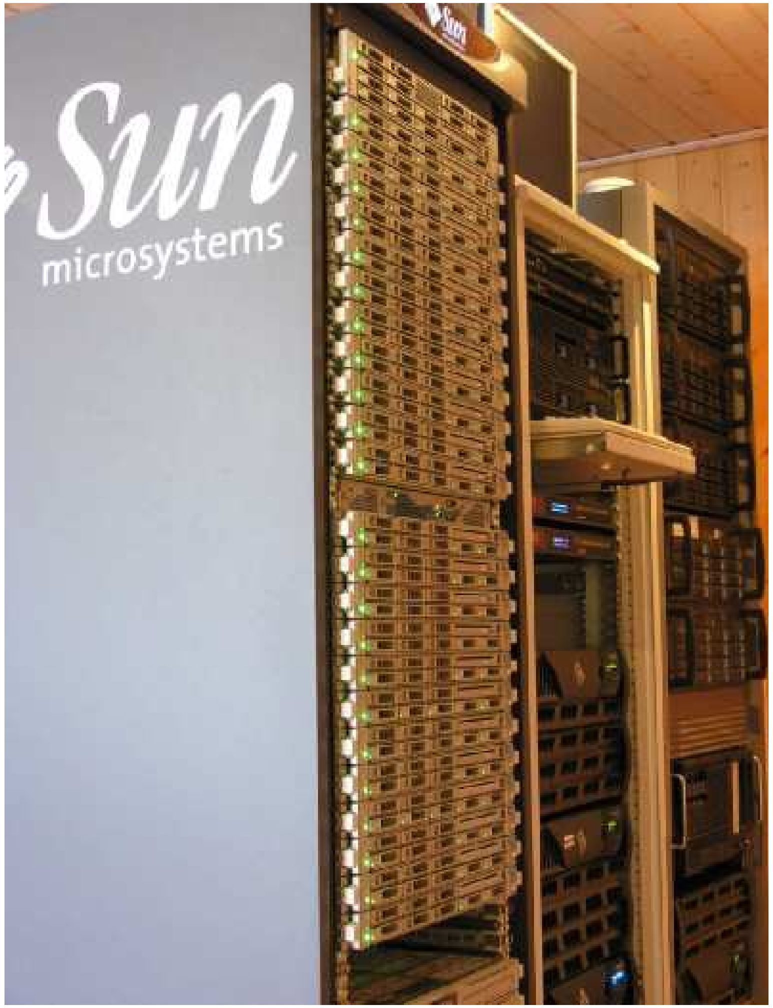


Current status of HPC Computing



Source: AnandTech

Tools (Hardware: Oct 2010)



- Compute Servers
  - 8 Intel Harpertown systems: (SUN X4150) total of 64 cores, 496 GB RAM
  - 16 Intel Nehalem systems: (SUN X4170) total of 128 cores, 1140 GB RAM
  - Memory / core typ. 8GB
  - Infiniband interconnect across servers
  - Each with local Raid 0 disk array
  - Operating System: SUSE Linux Enterprise Server
  - Latest addition: 1 AMD Opteron 6172 System (Magny Cours) 48 cores, 192 GB RAM
  - UPS, Air conditioning
  - Max. power consumption ~ 18kW
- Applications: ANSYS Mechanical, optiSLang

Interconnect: FDR Performance

Latencies	Bandwidth
Latency time from master to core 1 = 1.259 µs	Communication speed from master to core 1 = 8077.06 MB/sec
Latency time from master to core 2 = 1.175 µs	Communication speed from master to core 2 = 8857.00 MB/sec
Latency time from master to core 3 = 1.235 µs	Communication speed from master to core 3 = 9372.93 MB/sec
…	
Latency time from master to core 9 = 2.183 µs	Communication speed from master to core 9 = 5312.38 MB/sec
Latency time from master to core 10 = 2.393 µs	Communication speed from master to core 10 = 5377.34 MB/sec
Latency time from master to core 11 = 1.836 µs	Communication speed from master to core 11 = 5081.82 MB/sec
…	
Latency time from master to core 16 = 1.979 µs	Communication speed from master to core 16 = 5121.90 MB/sec
Latency time from master to core 17 = 2.012 µs	Communication speed from master to core 17 = 5313.56 MB/sec
Latency time from master to core 18 = 2.008 µs	Communication speed from master to core 18 = 5249.56 MB/sec
…	…
Latency time from master to core 28 = 1.993 µs	Communication speed from master to core 28 = 4939.63 MB/sec
Latency time from master to core 29 = 2.366 µs	Communication speed from master to core 29 = 4939.24 MB/sec
Latency time from master to core 30 = 2.333 µs	Communication speed from master to core 30 = 4765.06 MB/sec
Latency time from master to core 31 = 2.119 µs	Communication speed from master to core 31 = 4925.74 MB/sec

core – core on die

socket - socket

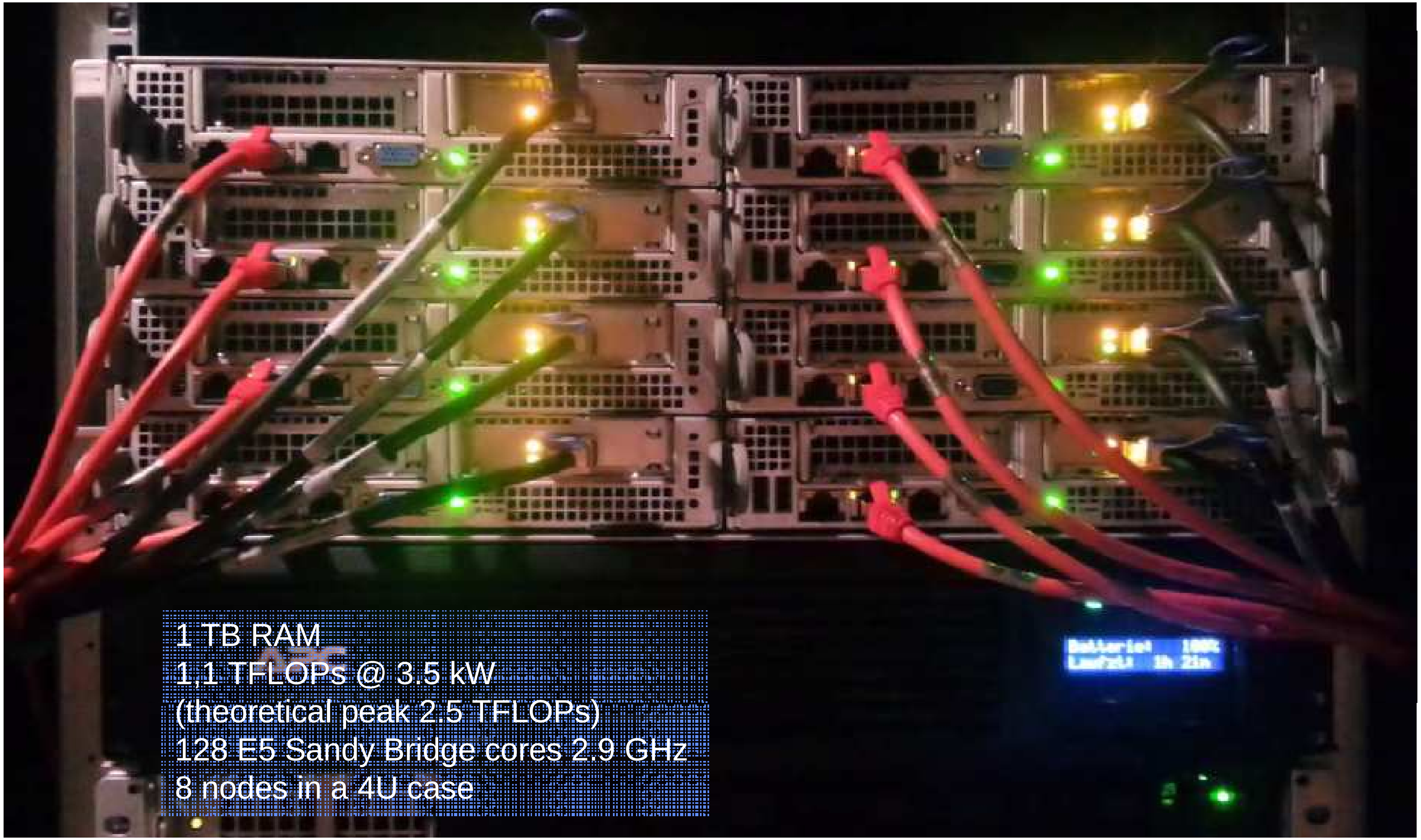
node - node

Tools (Hardware: Jan 2013)

128 E5 Sandy Bridge cores 2.9 GHz
Up to 4 GPUs per node

156 Westmere cores 2.9 GHz
Up to 2 GPUs per node

Tools (Hardware: April 2013)



Tools (Hardware: June 2013)

Comparison for 5 MDOF model (R14.5.7)

w/o GPUs (16x E5 2690)

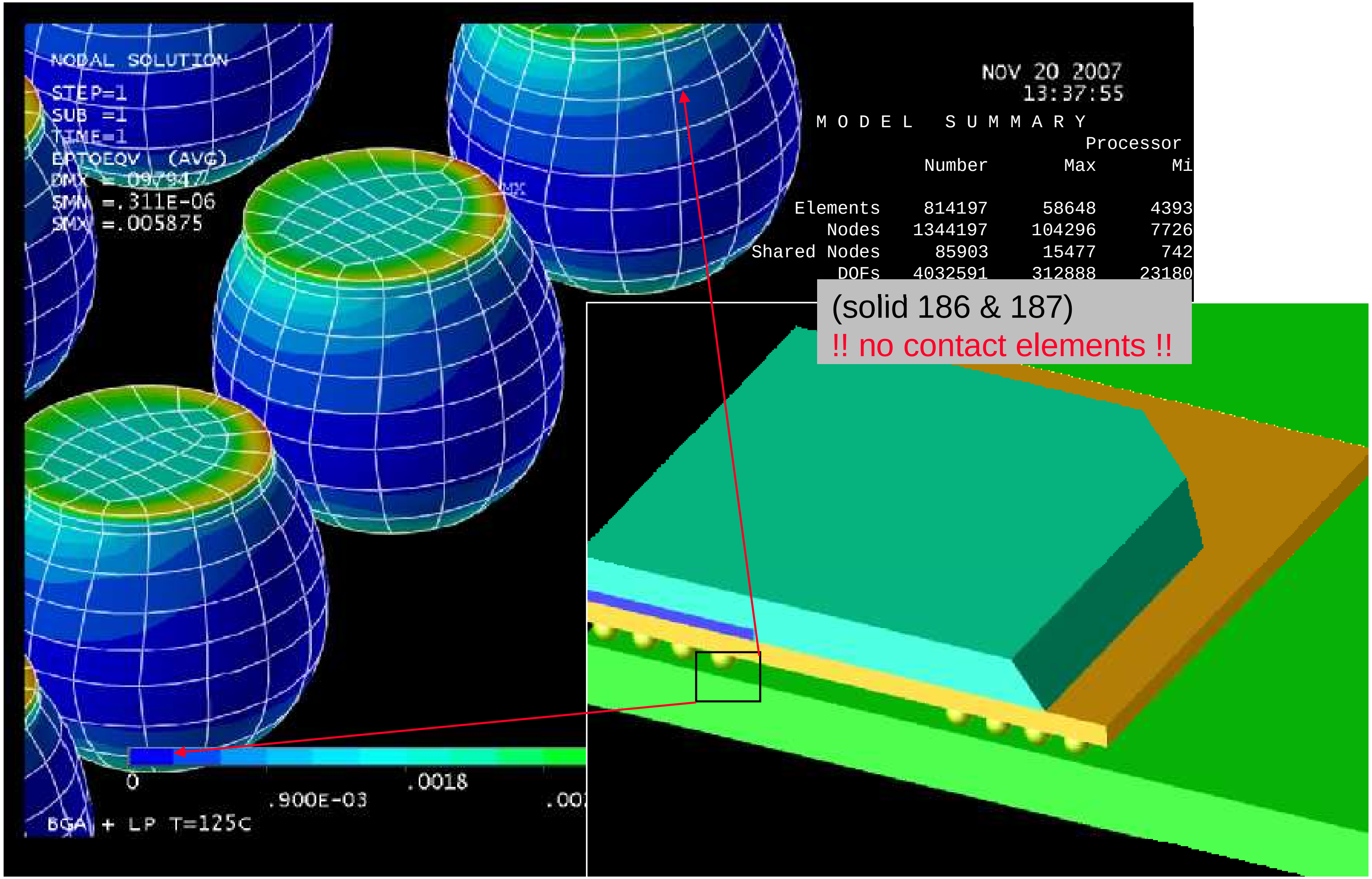
```
==========================
= multifrontal statistics =
==========================

    number of equations                     =         5162979
    no. of nonzeroes in lower triangle of a =       406199623
    no. of nonzeroes in the factor l        =     12023957495
    ratio of nonzeroes in factor (min/max)  =          0.0165
    number of super nodes                   =           64330
    maximum order of a front matrix         =           13764
    maximum size of a front matrix          =        94730730
    maximum size of a front trapezoid       =        62166852
    no. of floating point ops for factor    =      8.3112D+13
    no. of floating point ops for solve     =      4.7458D+10
    ratio of flops for factor (min/max)     =          0.0200
    near zero pivot monitoring activated
    number of pivots adjusted               =               0
    negative pivot monitoring activated
    number of negative pivots encountered   =               0
    factorization panel size                =             128
    number of cores used                    =             128


    time (cpu & wall) for structure input   =        1.080000        1.081304
    time (cpu & wall) for ordering          =       15.326202       15.326202
    time (cpu & wall) for other matrix prep =        7.683798        7.611568
    time (cpu & wall) for value input       =        1.020000        1.011115
    time (cpu & wall) for matrix distrib.   =        3.040000        3.053448
    time (cpu & wall) for numeric factor    =       73.570000       73.376386
    computational rate (mflops) for factor  =  1129695.075130  1132675.943119
    time (cpu & wall) for numeric solve     =        0.760000        0.833381
    computational rate (mflops) for solve   =    62444.779013    56946.350716
    effective I/O rate (MB/sec) for solve   =   237914.604467   216965.592968

 Memory allocated on core    0        =    1587.283 MB
 Memory allocated on core    1        =    1581.346 MB
…
 Memory allocated on core  126        =    1460.529 MB
 Memory allocated on core  127        =    1074.435 MB
 Total Memory allocated by all cores  =  209273.372 MB

 DSP Matrix Solver         CPU Time (sec) =        103.790
 DSP Matrix Solver     ELAPSED Time (sec) =        103.691
 DSP Matrix Solver      Memory Used ( MB) =       1587.283
```

w dual GPUs (4x E5 2690 / 8x Kepler K20x)

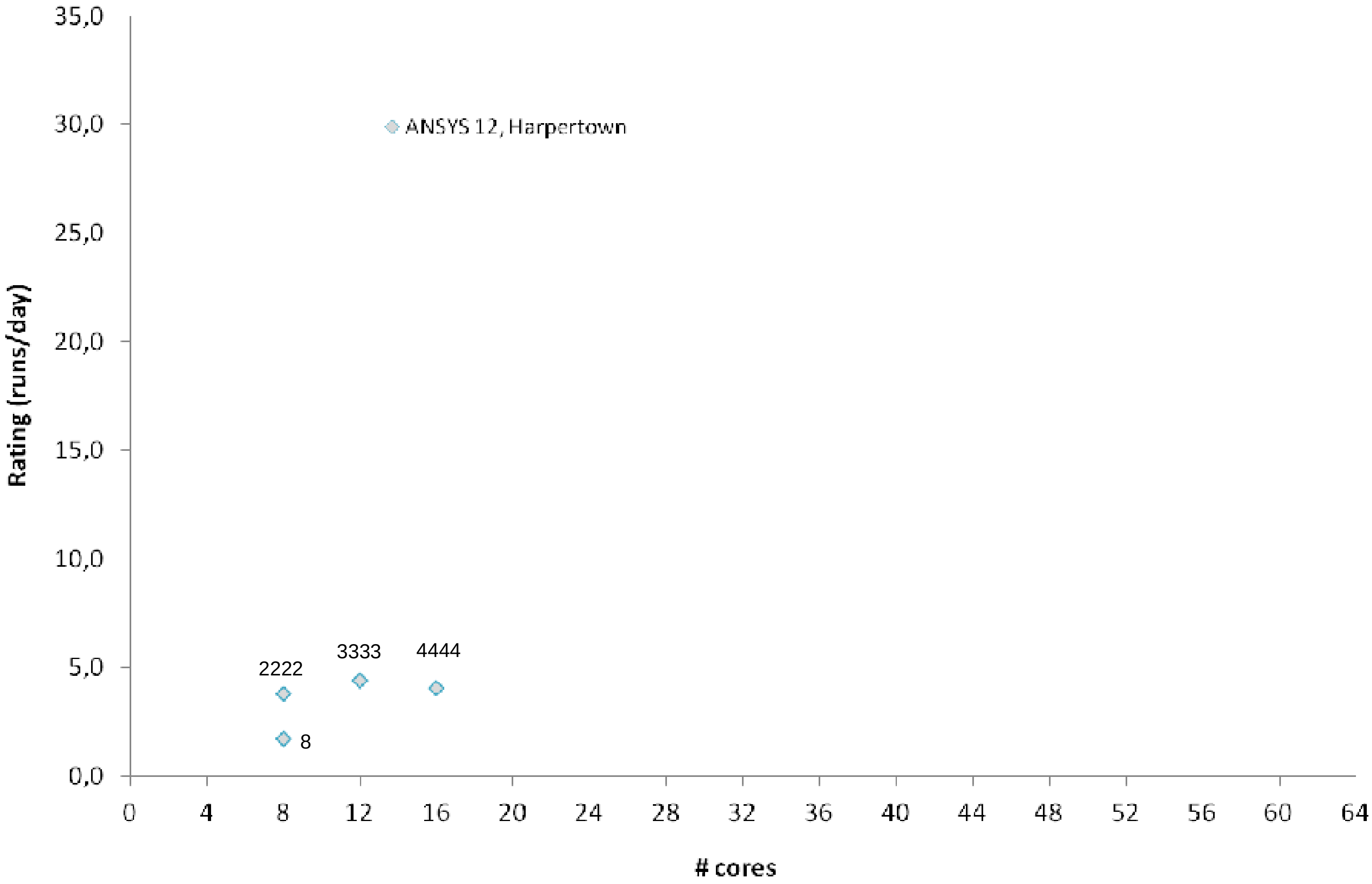
```
==========================
= multifrontal statistics =
==========================

    number of equations                     =         5162979
    no. of nonzeroes in lower triangle of a =       406199677
    no. of nonzeroes in the factor l        =     11901546940
    ratio of nonzeroes in factor (min/max)  =          0.3530
    number of super nodes                   =           64176
    maximum order of a front matrix         =           14307
    maximum size of a front matrix          =       102352278
    maximum size of a front trapezoid       =        65659593
    no. of floating point ops for factor    =      7.9920D+13
    no. of floating point ops for solve     =      4.7105D+10
    ratio of flops for factor (min/max)     =          0.4342
    near zero pivot monitoring activated
    number of pivots adjusted               =               0
    negative pivot monitoring activated
    number of negative pivots encountered   =               0
    factorization panel size                =             128
    number of cores used                    =              32
    GPU acceleration activated
    percentage of GPU accelerated flops     =         98.2176
    time (cpu & wall) for structure input   =        1.590000        1.590859
    time (cpu & wall) for ordering          =       18.430191       18.430191
    time (cpu & wall) for other matrix prep =        8.659809       12.335555
    time (cpu & wall) for value input       =        1.130000        1.130927
    time (cpu & wall) for matrix distrib.   =        9.680000        9.896982
    time (cpu & wall) for numeric factor    =       57.260000       58.358889
    computational rate (mflops) for factor  =  1395732.069696  1369450.643294
    time (cpu & wall) for numeric solve     =        2.560000        2.563528
    computational rate (mflops) for solve   =    18400.567684    18375.247013
    effective I/O rate (MB/sec) for solve   =    70106.161822    70009.690068

 Memory allocated on core    0        =    4952.022 MB
 Memory allocated on core    1        =    4303.400 MB
 ..
 Memory allocated on core   30        =    3476.934 MB
 Memory allocated on core   31        =    5166.779 MB
 Total Memory allocated by all cores  =  153019.380 MB


 DSP Matrix Solver         CPU Time (sec) =        105.440
 DSP Matrix Solver     ELAPSED Time (sec) =        110.549
 DSP Matrix Solver      Memory Used ( MB) =       4952.022
```

Applications

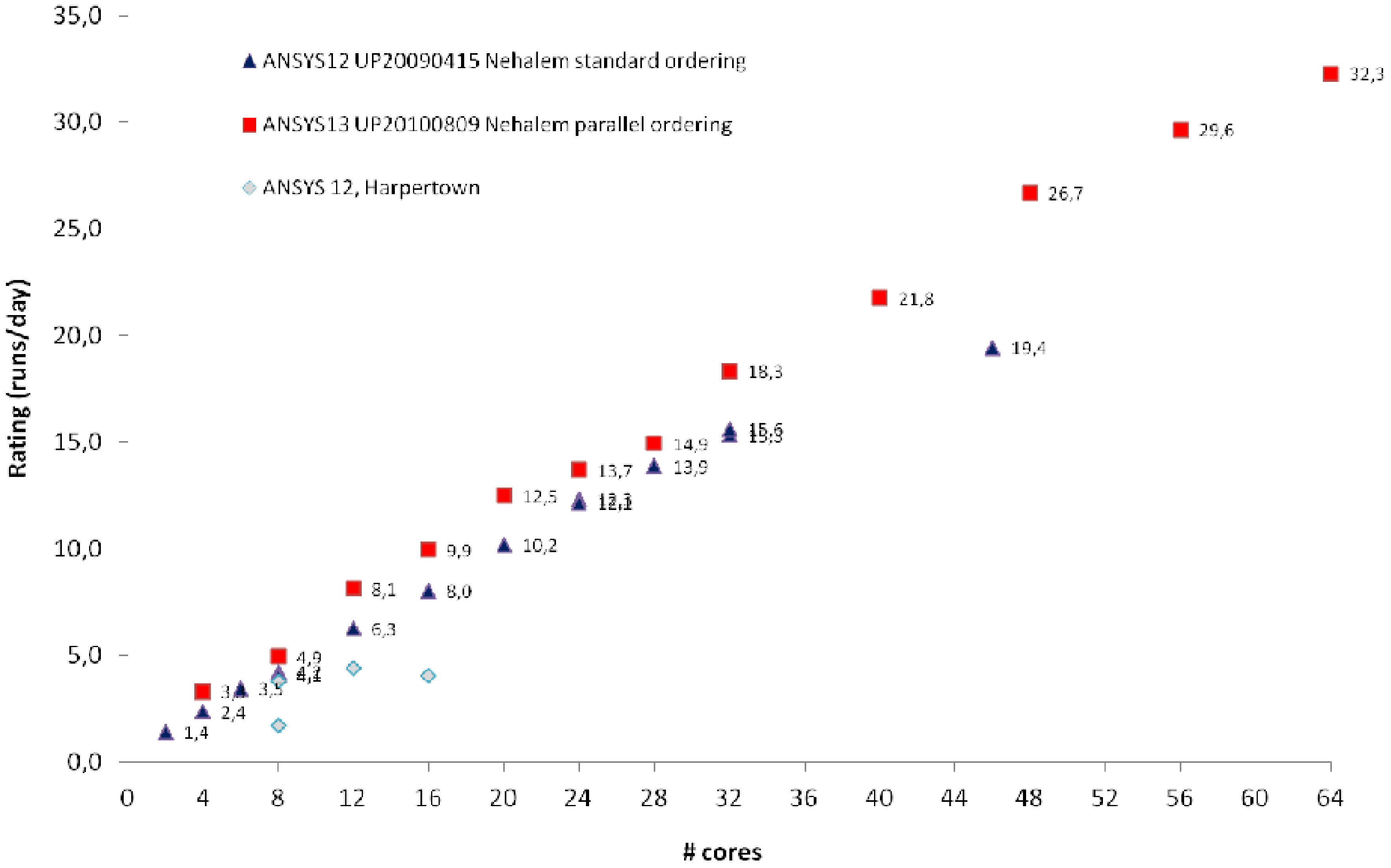
Example: Ball grid array



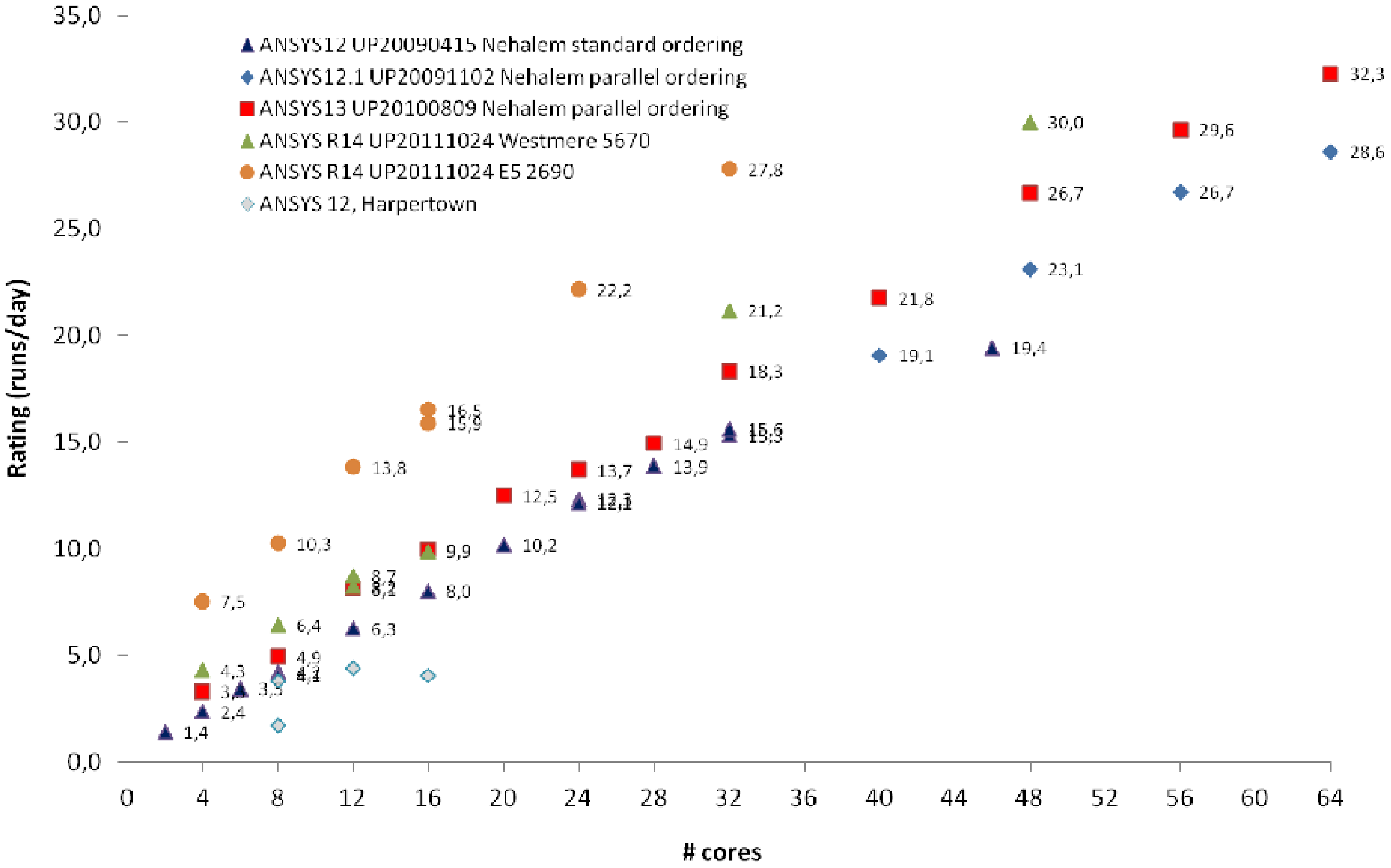
HPC mit ANSYS 14.0



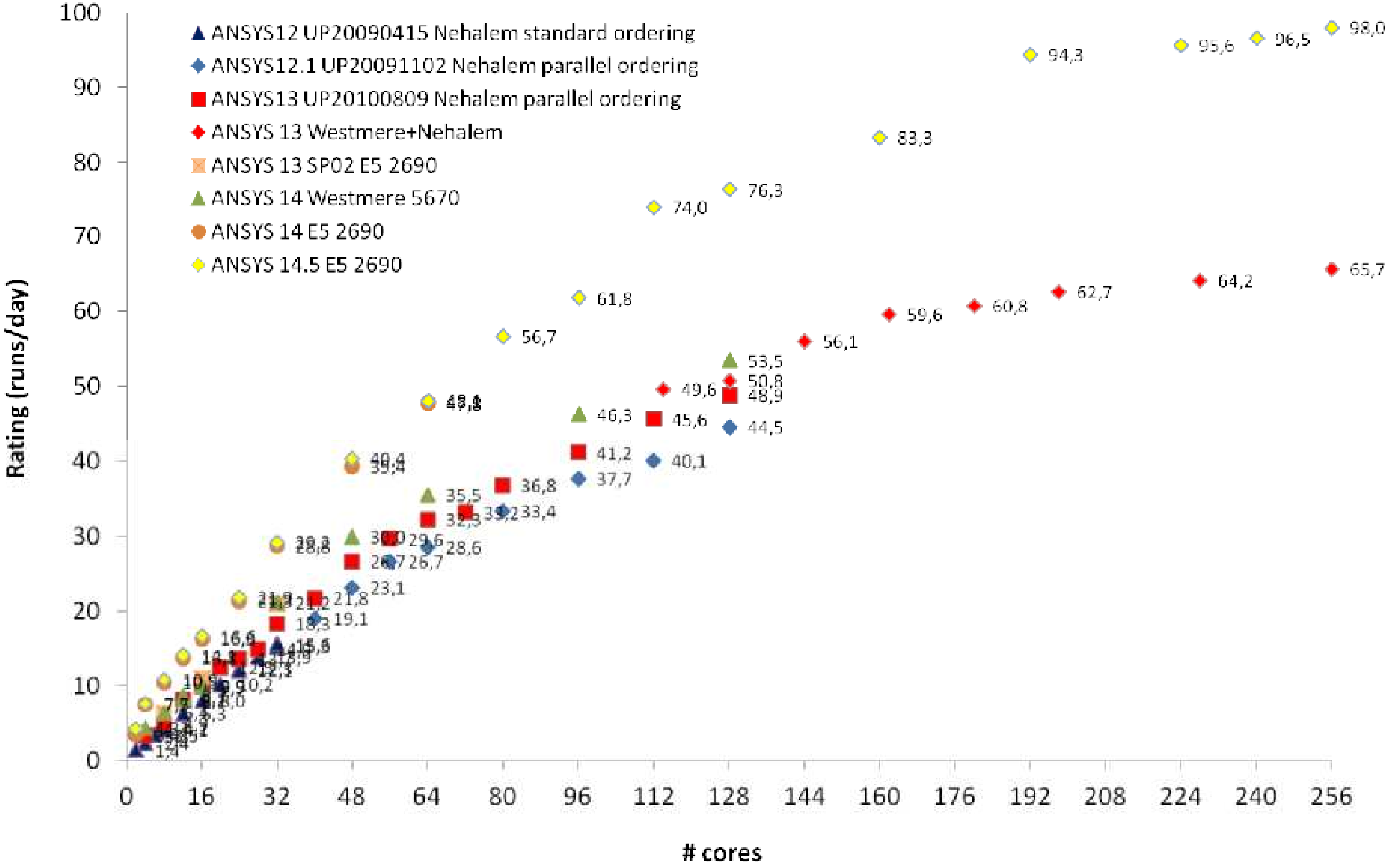
HPC mit ANSYS 14.0



HPC mit ANSYS 14.0



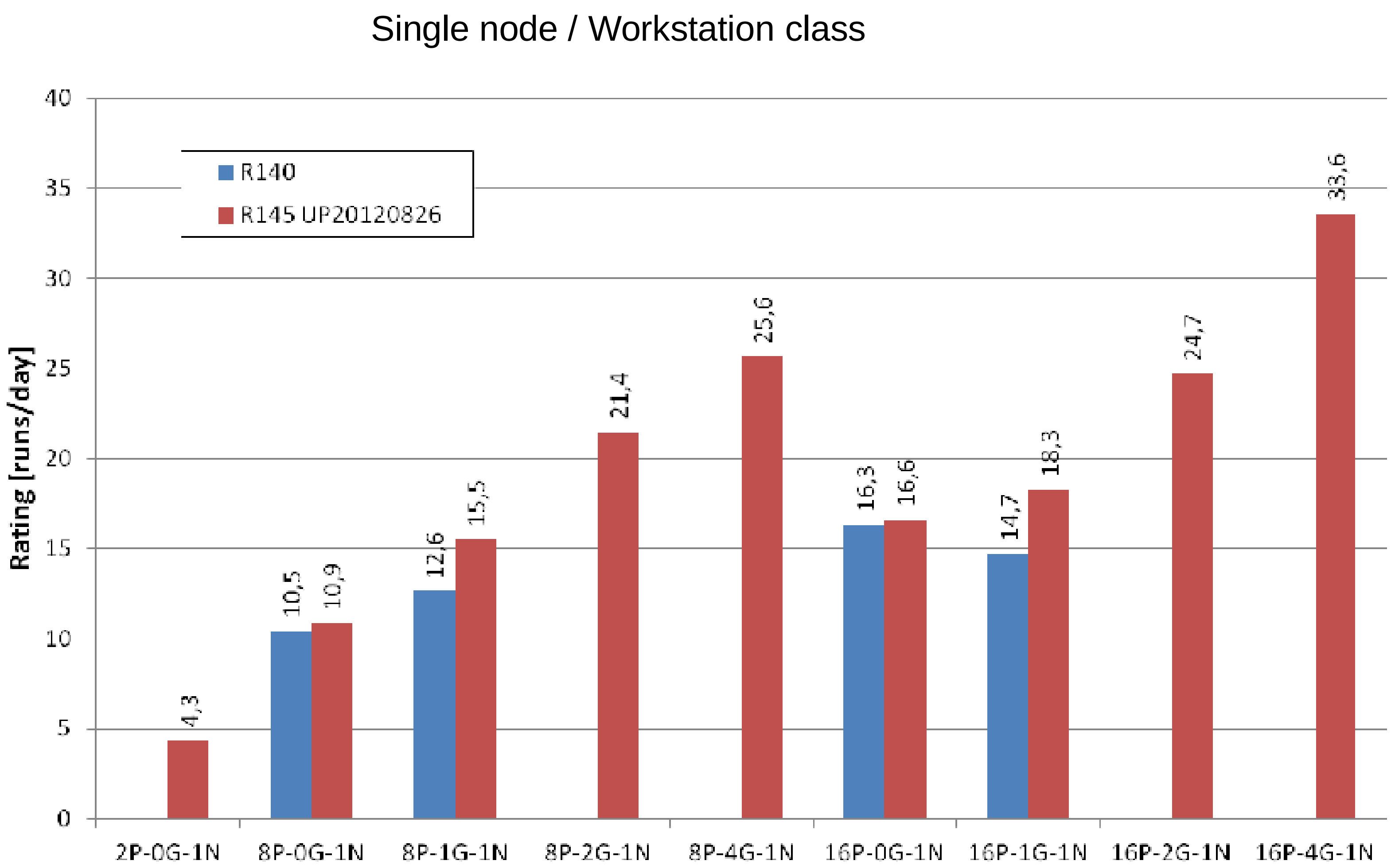
HPC mit ANSYS 14.5



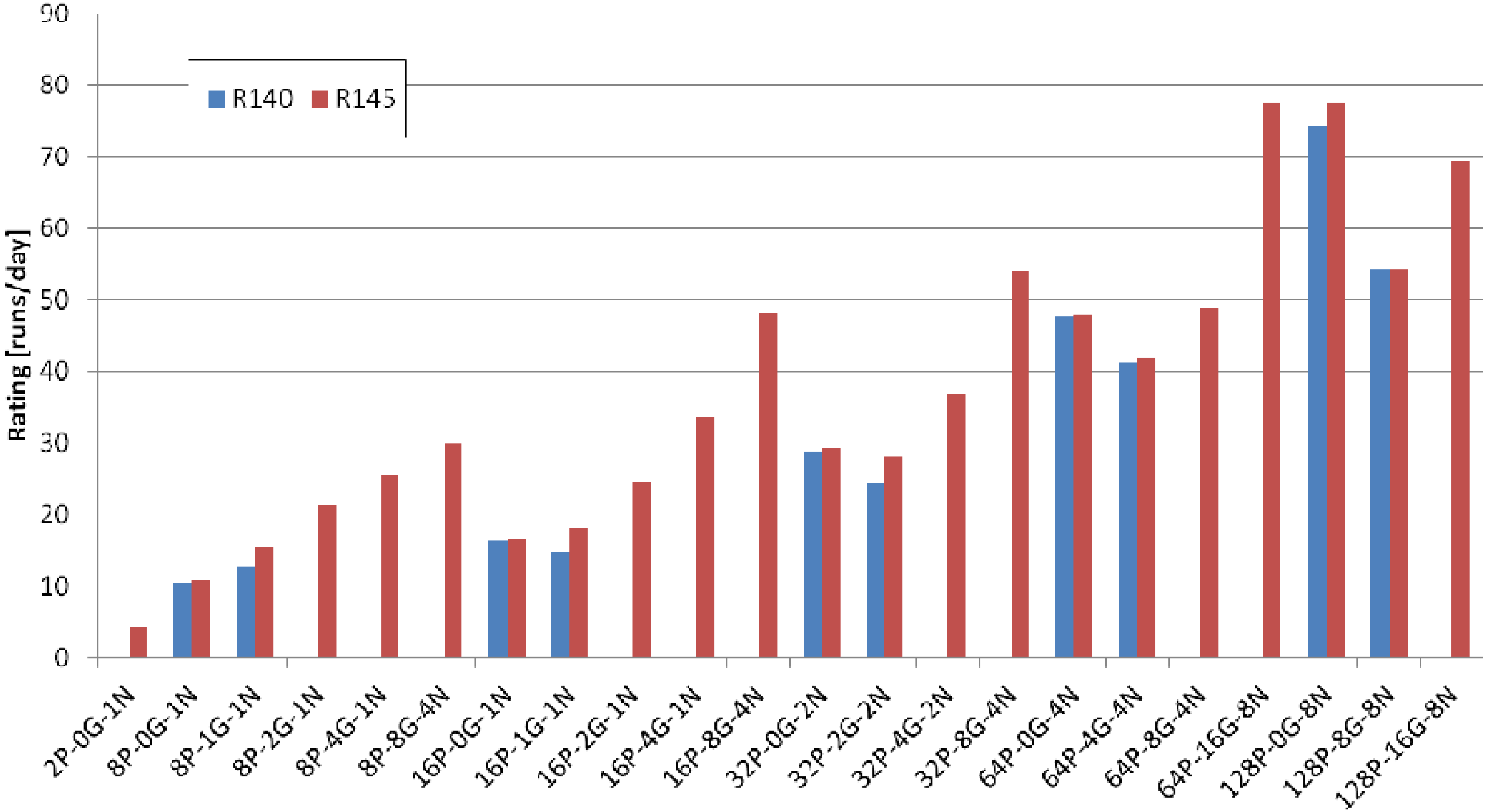
MicroConsult Engineering

ANSYS CONFERENCE &
31. CADFEM USERS' MEETING
19. - 21. Juni 2013, Mannheim

BGA Benchmark with R14.5 on Sandy Bridge Xeons + GPUs



BGA Benchmark with R145 (compilation of all results)



GPU Acceleration

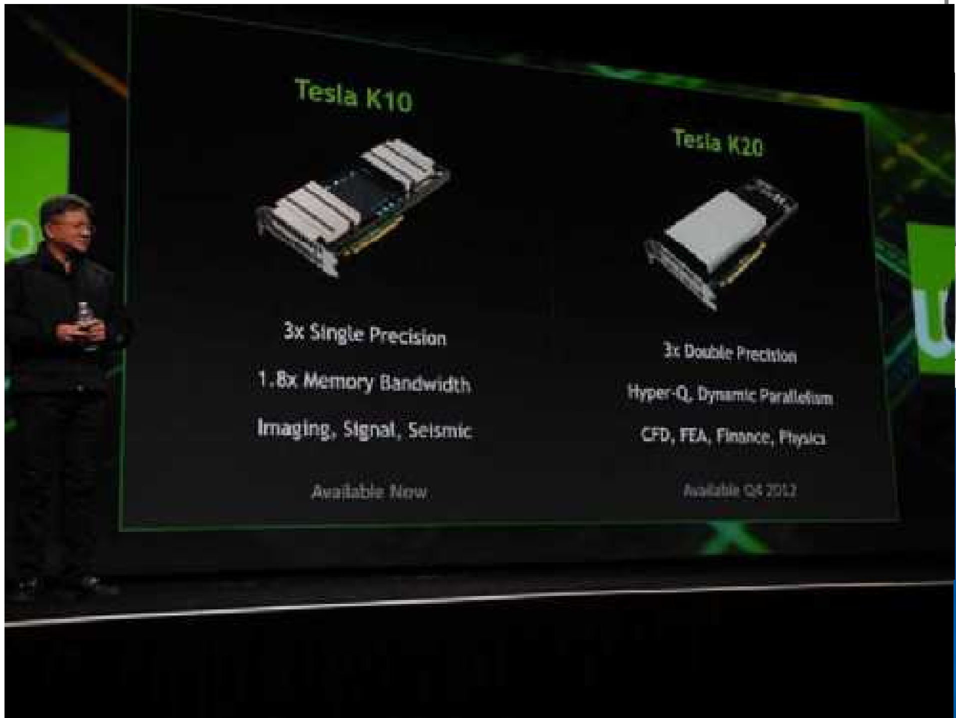
Real life test @ MicroConsult:
Hardware: E5 2690 4x Tesla K20X Accelerator, DSPARSE

```
Tue Jun 18 12:09:58 2013
+------------------------------------------------------+
| NVIDIA-SMI 5.319.23   Driver Version: 319.23         |
|-------------------------------+----------------------+----------------------+
| GPU  Name        Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|         Memory-Usage | GPU-Util  Compute M. |
|===============================+======================+======================|
|   0  Tesla K20Xm         Off  | 0000:02:00.0     Off |                  Off |
| N/A   29C    P0  130W / 235W  |   4118MB /  6143MB   |     99%      Default |
+-------------------------------+----------------------+----------------------+
|   1  Tesla K20Xm         Off  | 0000:03:00.0     Off |                  Off |
| N/A   29C    P0  103W / 235W  |   1990MB /  6143MB   |     99%      Default |
|                               |                      |                      |
|   2  Tesla K20Xm         Off  | 0000:83:00.0     Off |                  Off |
| N/A   28C    P0  109W / 235W  |   3764MB /  6143MB   |     99%      Default |
+-------------------------------+----------------------+----------------------+
|   3  Tesla K20Xm         Off  | 0000:84:00.0     Off |                  Off |
| N/A   30C    P0  120W / 235W  |   3680MB /  6143MB   |     99%      Default |
+-------------------------------+----------------------+----------------------+

+-----------------------------------------------------------------------------+
| Compute processes:                                               GPU Memory |
|  GPU       PID  Process name                                     Usage      |
|-----------------------------------------------------------------------------|
|    0     26177  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145     1316MB  |
|    0     26183  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145     1362MB  |
|    0     26180  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    0     26111  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145     1331MB  |
|    1     26102  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145      519MB  |
|    1     26179  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145     1290MB  |
|    1     26187  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    1     26102  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    2     26191  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145      800MB  |
|    2     26178  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145      875MB  |
|    2     26181  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145      888MB  |
|    2     26155  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145     1092MB  |
|    3     26176  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145     1014MB  |
|    3     26122  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145     1297MB  |
|    3     26167  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    3     26153  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145     1259MB  |
```

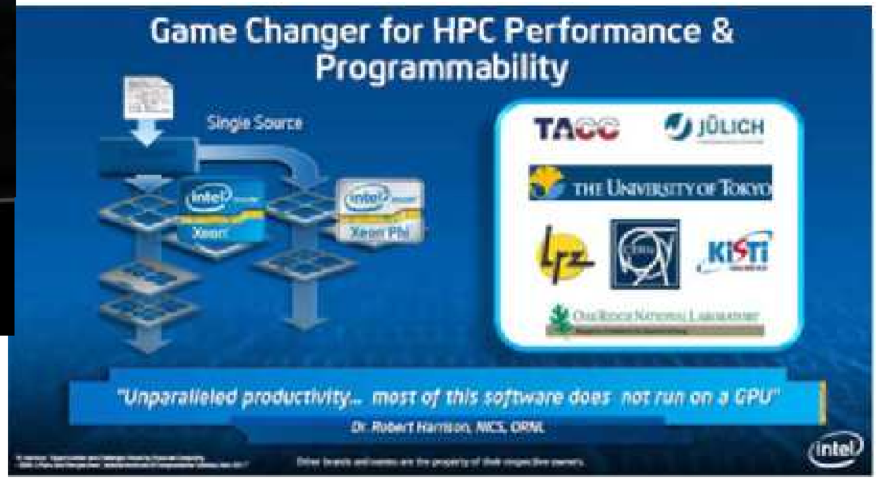
```
Tue Jun 18 12:10:28 2013
+------------------------------------------------------+
| NVIDIA-SMI 5.319.23   Driver Version: 319.23         |
|-------------------------------+----------------------+----------------------+
| GPU  Name        Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|         Memory-Usage | GPU-Util  Compute M. |
|===============================+======================+======================|
|   0  Tesla K20Xm         Off  | 0000:02:00.0     Off |                  Off |
| N/A   28C    P0   57W / 235W  |    349MB /  6143MB   |      0%      Default |
+-------------------------------+----------------------+----------------------+
|   1  Tesla K20Xm         Off  | 0000:03:00.0     Off |                  Off |
| N/A   28C    P0   57W / 235W  |    349MB /  6143MB   |      0%      Default |
|                               |                      |                      |
|   2  Tesla K20Xm         Off  | 0000:83:00.0     Off |                  Off |
| N/A   27C    P0   55W / 235W  |    349MB /  6143MB   |      0%      Default |
+-------------------------------+----------------------+----------------------+
|   3  Tesla K20Xm         Off  | 0000:84:00.0     Off |                  Off |
| N/A   30C    P0   57W / 235W  |    349MB /  6143MB   |      0%      Default |
+-------------------------------+----------------------+----------------------+

+-----------------------------------------------------------------------------+
| Compute processes:                                               GPU Memory |
|  GPU       PID  Process name                                     Usage      |
|=============================================================================|
|    0     26177  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    0     26183  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    0     26180  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    0     26111  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    1     26102  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    1     26179  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    1     26187  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    1     26182  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    2     26191  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    2     26178  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    2     26181  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    2     26155  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    3     26176  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    3     26122  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    3     26167  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
|    3     26153  /usr/ansys_inc/v145/ansys/bin/linx64/ansys.e145       80MB  |
```

Duty Cycle ca. 20-30%

Next steps


Tesla K10
3x Single Precision
1.8x Memory Bandwidth
Imaging, Signal, Seismic
Available Now
Tesla K20
3x Double Precision
Hyper-Q, Dynamic Parallelism
CFD, FEA, Finance, Physics
Available Q4 2012


intel inside
Xeon Phi
intel


Game Changer for HPC Performance & Programmability
Single Source
TACC
JÜLICH
THE UNIVERSITY OF TOKYO
KISTI
"Unparalleled productivity… most of this software does not run on a GPU"
Dr. Robert Harrison, NICS, ORNL
intel

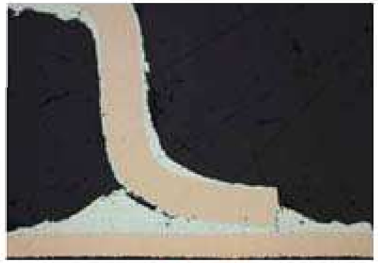
Applications:

- BGA, LQFP
- Einzelbauteile & Systembetrachtung
- Schwerpunkt Lotkriechen

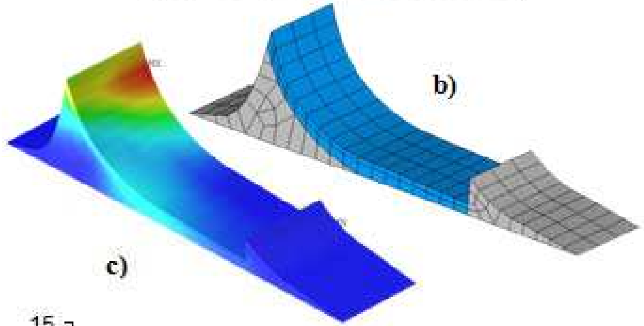
Development of a submodel technique for the simulation of solder joint fatigue of electronic devices mounted within an assembled ECU.

Robert Bosch GmbH, Schwieberdingen, Germany

MicroConsult GmbH, Germany

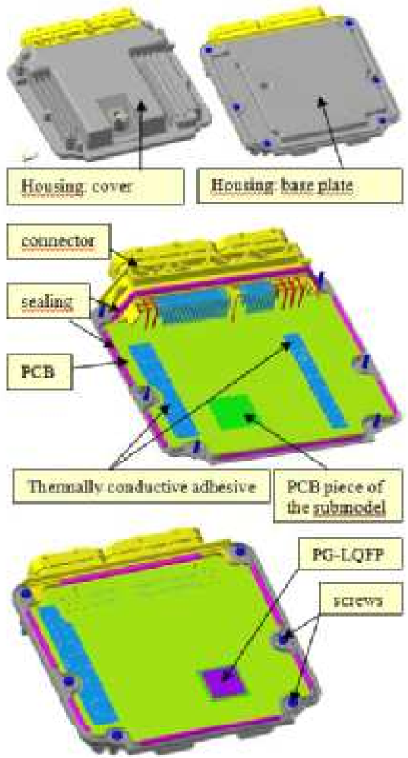


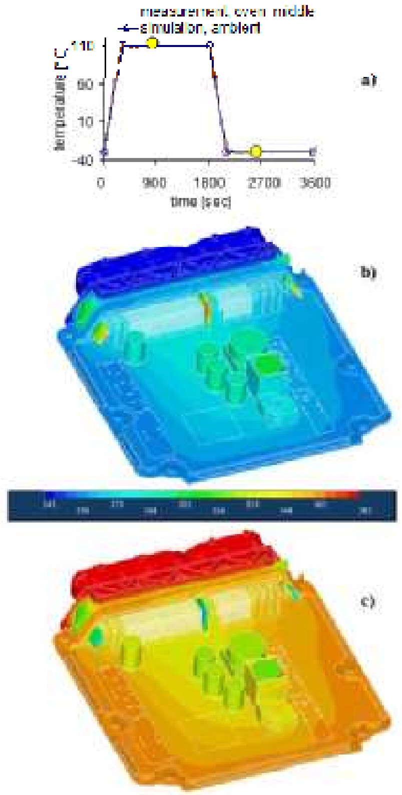
a)

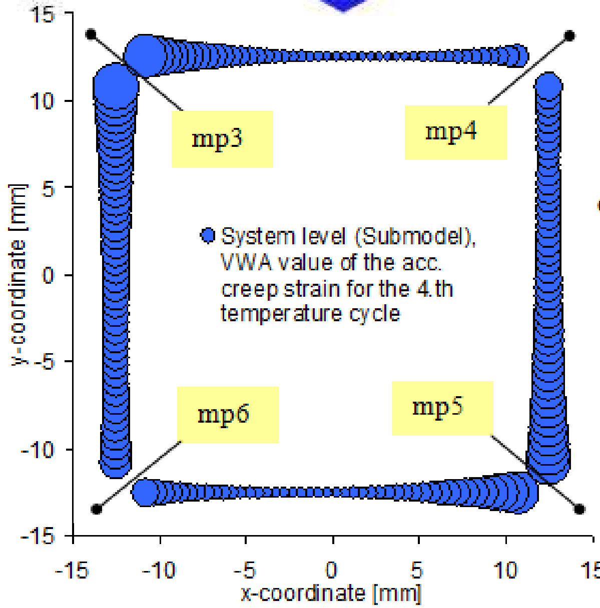


b)

c)







d)

MicroConsult Engineering

ANSYS CONFERENCE &
31. CADFEM USERS' MEETING
19. - 21. Juni 2013, Mannheim

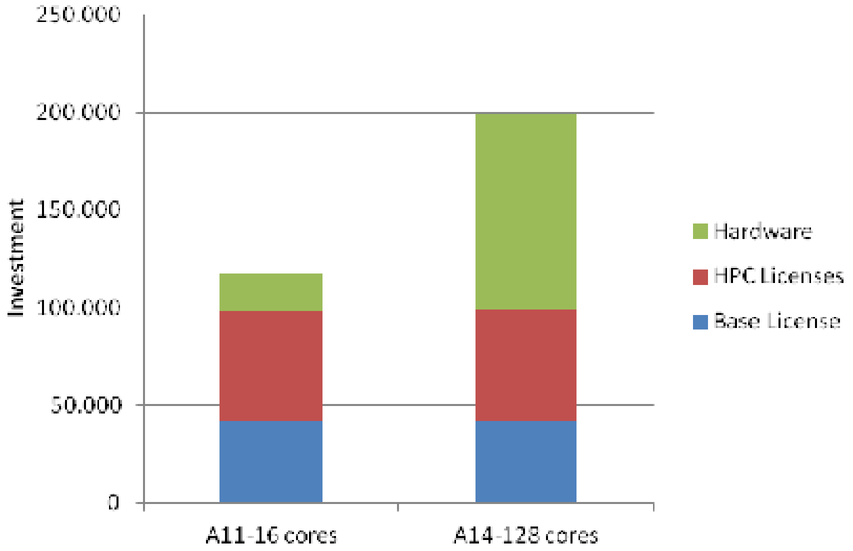
Benchmark Results: Leda Benchmark

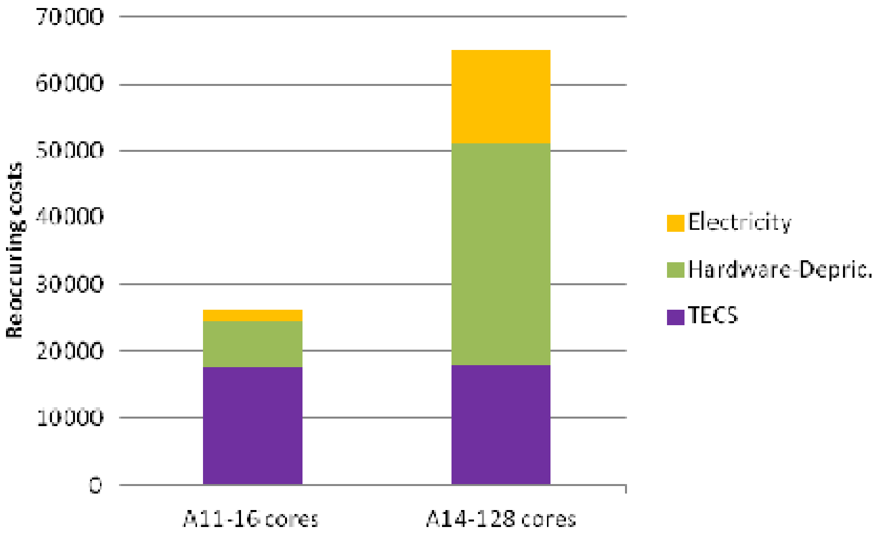
Procedure	ANSYS 11	ANSYS12	ANSYS12.1	ANSYS13 SP02	ANSYS 14 *(UP20110901)*	ANSYS 14.5	ANSYS 14.5 (UP20120826)
Thermal (full model) 3 MDOF	4h (8 cores)				***1h (8 cores + 1 GPU) 0.8h (32 cores)***		
Thermo-mechanical Simulation (full model) 7.8 MDOF	~ 5.5 days for 163 iterations (8 cores)	34.3h for 164 iterations (20 cores)	12.5h for 195 iterations (64 cores)	***9.9h for 195 iterations (64 cores)***	***7.5h for 195 iterations (128 cores)***	***6.4h for 196 iterations (128 E5 cores)***	***7.2h for 196 iterations (72 cores + 12 GPUs)***
Interpolation of boundary conditions	37h for 16 Loadsteps	Identical to ANSYS 11	Identical to ANSYS 11	***0.2h (improved algorithm)***	***0.2h***		
Submodell: Creep Strain Analysis 5.5 MDOF	~ 5.5 days for 492 iterations (16 cores)	38.5h for 492 iterations (16 cores)	8.5h for 492 iterations (76 cores)	***6.1h for 488 iterations (128 cores)***	***5.9h for 498 iterations (64 cores + 8GPUs) 4.2h (256 cores)***	***4h for 498 iterations (128 E5 cores) 4.8h (128 E5 cores + 16 GPUs)***	***5.5h for 498 iterations (72 cores + 12 GPUs)***
	2 weeks	5 days	2 days	1 day	½ day		

Best Performance with E5 Xeons

All runs with SMP Sparse or DSPARSE solver
Hardware 11 & 12: Dual X5460 (3.16 GHz Harpertown Xeon)
Hardware 12.1...14.5: Dual X5570 (2.93 GHz Nehalem Xeon) or Dual X5670 (2.93 GHz Westmere Xeon), M207x Nvidia GPUs, 14.5 results also with Dual E5 2690 (2.9 GHz Sandy Bridge Xeon)
ANSYS 14 + 14.5 creep runs with NROPT,,crpl + DDOPT, metis
ANSYS 12...14.5 runs with Infiniband interconnect

Comparison: 2009 vs. 2012

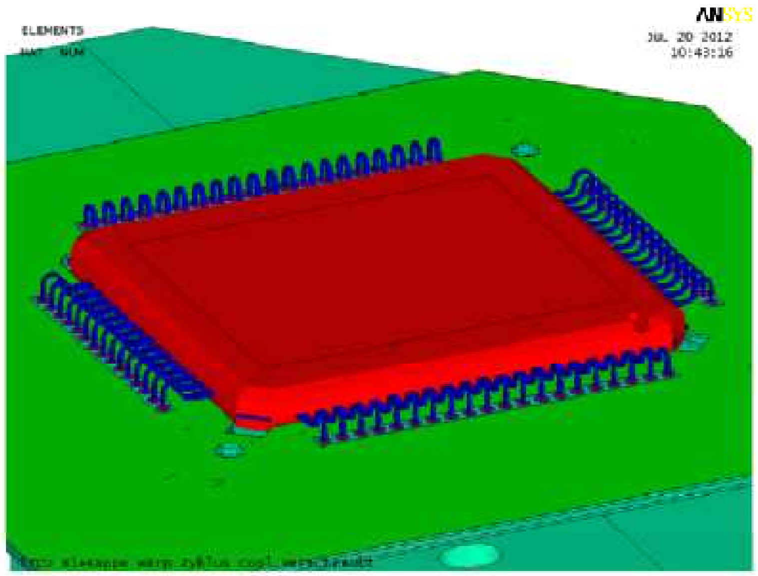


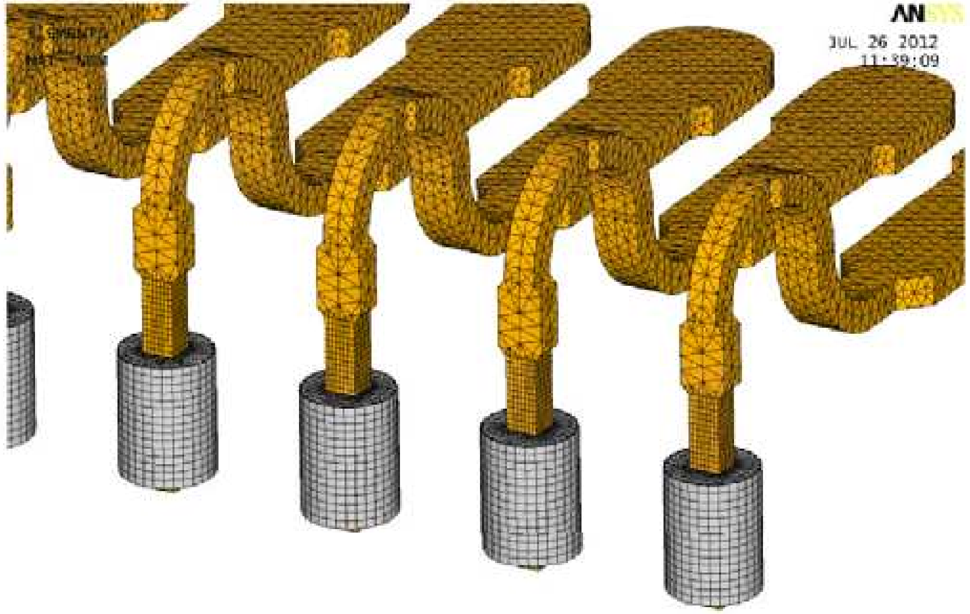


Update 2013: software costs dominate,
even @ 128 cores.

Examples

periodic structure, identical pins





Comparison for 5 MDOF model (w. contacts; R14.5)

w/o GPUs (E5 2690)

```
===========================
= multifrontal statistics =
===========================

    number of equations                          =         5162979
    no. of nonzeroes in lower triangle of a      =       406162318
    no. of nonzeroes in the factor l             =     11975754349
    ratio of nonzeroes in factor (min/max)       =          0.0201
    number of super nodes                        =           64317
    maximum order of a front matrix              =           18405
    maximum size of a front matrix               =       169381215
    maximum size of a front trapezoid            =       109457337
    no. of floating point ops for factor         =      8.1909D+13
    no. of floating point ops for solve          =      4.7268D+10
    ratio of flops for factor (min/max)          =          0.0388
    near zero pivot monitoring activated
    number of pivots adjusted                    =               0
    negative pivot monitoring activated
    number of negative pivots encountered        =               0
    factorization panel size                     =             128
    number of cores used                         =             128


    time (cpu & wall) for structure input        =        1.280000        1.284794
    time (cpu & wall) for ordering               =       15.169800       15.169800
    time (cpu & wall) for other matrix prep      =        6.890200        6.938609
    time (cpu & wall) for value input            =        1.420000        1.429186
    time (cpu & wall) for matrix distrib.        =        2.640000        2.653399
    time (cpu & wall) for numeric factor         =       74.380000       74.574193
    computational rate (mflops) for factor       =  1101222.924216  1098355.317452
    time (cpu & wall) for numeric solve          =        0.850000        0.858288
    computational rate (mflops) for solve        =    55609.392045    55072.387450
    effective I/O rate (MB/sec) for solve        =   211871.780508   209825.793035

 Memory allocated on core    0        =     1587.147 MB
 Memory allocated on core    1        =     1482.573 MB
…
 Memory allocated on core  126        =     1491.825 MB
 Memory allocated on core  127        =     2185.482 MB
 Total Memory allocated by all cores  =   197172.640 MB


 DSP Matrix Solver         CPU Time (sec) =        103.130
 DSP Matrix Solver     ELAPSED Time (sec) =        103.425
 DSP Matrix Solver      Memory Used ( MB) =       1587.147
```

w dual GPUs (E5 2690)

```
===========================
= multifrontal statistics =
===========================

    number of equations                          =         5162979
    no. of nonzeroes in lower triangle of a      =       406162561
    no. of nonzeroes in the factor l             =     11877671289
    ratio of nonzeroes in factor (min/max)       =          0.1692
    number of super nodes                        =           64238
    maximum order of a front matrix              =           13914
    maximum size of a front matrix               =        96806655
    maximum size of a front trapezoid            =        59723199
    no. of floating point ops for factor         =      7.8384D+13
    no. of floating point ops for solve          =      4.6973D+10
    ratio of flops for factor (min/max)          =          0.4117
    near zero pivot monitoring activated
    number of pivots adjusted                    =               0
    negative pivot monitoring activated
    number of negative pivots encountered        =               0
    factorization panel size                     =             128
    number of cores used                         =              64
    GPU acceleration activated
    percentage of GPU accelerated flops          =         99.7144
    time (cpu & wall) for structure input        =        1.060000        1.058888
    time (cpu & wall) for ordering               =       14.095459       14.095459
    time (cpu & wall) for other matrix prep      =        6.354541        7.541591
    time (cpu & wall) for value input            =        0.950000        0.944036
    time (cpu & wall) for matrix distrib.        =        3.340000        3.326019
    time (cpu & wall) for numeric factor         =       35.900000       37.974798
    computational rate (mflops) for factor       =  2183395.338670  2064102.972591
    time (cpu & wall) for numeric solve          =        0.730000        0.723194
    computational rate (mflops) for solve        =    64346.638293    64952.225467
    effective I/O rate (MB/sec) for solve        =   245160.688215   247467.975314

 Memory allocated on core    0        =     2420.815 MB
 Memory allocated on core    1        =     2283.246 MB
…
 Memory allocated on core   62        =     2758.796 MB
 Memory allocated on core   63        =     3805.647 MB
 Total Memory allocated by all cores  =   156488.815 MB


 DSP Matrix Solver         CPU Time (sec) =         62.870
 DSP Matrix Solver     ELAPSED Time (sec) =         66.165
 DSP Matrix Solver      Memory Used ( MB) =       2420.815
```

GPU Performance tested with mold injected part (w. fibers)

Objective

For a plastic cover generated via mold injection from a fiber reinforced plastic (PA66GF30) there is a considerable variation of the material properties caused by a variation in the direction of the fiber orientation. Furthermore, the degree of orientation will vary locally.

The fiber orientation can be calculated outside of ANSYS and mapped onto the model. However, a much finer mesh is needed to represent the locally varying material accurately, compared to the situation with a homogenous material.

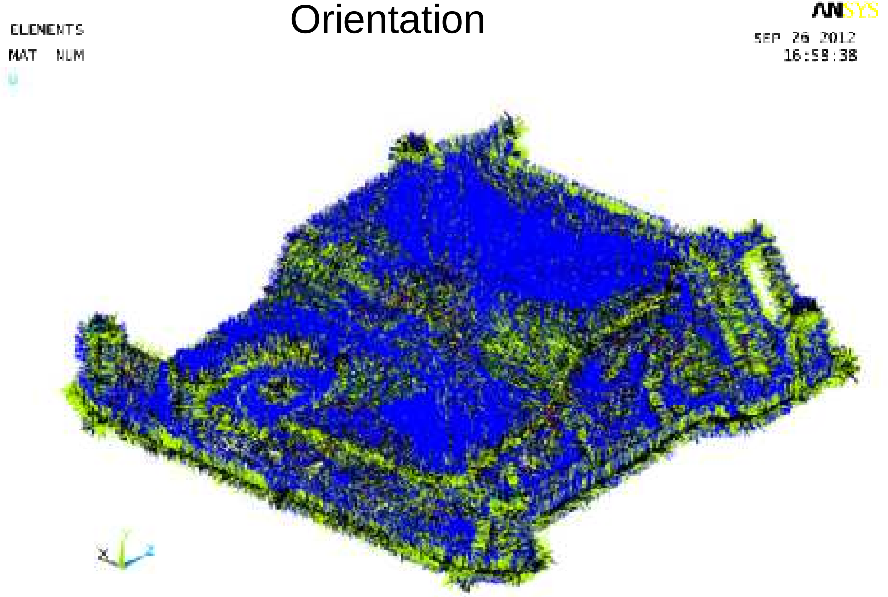
During a customer project we made a study with models of different meshing density (meshed inside workbench) to investigate the displacements under thermal load

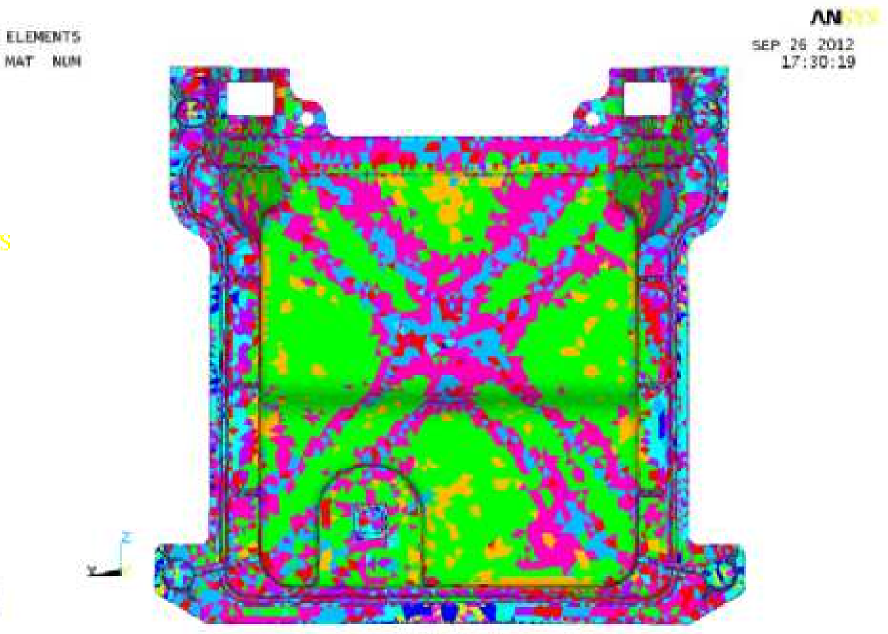
The model is a simple bulk model (solid 186), no contacts, no material non-linearities.

Coarse model (2mm Tets): 0.7 MDOF
Medium model (0.5mm HexDom): 5.9 MDOF

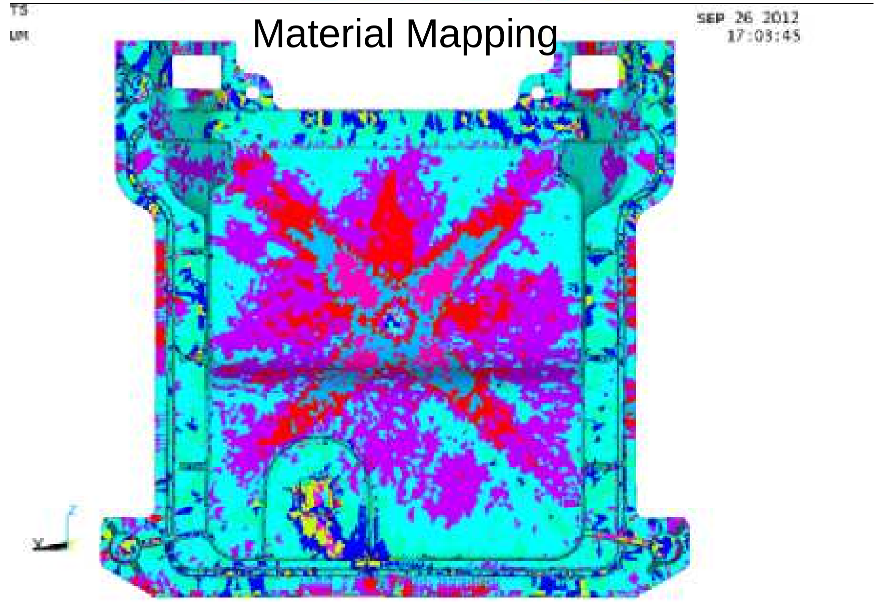
Objective

Orientation



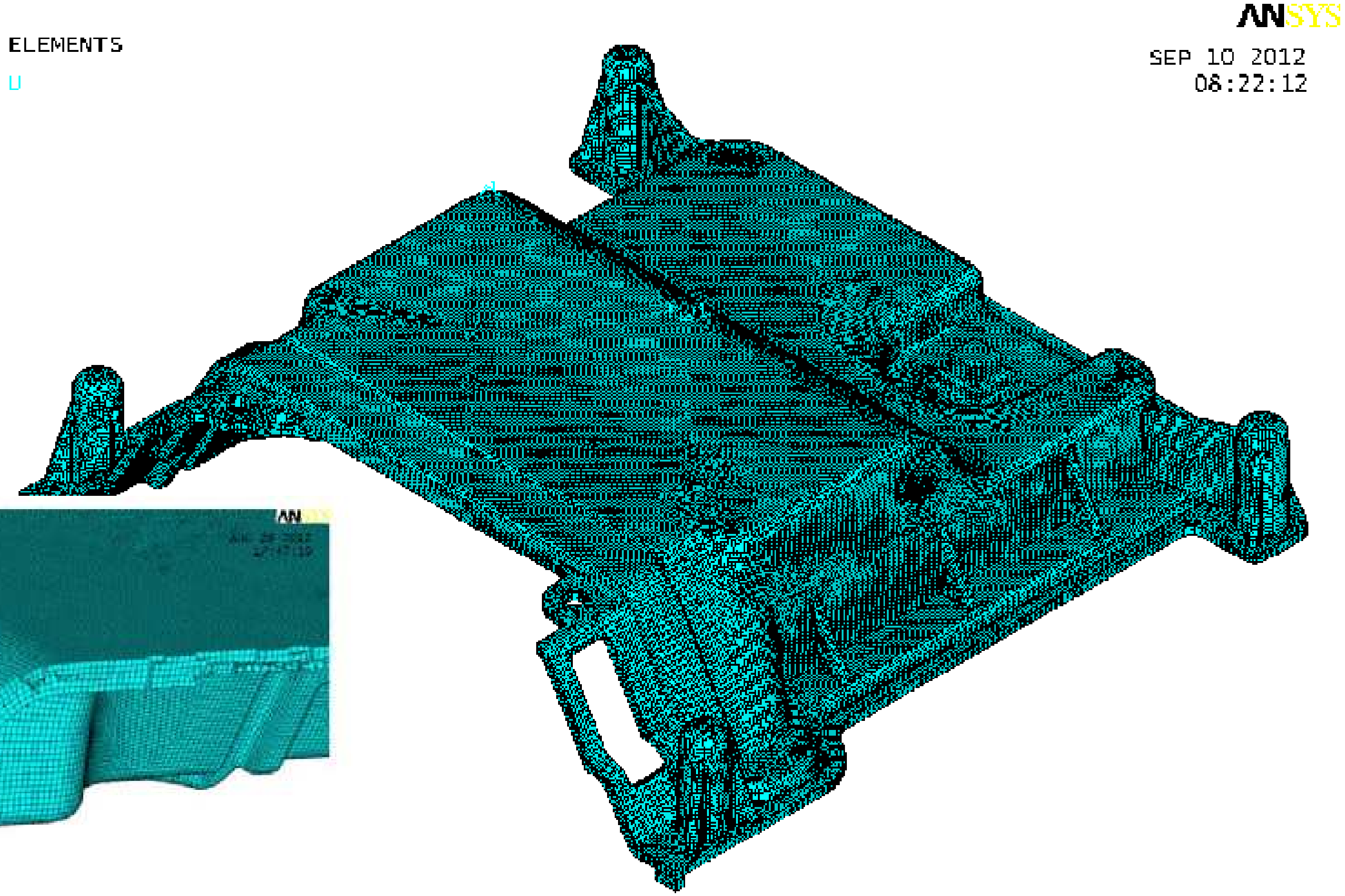


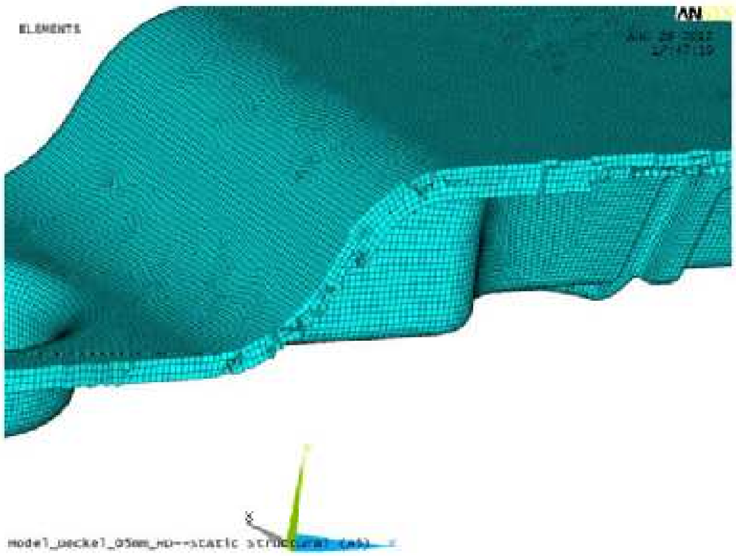
Material Mapping



MicroConsult Engineering

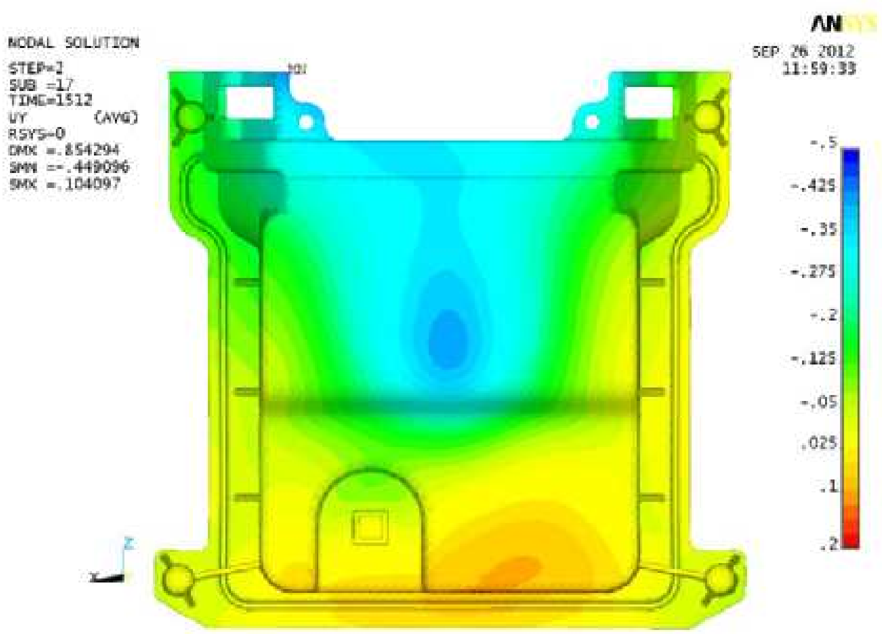
Model 0.5 mm Hex Dominant



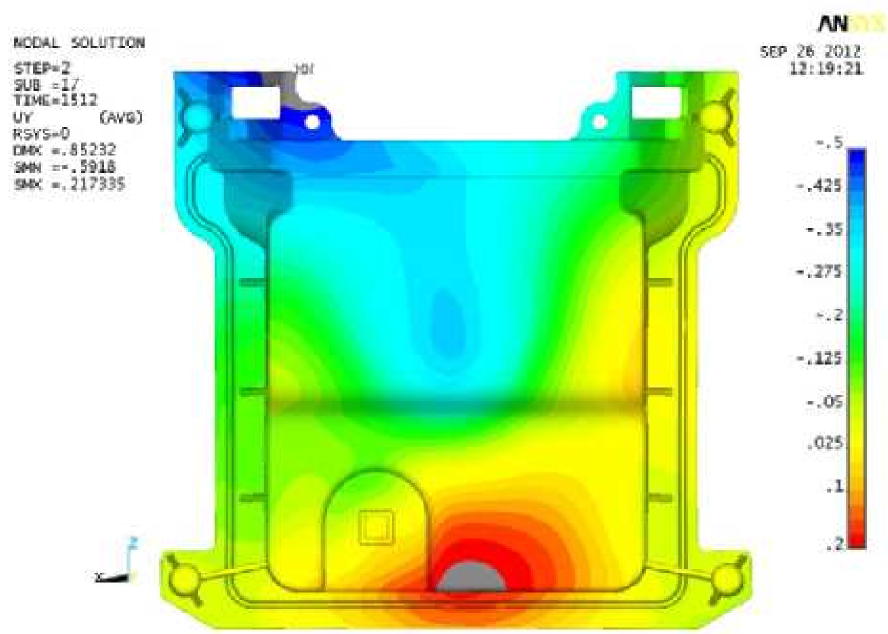


Difference in Displacements (free expansion)

2mm Tet Mesh



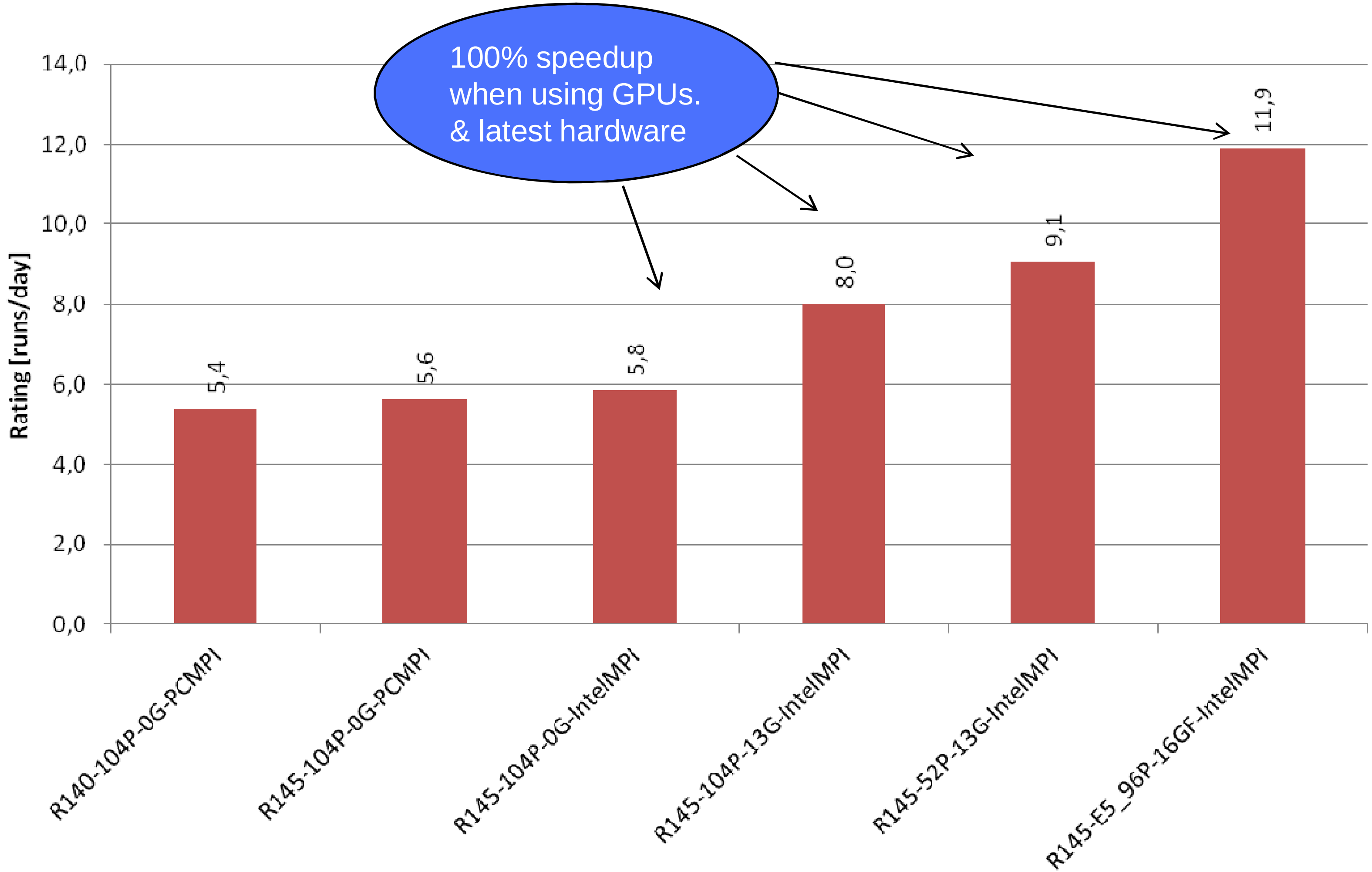
0.5mm Hex Dom Mesh



Coarse model (2mm Tets): 0.7 MDOF
Medium model (0.5mm HexDom): 5.9 MDOF

Displacement range off by about 50%

Results for 0.5mm HexDom model



Conclusions

- ANSYS Mechanical routinely deliver TFLOP per second performance in a HPC environment!
  - Highest Peak performance with GPUs (and suitable case)
  - Conventional solution provides similar performance with fewer surprises.
  - GPU licensing & stability critical for adoption

Acknoweldgements

- Jeff Beisheim, ANSYS Inc.
- Erke Wang, Peter Tiefenthaler, CADFEM GmbH
- Natalja Schafet, Wolfgang Müller-Hirsch, Robert Bosch GmbH
- Philipp Schmid, Holger Mai, MicroConsult Engineering GmbH

SELECT THE RIGHT TESLA GPU

Features	Tesla K20X	Tesla K20	Tesla K10	Tesla M2090	Tesla M2075
Number and Type of GPU	1 Kepler GK110		2 Kepler GK104s	1 Fermi GPU	1 Fermi GPU
GPU Computing Applications	Seismic processing, CFD, CAE, Financial computing, Computational chemistry and Physics, Data analytics, Satellite imaging, Weather modeling		Seismic processing, signal and image processing, video analytics	Seismic processing, CFD, CAE, Financial computing, Computational chemistry and Physics, Data analytics, Satellite imaging, Weather modeling	
Peak double precision floating point performance	1.31 Tflops	1.17 Tflops	190 Gigaflops (95 Gflops per GPU)	665 Gigaflops	515 Gigaflops
Peak single precision floating point performance	3.95 Tflops	3.52 Tflops	4577 Gigaflops (2288 Gflops per GPU)	1331 Gigaflops	1030 Gigaflops
Memory bandwidth (ECC off)	250 GB/sec	208 GB/sec	320 GB/sec (160 GB/sec per GPU)	177 GB/sec	150 GB/sec
Memory size (GDDR5)	6 GB	5 GB	8GB (4 GB per GPU)	6 GigaBytes	6 GigaBytes
CUDA cores	2688	2496	3072 (1536 per GPU)	512	448