ANSYS HPC Seminar Series

CFD

- Prepared and presented by Alex Pickard

apickard@simutechgroup.ca

Who Are We

- 120+ Employees
- 12 Local Offices
- 2000+ Customers
- 80% Engineering Staff with Advanced degree

Today’s Adenda

- Introduction to CFD HPC, solvers, and licenses
- Computer platform recommendations (CPU models & features, RAM quantity, memory bandwidth, storage, networking)
- Hardware recommendations for users (solver and preprocessor)
- Performance diagnostic information from reading solver logs
- Simulation best practices for performance and scalability
- Sizing resource request for your model
- Intro to clusters and Job Schedulers
- Workflow recommendations based on analysis type
- Demos / Workshops
  - Job submission techniques
  - Multi-step simulation (3D Printing)
  - Improving performance and scalability of contacts

What is HPC?

High-performance computing is the use of parallel processing techniques for solving complex computational problems.

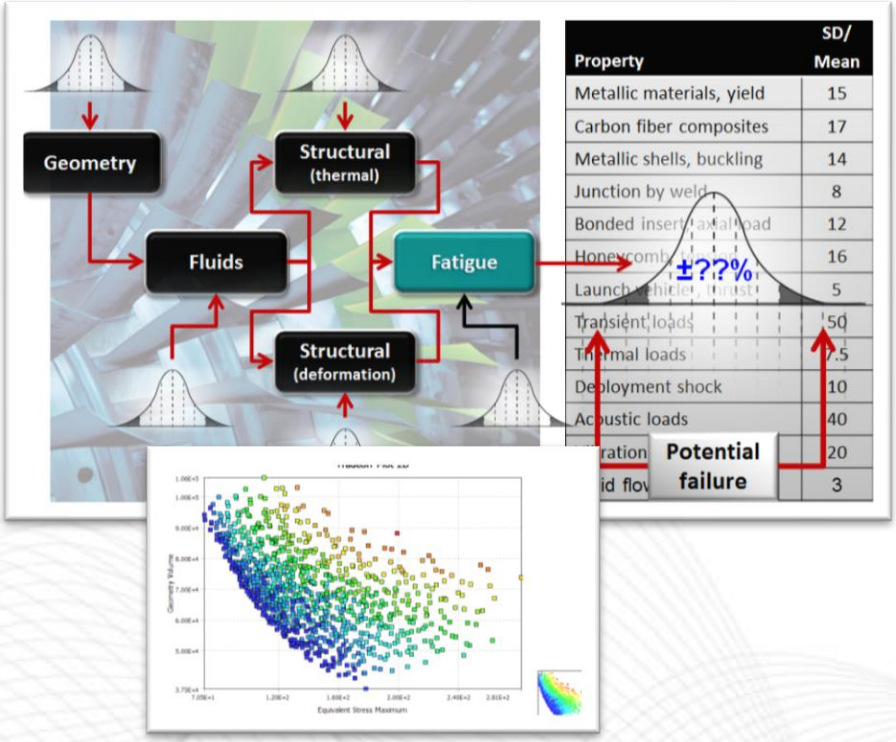
It especially refers to using multiple computers to work together on a single problem (clusters). It does not necessarily mean working on a single solution.

What is HPC?

Having HPC capability increases throughput:

- Faster results
- More design iteration
- Hit hard deadlines
- Greater engineering efficiency
- Parametric analysis & optimization

HPC enables more thorough design and analysis on a tighter deadline.

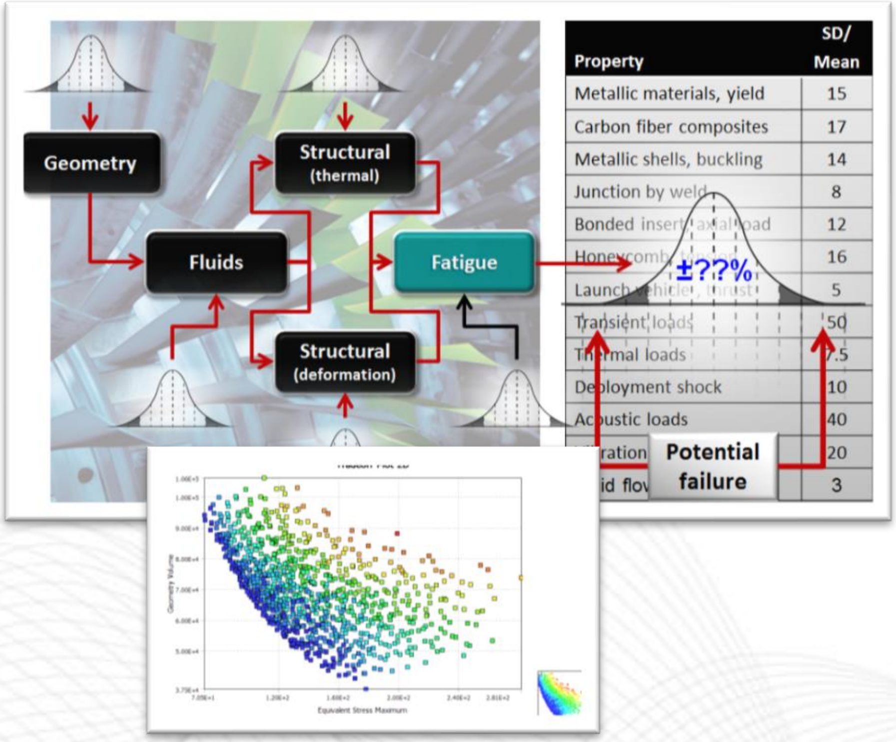


What is HPC?

Having HPC capability increases capability:

- More complex models
- More physics
- Less risky simplifications
- Greater detail
- System level analysis
- Discovery of new insight

HPC unlocks new capability within the ANSYS products your already have.



ANSYS CFD

- The ANSYS CFD suite consists of multiple high-end CFD codes and programs:

Fluent: A general CFD code considered our go-to.

- Excellent scalability and flexibility with UDFs

CFX: Another general use code with specialization in turbomachinery

- Slightly easier to use, greater starter

Icepak: A program made to do heat transfer, especially electronics

- Built on Fluent's solver and gains thus has similar HPC characteristics

Others which won't be covered:

- Forte (IC Engines), Fensap-ICE (Icing), Polyflow (Polymers+)

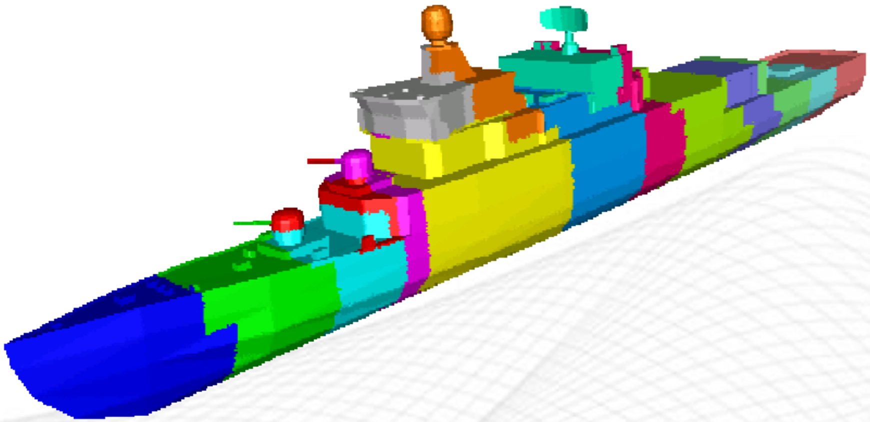
ANSYS Academic & HPC Licence Changes

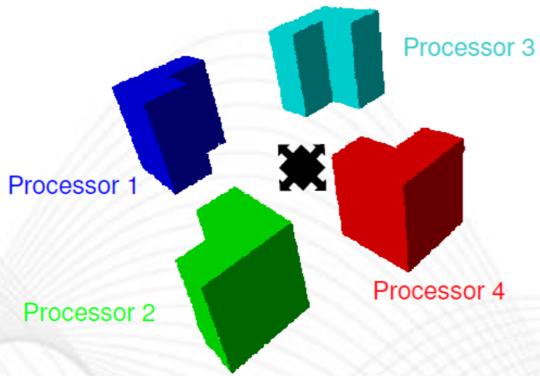
- ANSYS HPC licencing has changed multiple times in the last 5 years.
- Overall trend is to become (slightly) simpler.
- As of 2021 Academic and Commercial licencing is identical:
  - Previously there were dedicated research licenses that provided 16 cores standard.
  - Now academic and campus licence bundles contain **commercial** licenses.
  - Campus bundles contain more HPC to compensate, more than makes up for the change.
    - Old 10 Research bundle had 10 x 16 core solvers + 64 HPC (no solve bigger than 80 cores)
    - New 10 Research bundle has 10 x 4 core solvers + 180 HPC (allows up to 184 cores in one solve)
- Sovlers enable 4 cores standard + HPC licence to add individual cores.
- HPC Packs are also available which have an exponential effect:
  - First pack **triples** the allowed cores: 4 -> 12 = 3X
  - Second pack **triples** it again: 12 -> 36 = 3X
  - Next ones is nearly **quadruple**: 36 -> 132 = ~3.7X
  - Note: Uncommon in academia… 132 -> 516 = ~3.9X

Packs	Added Cores	Total Cores
1	8	12
2	32	36
3	128	128
4	512	516
5	2048	2052

Distributed Memory Mode

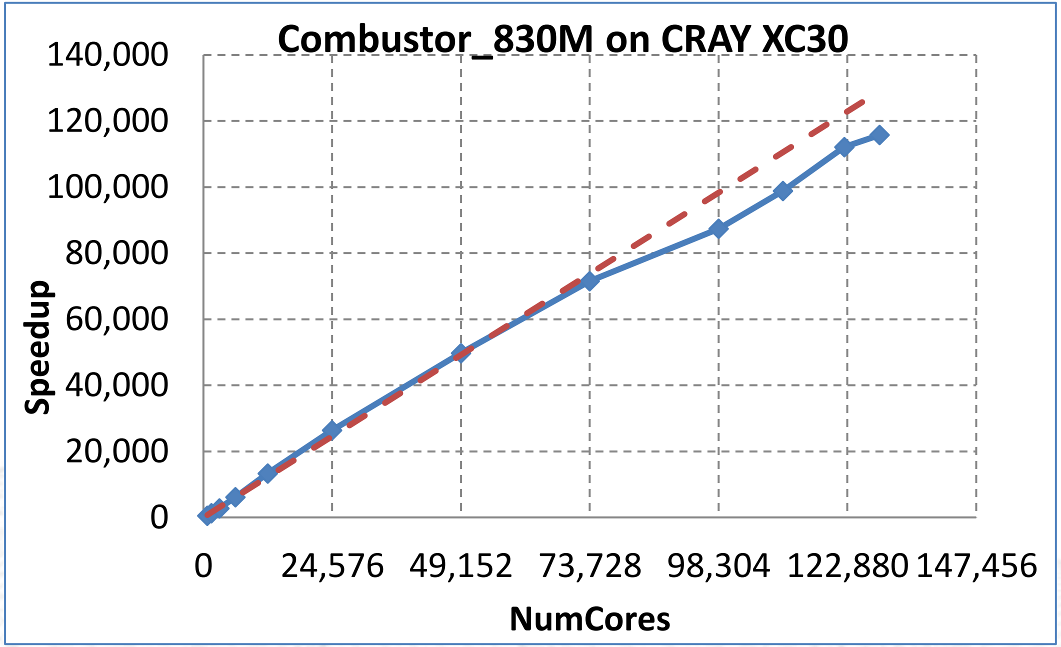
- Distributed solvers are standard across many ANSYS products.
- They scale and perform better than shared memory solvers.
- They enable each CPU core to have it's own solver process and work on and independent chunk of the problem
- Requires substantial coordination and communication between processes.
- Enables problems to be run across multiple computers.
- Many technologies are difficult to distribute across independent tasks (mesh changes)





CFD Scaling

- Both CFX and Fluent achieve excellent scaling to thousands of cores.
- Fluent still has greater potential is this regard:



CFX Solver Memory & CPU Requirements

- Memory usage and computation time are very feature driven.
- Help has detailed description of features that impact requirements

Chapter 15: CPU and Memory Requirements

This chapter provides information on typical increases in CPU (central processing unit) time and memory requirements incurred by some simulations and physical models:

- Tetrahedral Mesh
- Executable Selection
- Turbulence
- Energy Models
- CHT Regions
- Multicomponent Flows
- Multiphase Flows
- Additional Variables, Wall Distance Variables, and Boundary Distance Variables
- Combustion Modeling
- Radiation Modeling
- GGI Interfaces
- Transient Runs
- Mesh Deformation
- Bidirectional (Two-Way) Couplings with System Coupling

- Tets require 0.4x memory per element, or 2x per node, vs. Hex Mesh.
  - (Tets 5:1 with nodes, Hex 1:1)
- Double precision for large changes in grid dimension, aspect ratio, pressure range, multi-phase, etc.

CFX Solver Memory & CPU Requirements

- Large problem solver for 2^31 words of 4-bytes (~80M Hex or ~200M Tet).
- Hex meshes are better for multiphase, lower scaling penalty.

# of Phases	Memory Increases	
	(Hex Mesh)	(Tet Mesh)
1	1	1.80
2	2.15	3.40
3	3.50	5.70
4	5.15	8.05
5	7.00	10.60

- Mesh deformation has expensive computations once per timestep, plus extra RAM.
- Energy equations add 33% CPU to momentum and mass equations.
- Many other options (turbulence models, walls, combustion) require mostly CPU not much RAM.

CPU Instructions: AVX

- Many other ANSYS solvers, particularly direct matrix solves in Mechanical and EDT, use AVX instructions to solve their matrix.
- CFD solvers seems to have no or minimal AVX requirements.
- In a few years back an AVX2 binary of fluent was shipped that could allow minor gains (~5%), it seems to have disappeared.
- Performance is driven by core speed and data access bandwidth.

	Microarchitecture	Instruction Set	SP FLOPs / cycle	DP FLOPs / cycle
2017	Skylake	Intel® AVX-512 & FMA	64	32
2014	Haswell / Broadwell	Intel AVX2 & FMA	32	16
2012	Sandybridge	Intel AVX (256b)	16	8
2010	Nehalem	SSE (128b)	8	4

Memory Bandwidth

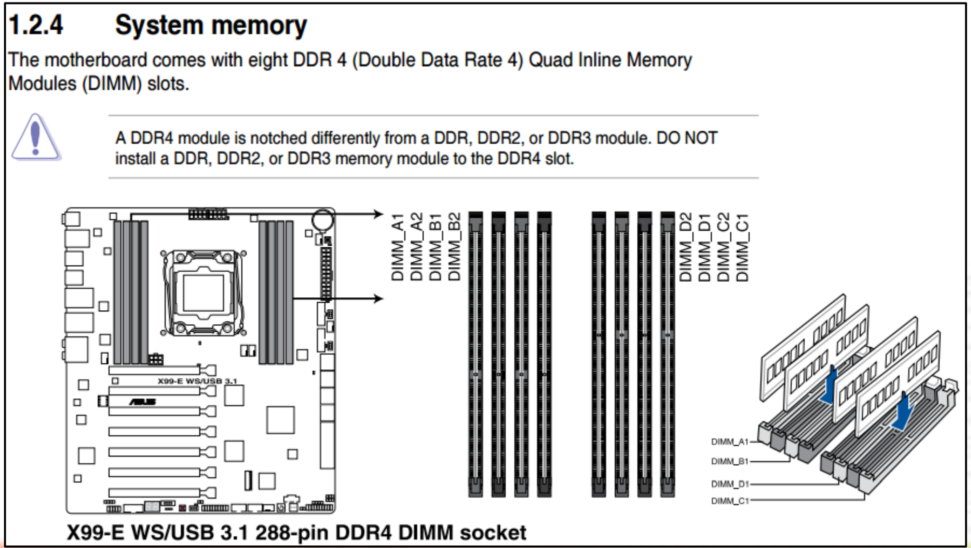
- Simulation software tends to do simple math on large pools of data.
- Data needs to fetched quickly, processed, and then the results stored quickly.
- The ability to fetch and store data is limited by the memory bandwidth and data cache on a system, which is inherent to the platform and CPU.
- Comparison of current server CPU models available:

Model	Intel Xeon Scalable (per socket)	AMD EPYC (per socket)
Memory Channels	6 x DDR4-2933 MHz	8 x DDR4-3200 MHz
L3 Data Cache	Up to 38.5 MB	Up to 256 MB
Cores	Up to 28	Up to 64

- As core count gets higher the bandwidth advantages exceed the core efficiency

Memory Bandwidth

- Be sure to populate all memory channels (not necessarily all slots)
- Aim between between 2-4 Cores / memory channel
- https://simutechgroup.com/maximizing-memory-performance-for-ansys-simulations/



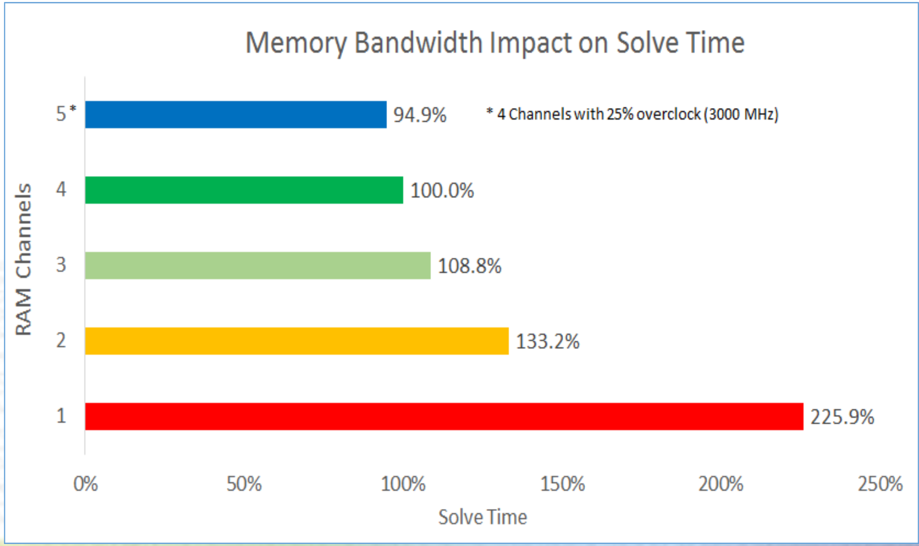
1.2.4 System memory

The motherboard comes with eight DDR 4 (Double Data Rate 4) Quad Inline Memory Modules (DIMM) slots.

A DDR4 module is notched differently from a DDR, DDR2, or DDR3 module. DO NOT install a DDR, DDR2, or DDR3 memory module to the DDR4 slot.

X99-E WS/USB 3.1 288-pin DDR4 DIMM socket

8 Core Test (CFX)

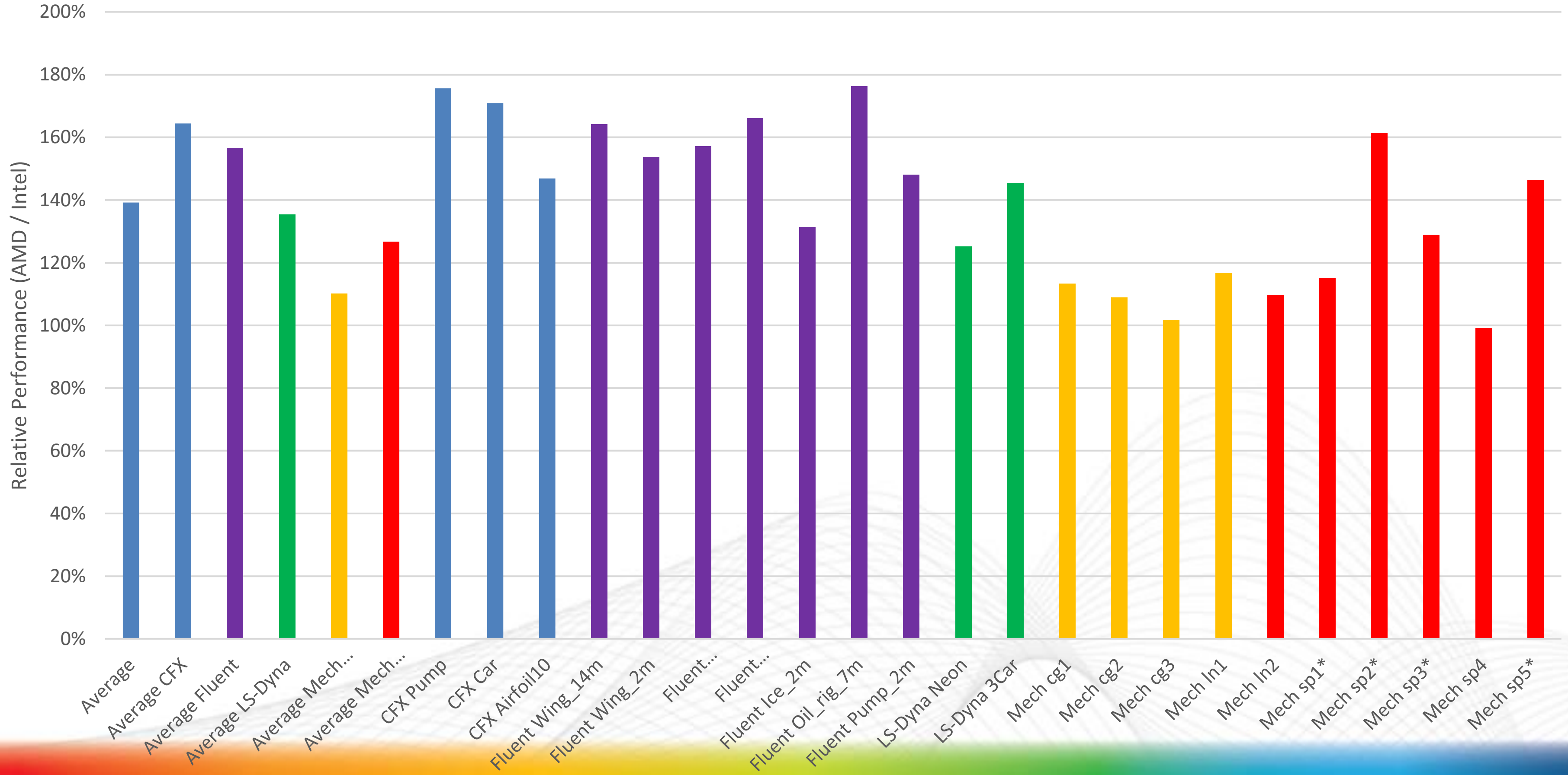


Which CPU Then?

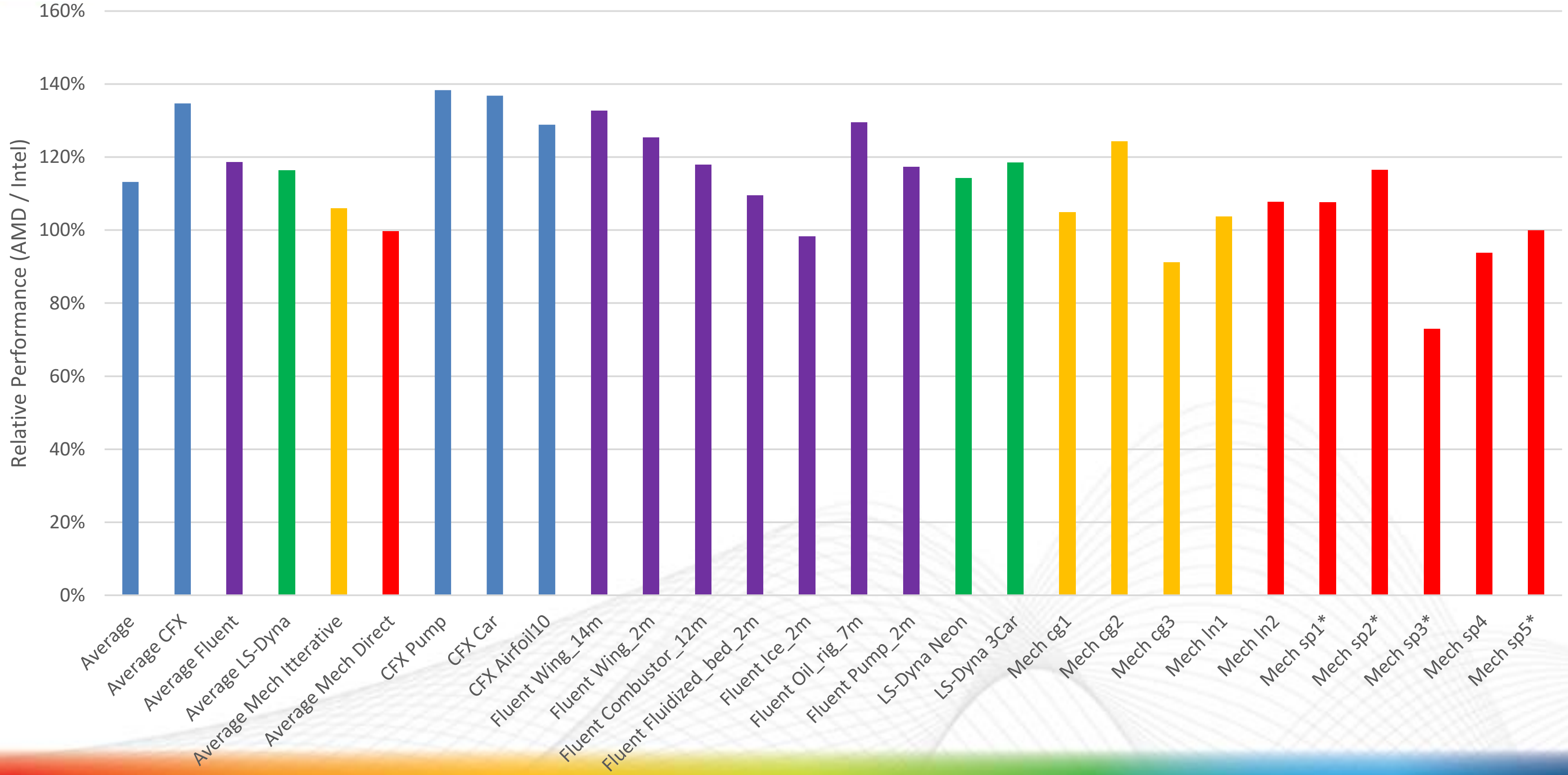
- AMD seems to come out ahead at every core count (but not price point).
- AMD CPUs have massive data cache advantage currently.
- Intel HEDT CPUs (quad memory channel) are cheaper than TR (no Ryzen 5000 data yet)
- Focus on frequency for preprocessing (meshing speed!)
- Either for laptops, focus on platform, power

Cores / Node	AMD	Intel
4-8	Ryzen 5800X	i9-10900k
12-16	Ryzen 5950X Threadripper 3960X	i9-10920X / i9-10980XE Xeon W Xeon Gold
24+	Threadripper Pro 1-2 AMD EPYC	Dual Xeon Gold (like 6242R)

Dual EPYC 7502 vs Dual Xeon 6242R
64 cores vs 40 cores



Dual EPYC 7502 vs Dual Xeon 6242R
36 cores vs 36 cores



Regarding RAM Sizing

- Giving a memory quantity recommendation is difficult because CFD models vary wildly in size, in particular transient vs. steady state.
- Extra RAM beyond what is needed to solve in RAM is not helpful.
- There really is no “page to disk” option, we need to have enough for our problem or compromise.
- We need maximum memory bandwidth (and speed!) from our platform, which frequently drives minimum quantity.
- Largely comes down to economics and appetite, RAM is fairly cheap (64GB+ ?).
- We use a cluster to scale our RAM and cores:
    - Need more RAM: request more nodes.
- SSD performance isn’t critical for solver, but does drive experience and post-processing.

GPUs

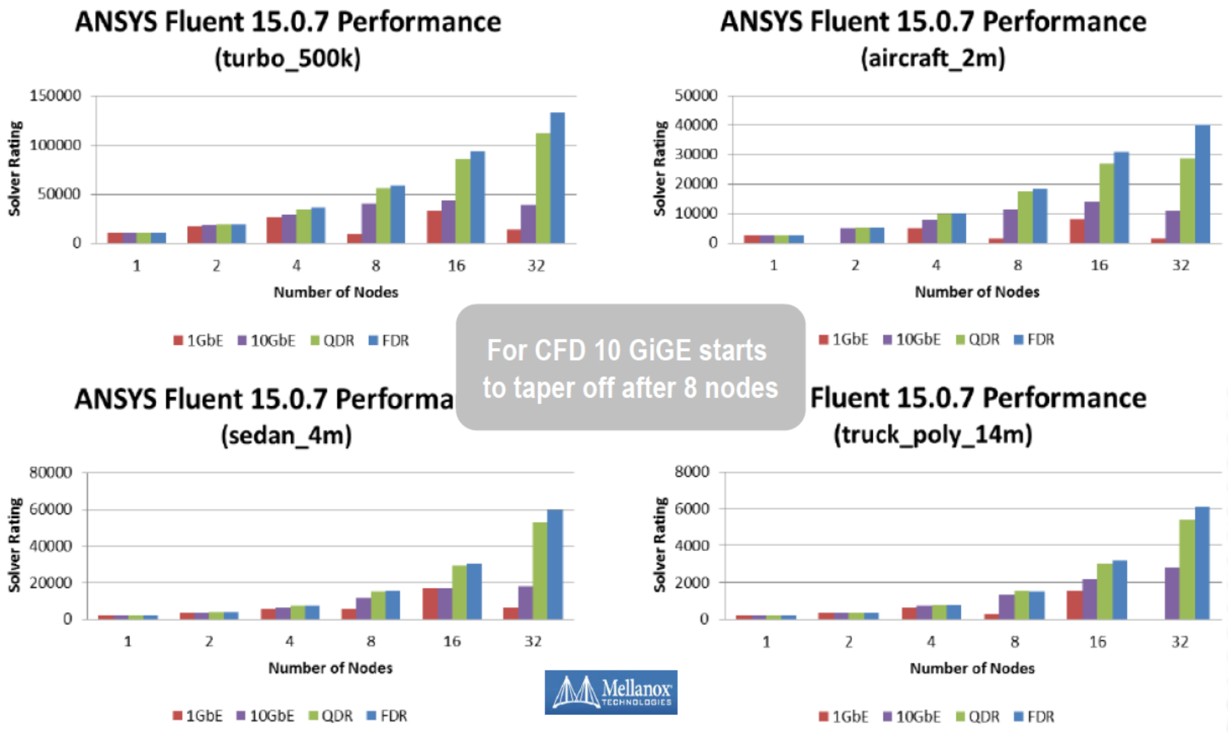
- Dedicated GPU is strongly recommended for GUI usage. (Quadro / Firepro)
- GPU’s for compute are not really recommended
  - Very expensive
  - Inconsistent benefit
  - Less benefit as more cores are available
  - Main use case: reduce licencing costs (they only count as 1 core)

Networking

- Networking nodes together in CFD work extremely well.
- Each node adds more cache, memory bandwidth, memory quantity, cores, etc.
- Regular 1gbps Ethernet can be used for small clusters of small machines (workstations, laptops)
- High speed interconnect is required for larger clusters.
- RDMA communication is effectively required to see maximum gains.
  - Traditionally infiniband was recommended, but ethernet has this too.
  - ~2 microsecond vs 30 microsecond latency, higher bandwidth, less overhead
- Linux easier to implement, but Windows 10 and Server also work.

Networking

- Old data comparing interconnects, but still relevant:



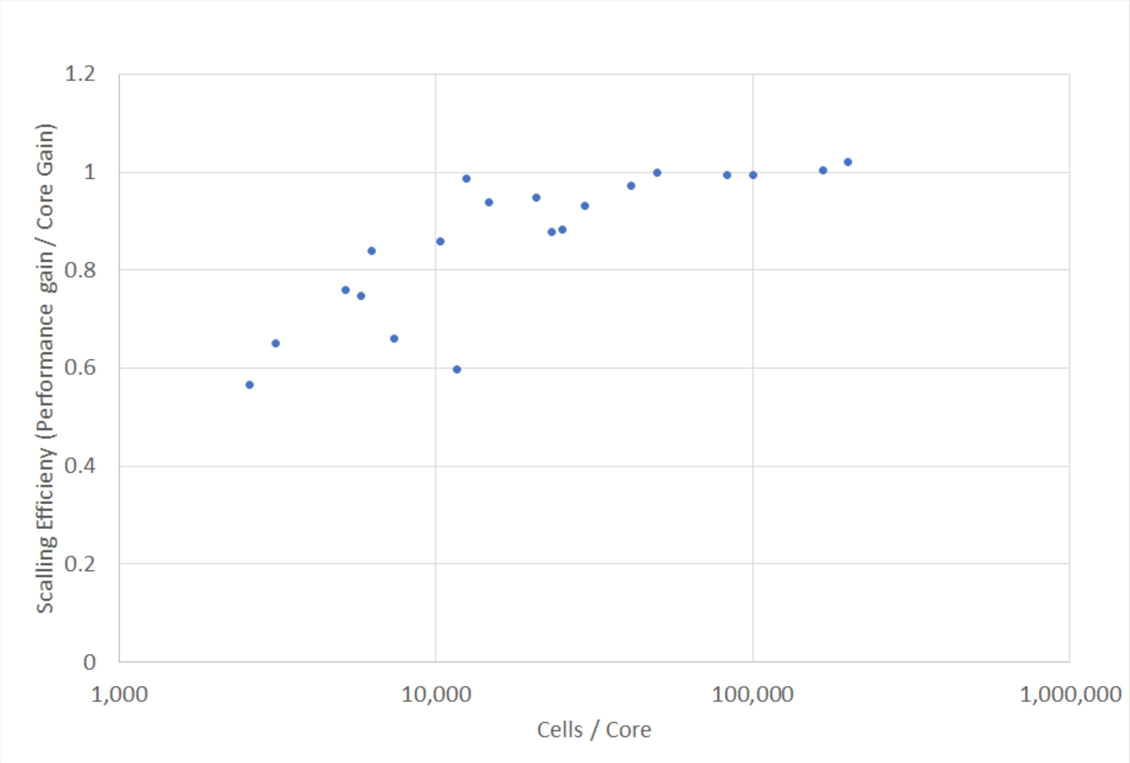
Networking

- Brand new 2021R1 data on Dual EPYC 7542 with 100 gbps Ethernet.

Cores	Oil Rig 7m SP	Oil Rig 7m DP	Combustor 12m SP	Combustor 12m DP	Aircraft Wing 14m SP	Aircraft Wing 14m DP
64	0.819	1.373	12.544	17.255	3.593	4.756
128 TCP	0.658	0.93	8.274	10.704	1.957	2.459
128 RDMA	0.51	0.674	6.563	8.477	1.794	2.314
128 vs 64 TCP mode Speedup	24%	48%	52%	61%	84%	93%
128 vs 64 RDMA mode Speedup	61%	104%	91%	104%	100%	106%

Nodes per Core and Scaling Efficiency

- Fluent data shown from tests at Argone National Labs (thousands of cores)
- Excellent efficiency at > 10k cells per core
- Good scaling down to 2500 cells/core



Load Balancing

- Check for even element distribution
  - Unless machines have dissimilar speed…

Fluent

```
>> 4 Active Partitions:
    P    Cells I-Cells Cell Ratio    Faces I-Faces Face Ratio Neighbors Load
    0     3520     142      0.040    11399     195      0.017         1    1
    1     3298     115      0.035    10678     151      0.014         1    1
    2     3451     305      0.088    11404     372      0.033         2    1
    3     3583     332      0.093    11586     416      0.036         2    1

   ----------------------------------------------------------------------
   Collective Partition Statistics:          Minimum    Maximum    Total
   ----------------------------------------------------------------------
   Cell count                                3298       3583       13852
   Mean cell count deviation                 -4.8%      3.5%
   Partition boundary cell count             115        332        894
   Partition boundary cell count ratio       3.5%       9.3%       6.5%

   Face count                                10678      11586      44500
   Mean face count deviation                 -5.2%      2.8%
   Partition boundary face count             151        416        567
   Partition boundary face count ratio       1.4%       3.6%       1.3%

   Partition neighbor count                  1          2
   ----------------------------------------------------------------------
   Partition Method                          Metis
   Stored Partition Count                    4
```

CFX

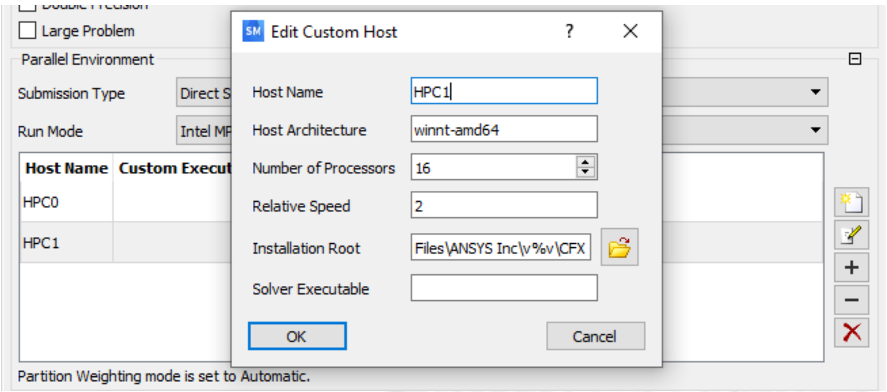
	Elements		Vertices			Faces	
Part	Number	%	Number	%	%Ovlp	Number	%
Full	5362055		1305718			431798	
1	299665	5.5	82340	6.0	2.8	26341	5.9
2	336639	6.2	84940	6.2	3.0	31574	7.1
3	372613	6.8	87696	6.4	5.4	26181	5.9
4	378777	6.9	86632	6.4	5.5	18345	4.1
5	343350	6.3	87688	6.4	3.4	40596	9.2
6	310761	5.7	83422	6.1	3.3	28994	6.5
7	304609	5.6	85408	6.3	4.4	24073	5.4
8	379209	6.9	84450	6.2	5.5	19174	4.3
9	355124	6.5	86655	6.4	3.2	43423	9.8
10	317689	5.8	87163	6.4	5.9	22519	5.1
11	374125	6.8	86617	6.4	4.8	18616	4.2
12	366937	6.7	85446	6.3	6.1	26490	6.0
13	312941	5.7	84959	6.2	4.2	30573	6.9
14	367926	6.7	83759	6.1	5.2	18682	4.2
15	348172	6.4	85002	6.2	2.8	32840	7.4
16	293859	5.4	81763	6.0	2.6	34783	7.8
Min	293859	5.4	81763	6.0	2.6	18345	4.1
(part)		(16)		(16	16)		(4)
Max	379209	6.9	87696	6.4	6.1	43423	9.8
(part)		(8)		(3	12)		(9)
Ave	341400	6.3	85246	6.2	4.3	27700	6.2
Sum	5462396	100.0	1363940	100.0		443204	100.0

Load Balancing

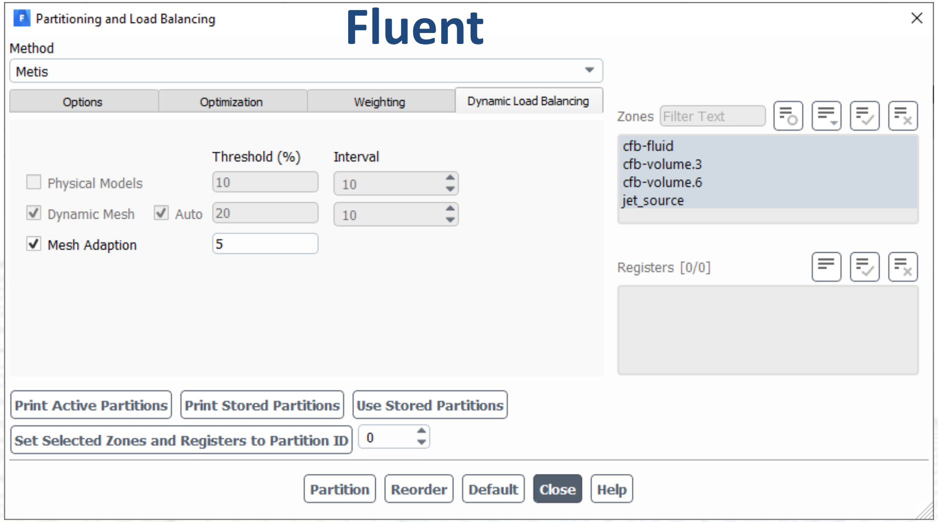
- Bias workload towards machines that are faster per core.
- Fluent has both manual methods and automatic methods for workload distribution.
- CFX can be balanced manually in run definition, or using CCL language.

```
> /par/part/set/load-distribution
()
load(1) [()] 1
load(2) [()] 1
load(3) [()] 1
load(4) [()] 1
load(5) [()] 2
load(6) [()] 2
load(7) [()] 2
load(8) [()] 2
```

CFX



Fluent

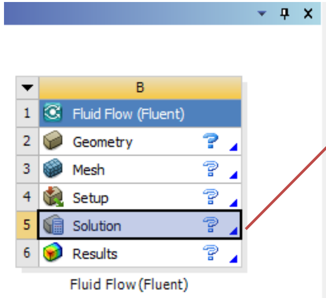


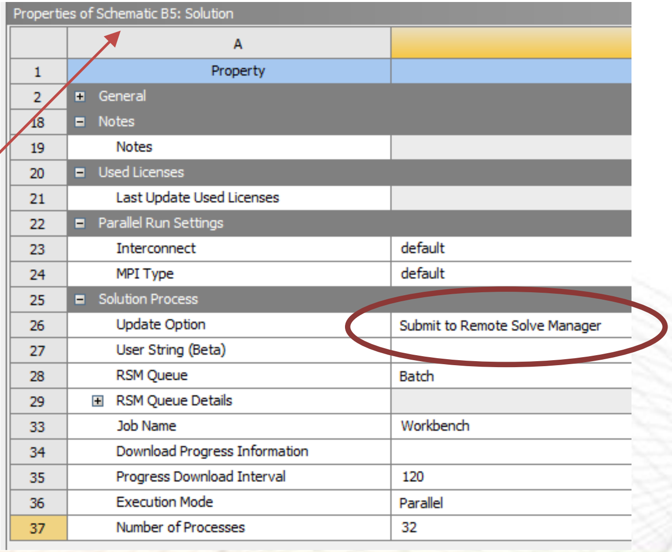
Job Schedulers & RSM

- Remote Solve Manager (RSM) is a background solving feature that allows solves to happen as a background task on your computer or others on your network.
- RSM can also integrate with the following Job Schedulers:
  - ANSYS RSM Cluster (ARC)
    - Windows and Linux, Free from ANSYS
  - PBS Pro (Linux)
  - Torque (Linux, very similar to PBS, just had official support dropped but still works)
  - Platform LSF (Linux)
  - SLURM (Linux, newly Supported)
  - Univa Grid Engine (Linux)
  - Windows HPC Server (Windows Server)
- Either the solvers or Workbench can be batch solved manually, allowing any scheduler.

Job Submission Techniques - RSM

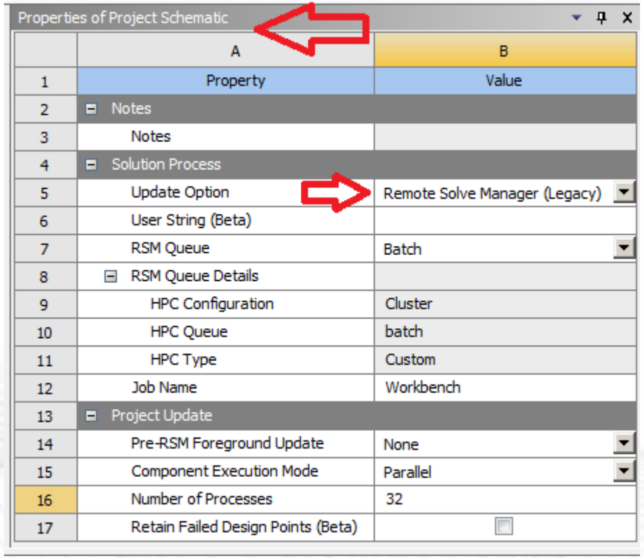
- RSM is easiest way to use remote resources, not necessarily the best.
- Using RSM you can:
  - Submit Simulation System from Workbench
  - Submit whole project from Workbench





Properties of Schematic B5: Solution

	A	B
1	Property	
2	General	
18	Notes	
19	Notes	
20	Used Licenses	
21	Last Update Used Licenses	
22	Parallel Run Settings	
23	Interconnect	default
24	MPI Type	default
25	Solution Process	
26	Update Option	Submit to Remote Solve Manager
27	User String (Beta)	
28	RSM Queue	Batch
29	RSM Queue Details	
33	Job Name	Workbench
34	Download Progress Information	
35	Progress Download Interval	120
36	Execution Mode	Parallel
37	Number of Processes	32

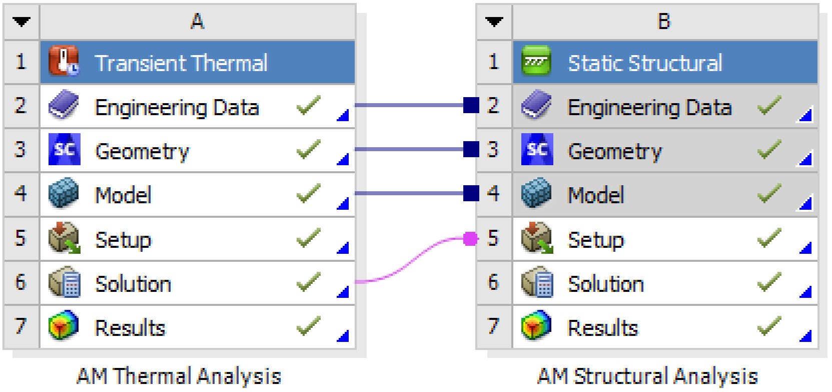


Properties of Project Schematic

	A	B
1	Property	Value
2	Notes	
3	Notes	
4	Solution Process	
5	Update Option	Remote Solve Manager (Legacy)
6	User String (Beta)	
7	RSM Queue	Batch
8	RSM Queue Details	
9	HPC Configuration	Cluster
10	HPC Queue	batch
11	HPC Type	Custom
12	Job Name	Workbench
13	Project Update	
14	Pre-RSM Foreground Update	None
15	Component Execution Mode	Parallel
16	Number of Processes	32
17	Retain Failed Design Points (Beta)	

Workbench Job Submission via RSM

- Submitting whole workbench project has significant benefit for interdependent systems.
- Normally System A work be solved, the results retrieved, System B would map those results, then be submitted, solved, and retrieved.
  - Lots of file transfer and user input required.
- Workbench project update allows a single submission that updates the whole project.
  - CFD + Structural, optimization, parametric models



Direct Batch Solve

- **Fluent:**

  fluent 3d -mpi=intel -t {cores} -g -cnf={hostfile} -i file.journal > solve.log

  3d, 3ddp, 2d etc.

-mpi	MPI selection (intel, ibm, Microsoft, intel2019)
-t	Threads (cores)
-cnf	List of hosts
-i	Input Journal File
-g	No Graphics
> solve.log	redirect output to log file for saving

Direct Batch Solve

- **CFX** (CFD):

```
cfx5solve -batch -def “%INPUT_DEF%” -par-dist $(cat hostfile | tr '\n' ',') -start-method
"Intel MPI Distributed Parallel”

-batch              batch Mode
-def                Job Definition File
-par-dist           Hosts (parallel distributed, see others)
-start-method       (MPI and local vs distributed)
-double             Double Precision
-size {x}           Multiplier for memory allocation from estimates
-large              Large Problem Solver
-ccl {file.ccl}              Command Language File for many advanced features
```

- All solvers have many optional arguments that should be checked and used, this is only a quick reference to start.
  - See ANSYS Help

Batch Workbench Job

- If you don't have RSM, you can still submit a whole workbench project as a single job.
  - Archive the project into a wbpz file (optional).
  - Move that file onto cluster
  - Extract using workbench (or just unzip it, wbpz is just a gzip file):

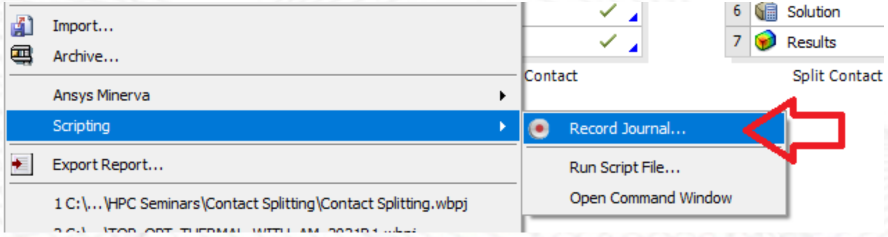
```
/ansys_inc/v202/Framework/bin/Linux64/runwb2 -B -E 'Unarchive(ArchivePath=r"CFD.wbpz",
ProjectPath=r"CFD.wbpj", Overwrite = True)'
```

  - Update the whole project:

```
/ansys_inc/v202/Framework/bin/Linux64/runwb2 -B -E 'Update(); Save(Overwrite=True)' -F "CFD.wbpj"
```

  - Learning Workbench scripting language is not covered here, but the "recording" feature is highly recommended.

```
-A addinsfile: Use addins file 'addins'
-B           : Run in batch
-C configfile: Use configuration 'config'
-D dataFile  : Import an application data file.
-E statement : Execute a journal statement at startup.  May be repeated.
-F WBProject : Open a Workbench project file (*.wbpj)
-I           : Run interactively
-L language  : Show UI in 'language'
-R replayfile: Use replay file 'replay'
-W workspace : Start UI with specified workspace
-X           : Run interactively & exit upon journal replay end
```



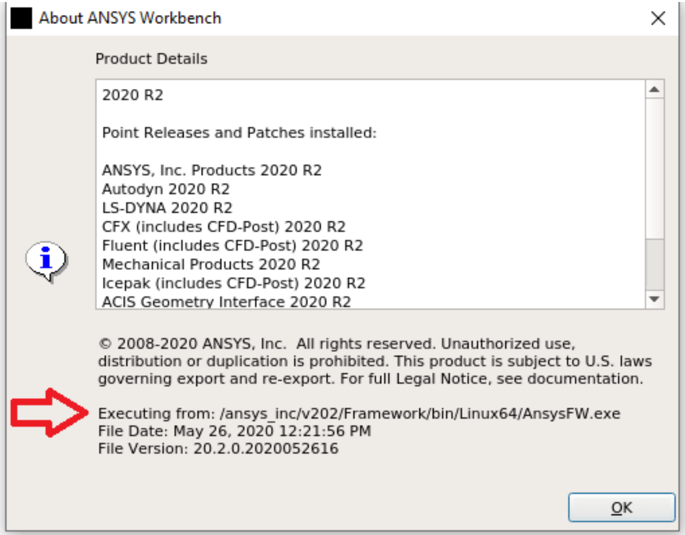
Interactive Cluster Jobs

- Many job schedulers allow interactive job sessions, including X forwarding (graphics).
- For PBS:

```
qsub -I -X -N Jobname -l select=16
```

- Then just open Workbench or CFD solver and use the GUI:

```
/ansys_inc/v202/Framework/bin/Linux64/runwb2
/ansys_inc/v202/fluent/bin/fluent
/ansys_inc/v202/CFX/bin/cfx5launch
```

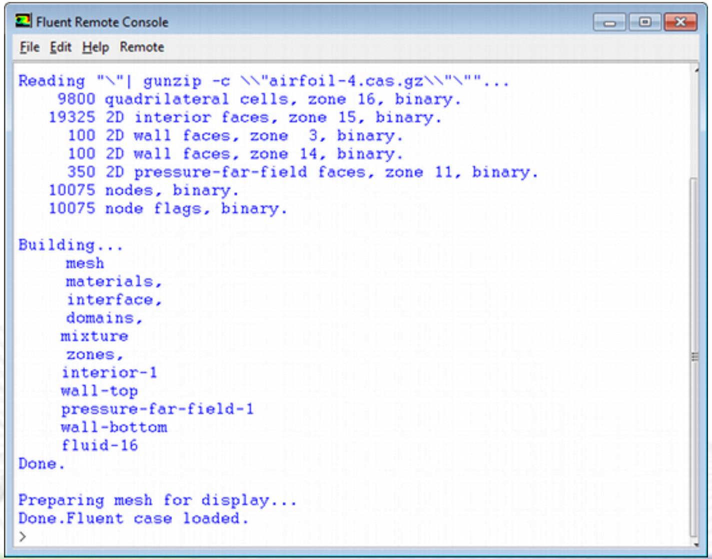


Other Solver Methods

- Fluent has hybrid Window GUI with Linux Solver mode (See Demo)
  - So does Icepak
- CFX has full job monitor for observing and editing batch jobs (see Demo)
- Fluent has solver “as a service” mode with remote console (not covered):
- Fluent has direct job scheduler integration with:

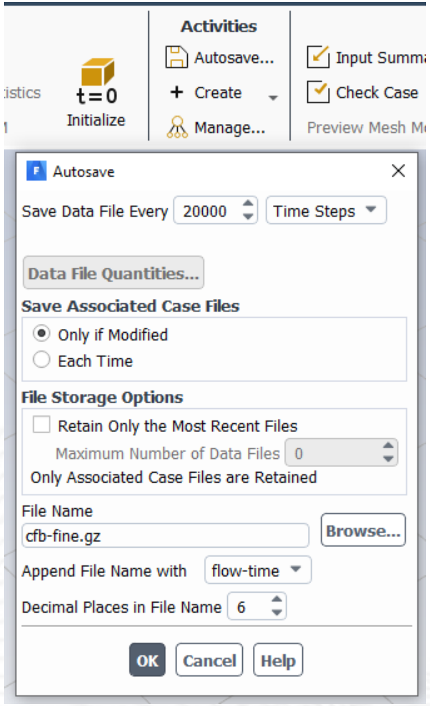
  -scheduler=pbs {pbs, lsf, sge}

  Relatively new since 2019 R2



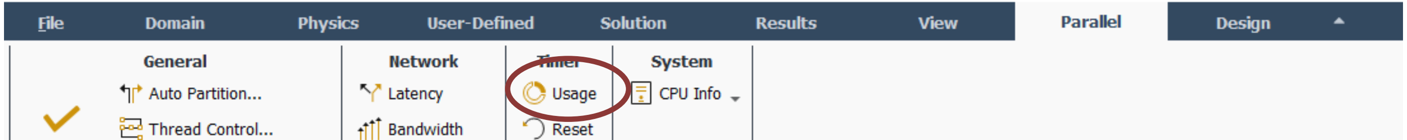
Restarts and Initialization

- CFD solves can be quite long, but are easy to restart.
- Don't forget to autosave incase of a crash
- Reload initial conditions from a previous solve to better initialize flow.
- Save mesh and partitioning setup for subsequent solves.



Measuring Performance

- Fluent has performance statistics option or /par/timer/usage command:



FLUENT

- You only need a few iterations to get a result you can extrapolate usually:
- Focus on solver time not total time:

CFX

```
CFD Solver wall clock seconds: 5.1197E+01
vs.
Total wall clock time: 6.299E+01 seconds
```

```
> /par/timer/usage

Performance Timer for 25 iterations on 36 compute nodes
  Average wall-clock time per iteration:                24.358 sec
  Global reductions per iteration:                        2074 ops
  Global reductions time per iteration:                  0.000 sec (0.0%)
  Message count per iteration:                         1050698 messages
  Data transfer per iteration:                         898.971 MB
  LE solves per iteration:                                   8 solves
  LE wall-clock time per iteration:                     16.900 sec (69.4%)
  LE global solves per iteration:                            9 solves
  LE global wall-clock time per iteration:               0.007 sec (0.0%)
  LE global matrix maximum size:                        1067
  AMG cycles per iteration:                             55.440 cycles
  Relaxation sweeps per iteration:                        4459 sweeps
  Relaxation exchanges per iteration:                        0 exchanges
  LE early protections (stall) per iteration:            0.000 times
  LE early protections (divergence) per iteration:       0.000 times

  Total wall-clock time:                               608.950 sec
```

Demonstrations

- Launch programs over SSH with PBS
- Fluent hybrid mode
- Monitor job in progress

Questions?

Thanks for listening!

Apickard@simutechgroup.ca

www.SimuTechGroup.com/Offices