rescale

Simulation cost analysis on Rescale HPC Cloud

Bosung Lee, Ph.D.
Director of Solutions Architect, Rescale
(bosung@rescale.com)

Contents

Backgrounds and Objectives

Backgrounds and objectives

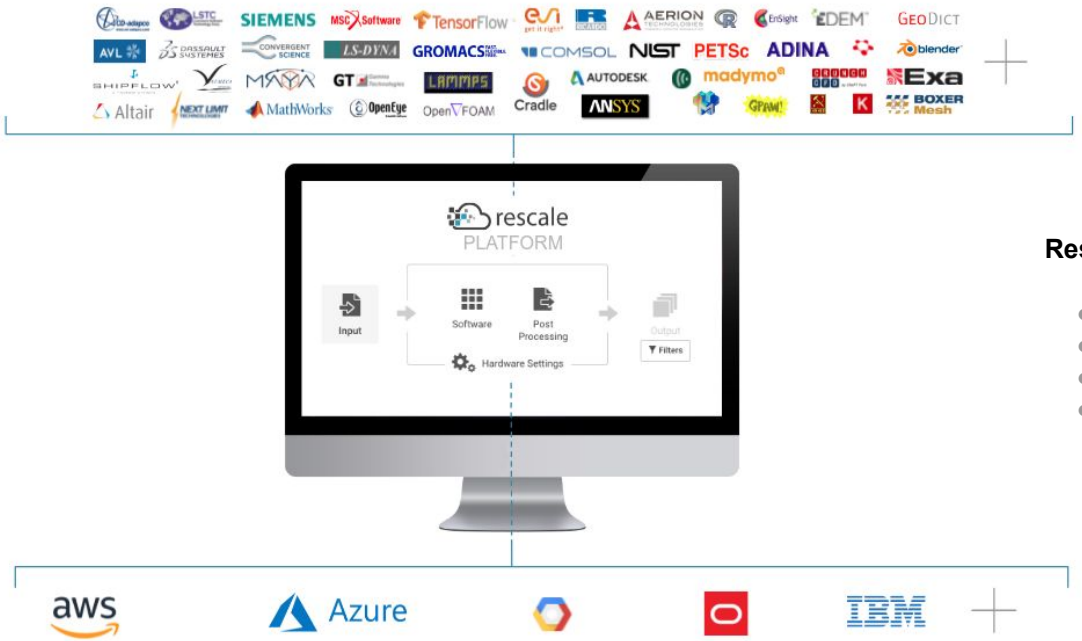
- Backgrounds
  - Simulation cost ($) = HPC hardware cost + simulation software license cost, and it depends on the simulation time
    - Limited simulation infrastructure (HW + SW) forces queuing the simulation
    - Simulation software license cost is more expensive than HPC hardware cost
  - HPC Cloud & open source software becomes the alternative for reducing the simulation cost

- Objectives
  - Simulation cost comparison using CFD / AI&ML on Rescale HPC Cloud
    - StarCCM+ & Fluent user cases on Rescale coretypes
    - DrivAer full car model on Fluent, StarCCM+, NextFOAM BARAM on the same coretype
    - ImageAI Object Detection on Rescale GPU coretypes
  - Discuss how to reduce the total cost of simulation

Introduction of Rescale HPC Cloud

All simulations on single HPC platform

- Fully managed HPC Cloud platform with public cloud hardware and pre-installed simulation software

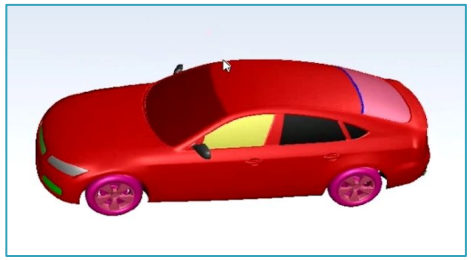


Rescale ScaleX Platform

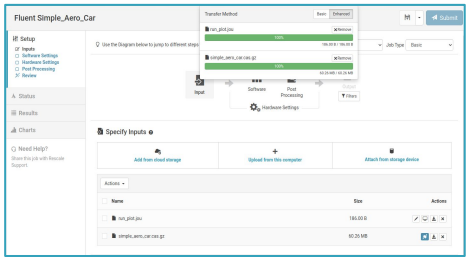
- Enterprise big compute
- Innovation acceleration
- User-first SaaS platform
- Security and admin controls

Simulation workflow of Rescale HPC cloud

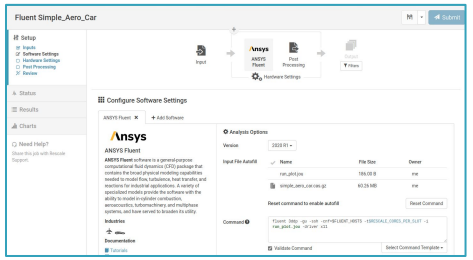
- Upload input files -> Software settings -> Hardware settings -> Submit job



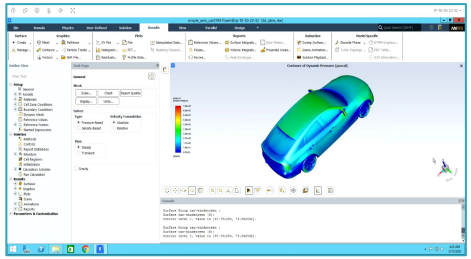
Model setup



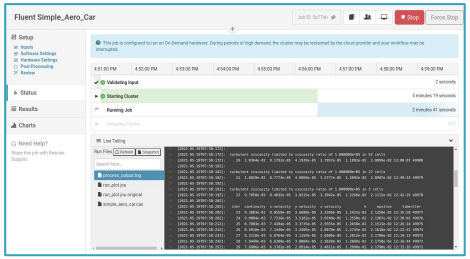
Input file upload



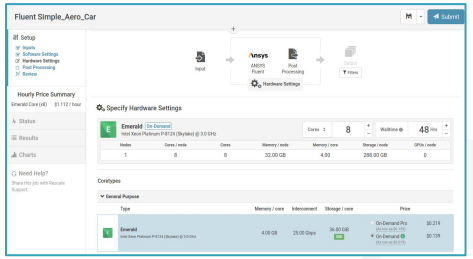
Software / license setup



Visualization on Rescale VDI



Solving



Coretype / #of cores setup

Pre-configured simulation optimized coretypes

General Purpose

Type	Memory / core	Interconnect	Storage / core
Emerald Intel Xeon Platinum P-8124 (Skylake) @ 3.0 GHz	4.00 GB	25.00 Gbps	36.00 GiB SSD
Luna 2nd Generation Intel Xeon Scalable Processors (Cascade Lake) @ 3.6 GHz	4.00 GB	25.00 Gbps	36.00 GiB SSD
Ferrite Intel Xeon Platinum 8168 (Skylake) @ 2.7 GHz	4.00 GB	10.00 Gbps	16.00 GiB SSD
Willow AMD EPYC 7002 series @ 2.3 GHz	4.00 GB	20.00 Gbps	12.00 GiB
Spruce AMD EPYC 7452 (Rome) @ 2.4 GHz	8.00 GB	30.00 Gbps	16.00 GiB SSD
Sequoia AMD EPYC 7452 (Rome) @ 2.4 GHz	14.00 GB	30.00 Gbps	32.00 GiB SSD
Ruby Intel Xeon Platinum 8175M (Skylake) @ 2.5 GHz	16.00 GB	25.00 Gbps	75.00 GiB SSD

High Memory

Type	Memory / core	Interconnect	Storage / core
Ruby Intel Xeon Platinum 8175M (Skylake) @ 2.5 GHz	16.00 GB	25.00 Gbps	75.00 GiB SSD
Oak AMD EPYC 7000 series @ 2.5 GHz	16.00 GB	20.00 Gbps	24.00 GiB
Moonstone Intel Xeon Platinum 8175M (Skylake) @ 2.5 GHz	16.00 GB	25.00 Gbps	40.00 GiB SSD
Diamond Intel Xeon Platinum 8151 @ 3.4 GHz	16.00 GB	25.00 Gbps	75.00 GiB SSD
Topaz Intel Xeon E7-8880 v3 (Haswell) @ 2.3 GHz	30.50 GB	25.00 Gbps	56.25 GiB SSD

High Interconnect

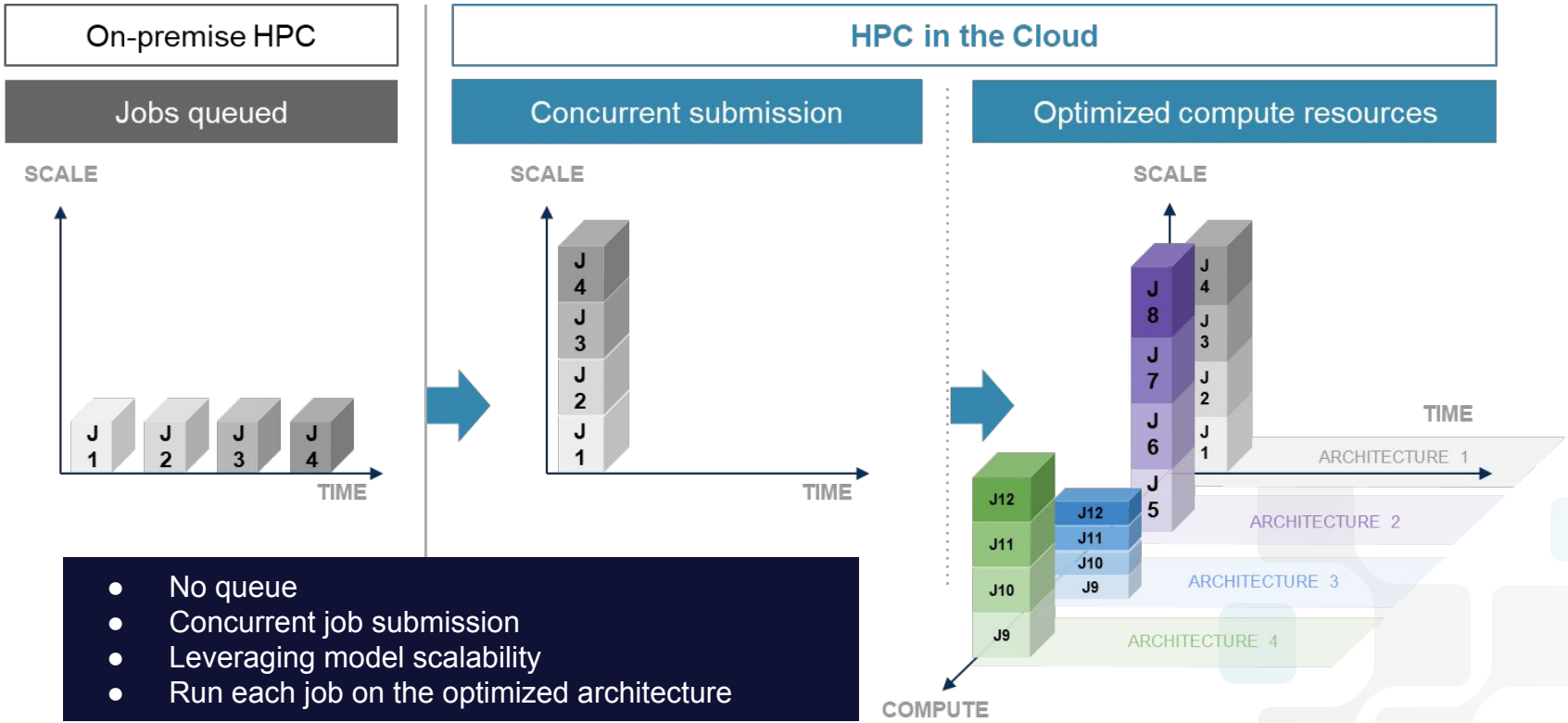
Type	Memory / core	Interconnect	Storage / core
Catseye Intel Xeon Platinum P-8124 (Skylake) @ 3.0 GHz	5.25 GB	100.00 Gbps	36.00 GiB SSD
Carbon Intel Xeon Platinum 8168 (Skylake) @ 2.7 GHz	8.00 GB	100.00 Gbps	15.91 GiB SSD
Amber v2 AMD EPYC 7742 (Rome) @ 2.4 GHz	4.00 GB	200.00 Gbps	4.00 GiB SSD
Amber AMD EPYC 7551 (Naples) @ 2.0 GHz	4.00 GB	100.00 Gbps	11.67 GiB SSD
Graphite V2 Intel Xeon Scalable (Skylake) processors @ 3.1 GHz	16.00 GB	100.00 Gbps	1250.00 GiB SSD

GPU

Type	Memory / core	Interconnect	Storage / core
Aquamarine v3 Intel Xeon E5-2690 v4 (Broadwell) CPUs @ 2.6 GHz, Tesla V100	18.67 GB	10.00 Gbps	122.83 GiB SSD
Rhodium NVIDIA V100-enabled w/ NVLink, Intel(R) Xeon(R) CPU E5-2686 v4 (Broadwell) @ 2.3 GHz, Tesla V100	15.25 GB	25.00 Gbps	37.50 GiB SSD
Dolomite NVIDIA V100-enabled w/ NVLink, Intel(R) Xeon(R) CPU E5-2686 v4 (Broadwell) @ 2.3 GHz, Tesla V100	15.25 GB	25.00 Gbps	75.00 GiB SSD
Obsidian Intel Xeon E5-2676 v3 (Haswell), Tesla K80	30.50 GB	25.00 Gbps	37.50 GiB SSD

Coretypes are continuously on-boarded

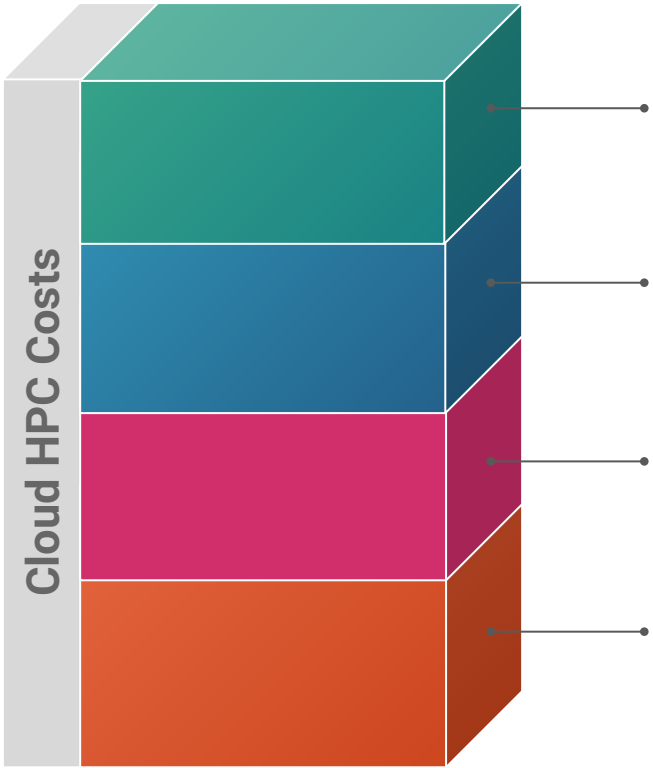
Faster simulation and shorter turnaround time on Rescale



- No queue
- Concurrent job submission
- Leveraging model scalability
- Run each job on the optimized architecture

Simulation cost on HPC Cloud

Key cost factors in simulation cost on HPC Cloud



Software / Application Licensing
Commercial ISV Licensing

Upfront Implementation
In-house IT (FTE), Systems Integrators, Outsourced Support, Cloud provider support

Ongoing Support & Maintenance
Utilities, Operations, Application Support, Cloud Provider Support, e.g. Security & Compliance Standards & Updates

Cloud Infrastructure
Compute hardware, memory, networking, storage

Simulation software licensing cost

- Perpetual license cost depends on number of cores / number of users or concurrent simulation session
  - Ansys : Solver (4 cores) + number of HPC Packs
    - 1 HPC Pack : 8 cores / 2 HPC Packs : 32 cores / 3 HPC Packs : 128 cores
  - StarCCM+ Power Session : no limitation of the number of cores

- On-demand license cost depends on the number of cores and/or license usage time
  - Ansys Elastic Licensing : AEC = 20 (solver) + INT(5 x $nCores^{0.57}$), 1 AEC = $1.4 / hour (estimated)
  - StarmCCM+ Power-on-Demand : $32.5 / hour (estimated)

- Perpetual license cost should be calculated by the license usage time
  - License cost / core-hour = Yearly cost / (24 * 365 hours) / utilization (%) / number of cores
    - $100,000 / year for 100 cores and HPC utilization is 60 % : $0.1903 / core-hour

- **License cost is mainly dependent on the simulation time**
  - **As the simulation time decreases, the license cost decreases also**

Typical coretype prices on Rescale HPC Cloud

Coretype	Spec	mem / core (GB)	storage / core (GB)	Interconnect (Gbps)	cores / node	ODE ($/hour)		ODP ($/hour)	
						Total Price	Node price	Total Price	Node price
Emerald	Xeon Platinum P-8124 (Skylake) @ 3.0 GHz	4	36	25	36	**0.1055**	3.798	**0.1473**	5.3028
Luna	Xeon SP (Cascade Lake) @ 3.6 GHz	4	36	25	48	**0.1055**	5.064	**0.1473**	7.0704
Moonstone	Xeon Platinum 8175M (Skylake) @ 2.5 GHz	16	40	25	48	**0.1158**	5.5584	**0.2083**	9.9984
Catseye	Intel Xeon Platinum P-8124 (Skylake) @ 3.0 GHz	5.25	36	100 (EFA)	36	**0.1055**	3.798	**0.1473**	5.3028
Pyrite Max	Xeon SP (Cascade Lake) Processors @ 3.8 GHz	8	80	100 (EFA)	24	**0.1757**	4.2168	**0.1757**	4.2168
Ferrite	Xeon Platinum 8168 (Skylake) @ 2.7 GHz	4	16	10	36	**0.0833**	2.9988	**0.1361**	4.8996
Calcite	Xeon Platinum 8272CL (Cascade Lake) @ 2.5 GHz	8	75	30	32	**0.0956**	3.0592	**0.179**	5.728
Chromium	Xeon Platinum 8272CL (Cascade Lake) @ 2.5 GHz	15.75	75	30	32	**0.1092**	3.4944	**0.213**	6.816
Carbon	Xeon Platinum 8168 (Skylake) @ 2.7 GHz	8	15.91	100 (IB)	44	**0.079**	3.476	**0.1373**	6.0412
AmberV2	AMD EPYC 7742 (Rome) @ 2.4 GHz	4	8	200 (IB)	120	**0.0567**	6.804	**0.0769**	9.228
Ammonite	AMD EPYC 7742 (Rome) @ 2.4 GHz	4.3	8.57	200 (IB)	112	**0.0579**	6.4848	**0.0796**	8.9152
Apatite	AMD EPYC 7742 (Rome) @ 2.4 GHz	5.45	10.91	200 (IB)	88	**0.0627**	5.5176	**0.0904**	7.9552
Jasper	AMD EPYC 7742 (Rome) @ 2.4 GHz	8	16	200 (IB)	60	**0.0733**	4.398	**0.1138**	6.828
Peridot	AMD EPYC 7742 (Rome) @ 2.4 GHz	10.91	21.82	200 (IB)	44	**0.0854**	3.7576	**0.1407**	6.1908
Rozenite	AMD EPYC 7763 (Milan) @ 2.4 GHz	3.75	16	200 (IB)	120	**0.0676**	8.112	**0.073**	8.76
Limonite	AMD EPYC 7763 (Milan) @ 2.4 GHz	4.67	20	200 (IB)	96	**0.0778**	7.4688	**0.0813**	7.8048
Hematite	AMD EPYC 7763 (Milan) @ 2.4 GHz	7	30	200 (IB)	64	**0.0939**	6.0096	**0.1019**	6.5216
Edenite	AMD EPYC 7763 (Milan) @ 2.4 GHz	14	60	200 (IB)	32	**0.1478**	4.7296	**0.1638**	5.2416
Albite	AMD EPYC 7763 (Milan) @ 2.4 GHz	28	120	200 (IB)	16	**0.2467**	3.9472	**0.2875**	4.6

Above prices are sample, and these prices are variable on the contract size, date, and regions

Simulation cost comparison on Rescale HPC Cloud

StarCCM+ simulation cost

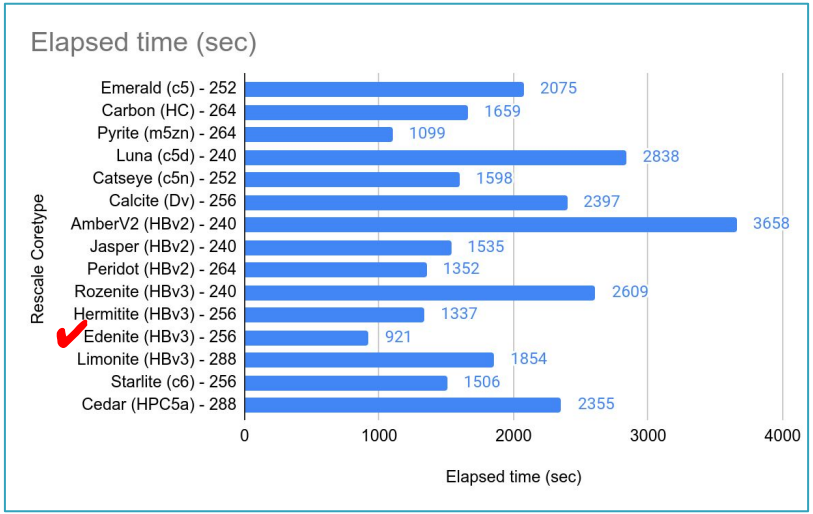
Ansys Fluent simulation cost

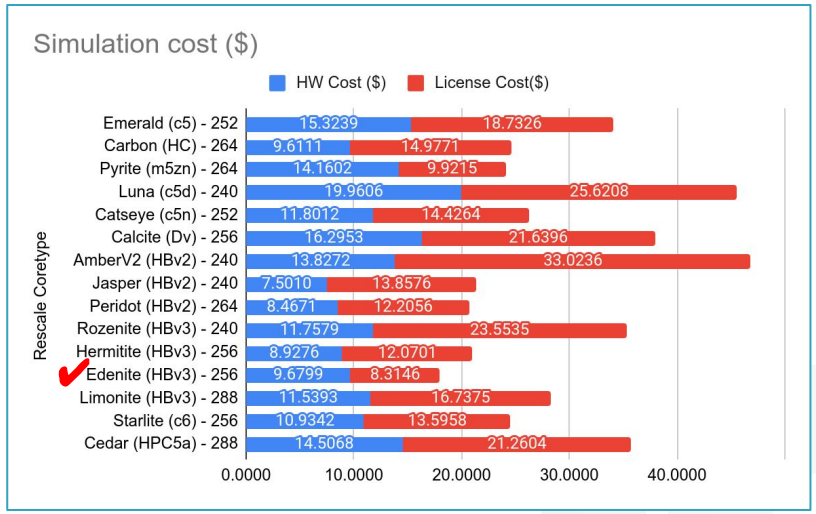
DrivAer full car model using Fluent/StarCCM+/BARAM

Cost of the object detection on GPU coretypes

StarCCM+ simulation cost
- Cells : 28M / 15.06.008 Double / Iteration = 200

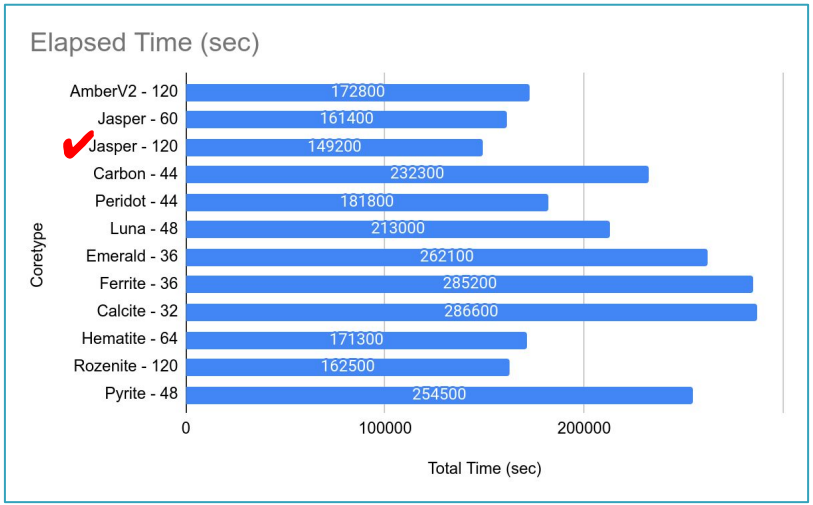
- Jasper is the minimum hardware cost : $7.5010
- Edenite shows the minimum total simulation cost (hardware + license cost) : $17.9945
    - The simulation cost of StarCCM+ PoD depends mainly on the simulation time

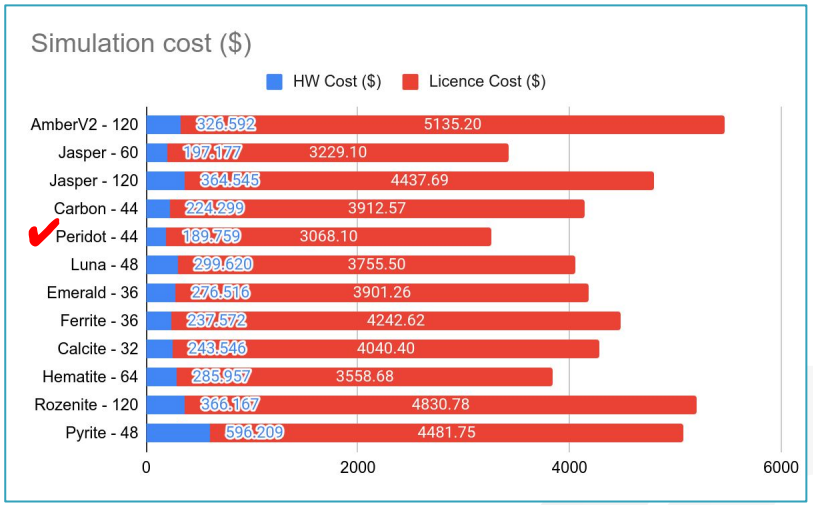




Ansys Fluent simulation cost
- Unsteady 2D supersonic nozzle (flow time = 1.2 sec)

- Jasper 120 core shows the fastest elapsed time : 149,200 sec
- Peridot 44 cores' hardware cost is the minimum : $189.759
- The minimum simulation cost is Peridot 44 cores' $3257.895
    - Ansys Fluent simulation cost depends on the simulation time and the number of cores





DrivAer full car model
- StarCCM+ / Fluent / BARAM (OpenFOAM based Solver)

- Simulation conditions

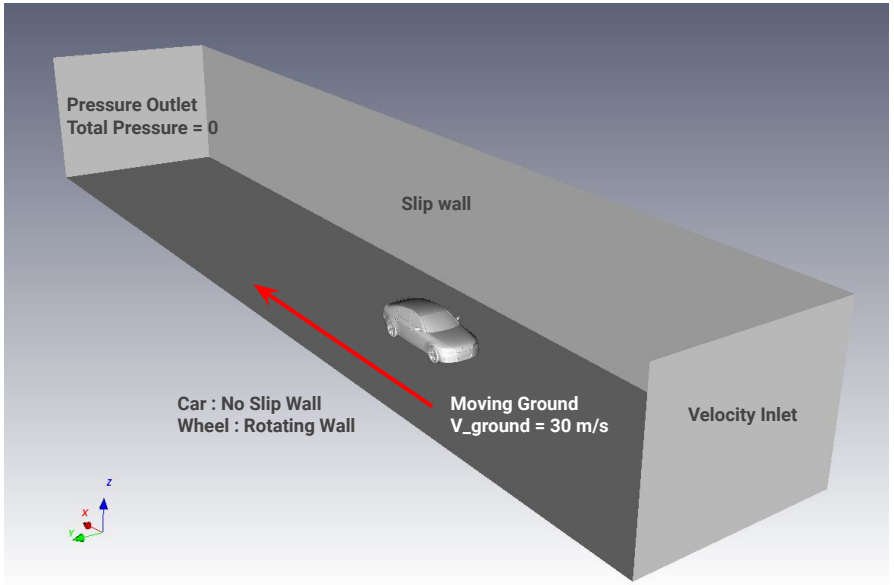
	Version	Solution scheme	Spatial Discretization	Relaxation Factors	Flow Time
BARAM	6.2	simpleNFoam Transient	Momentum : 2nd Turbulent : 1st	Pressure : 0.3 Momentum : 0.3 Turbulence : 0.7	t = 2.0 Δt = 0.001 sub iter = 10
Fluent	2020R1	SIMPLE Unsteady	Momentum : 2nd Turbulent : 1st	Pressure : 0.3 Momentum : 0.3 Turbulence : 0.7	t = 2.0 Δt = 0.001 sub iter = 10
StarCCM+	15.06.008 Double	Segregated Flow Implicit Unsteady	Momentum : 2nd Turbulent : 1st	Pressure : 0.3 Momentum : 0.3 Turbulence : 0.7	t = 2.0 Δt = 0.001 sub iter = 10

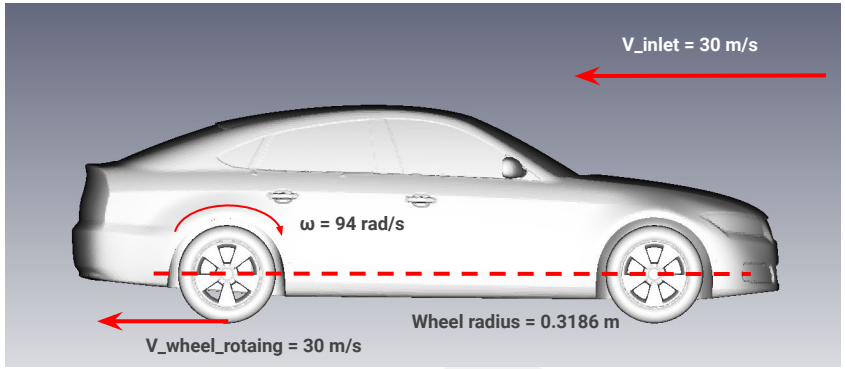
- Hardware specifications

Rescale Coretype	Processor	#of Cores / node	Memory	Interconnect	Hourly price / node
Jasper	AMD EPYC 7742 (Rome) @ 2.4 Ghz	60 / 120 / 180 / 240	8 GB /core 488 GB / node	200 Gbps Infiniband	$4.398

DrivAer full car model
- Boundary conditions / Flow conditions

- Same boundary conditions are applied to BARAM / Fluent / StarCCM+

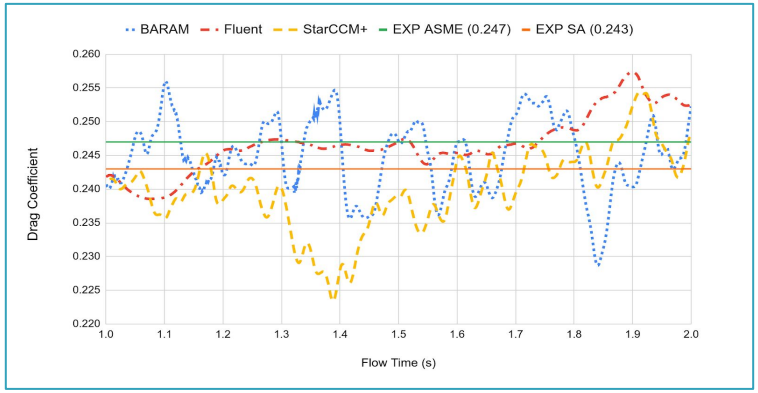


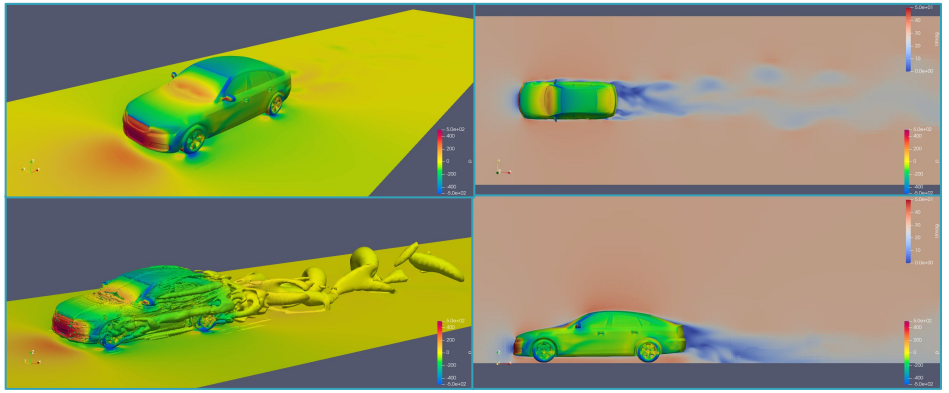


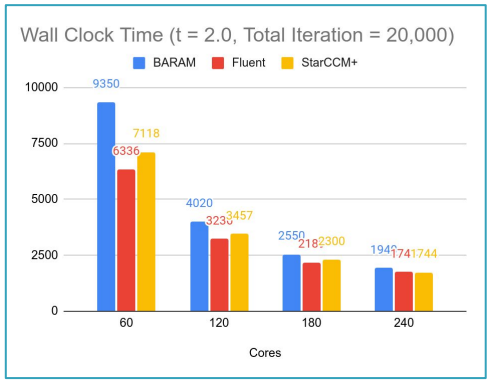
ρ (density) : 1.205
ɲ (viscosity) : 1.58e-5
Prt : 0.85
Turbulent intensity : 1%
Viscosity Ratio : 10

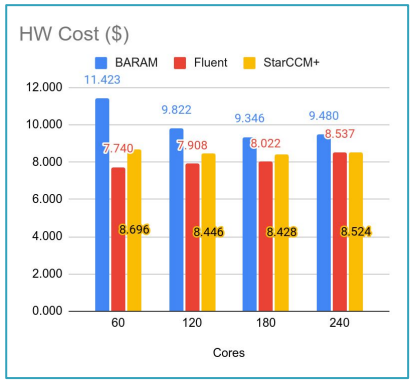
DrivAer full car model

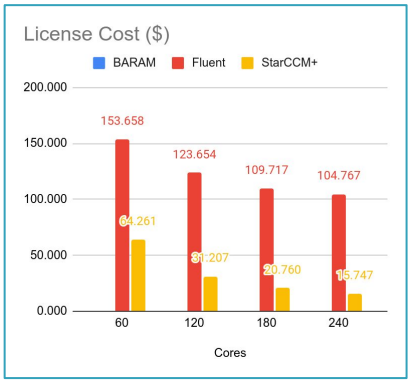
- StarCCM+ / Fluent / BARAM (OpenFOAM based Solver)

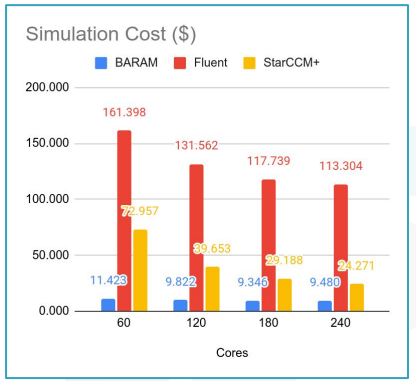












DrivAer full car model
- StarCCM+ / Fluent / BARAM (OpenFOAM based Solver)

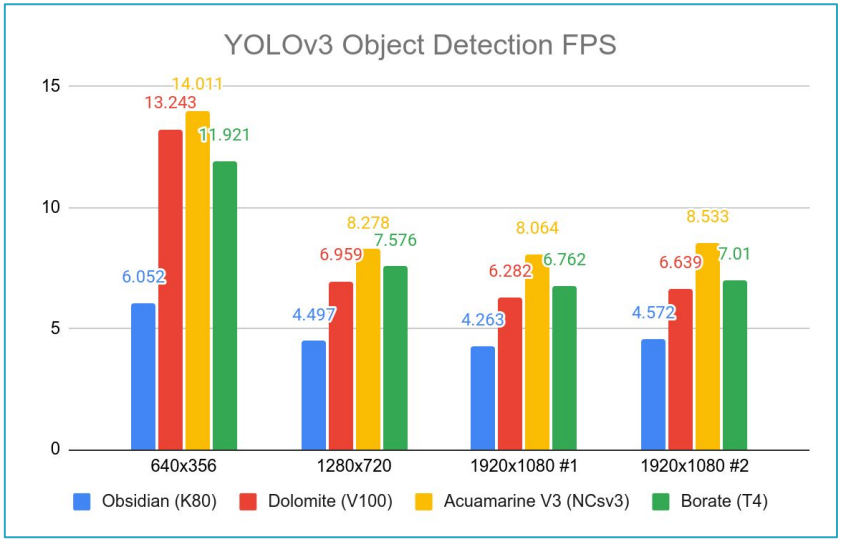
- NEXTfoam BARAM, Fluent and StarCCM+ show similar drag coefficients compared to experiments
- Elapsed time in NEXTfoam BARAM is slightly longer than Fluent and StarCCM+
  - Rescale cloud cost depends on the elapsed time
- Simulation license cost is very expensive than cloud hardware cost
  - StarCCM+ Power-session & Power-on-Demand license have no limit in the number of cores
  - Fluent AEC license cost requires the complex cost calculation depends on the problem
- Total simulation cost decreases as the number of cores increases
  - Reduced time using more number of cores can reduce the license cost
- Open source simulation software requires only hardware cost

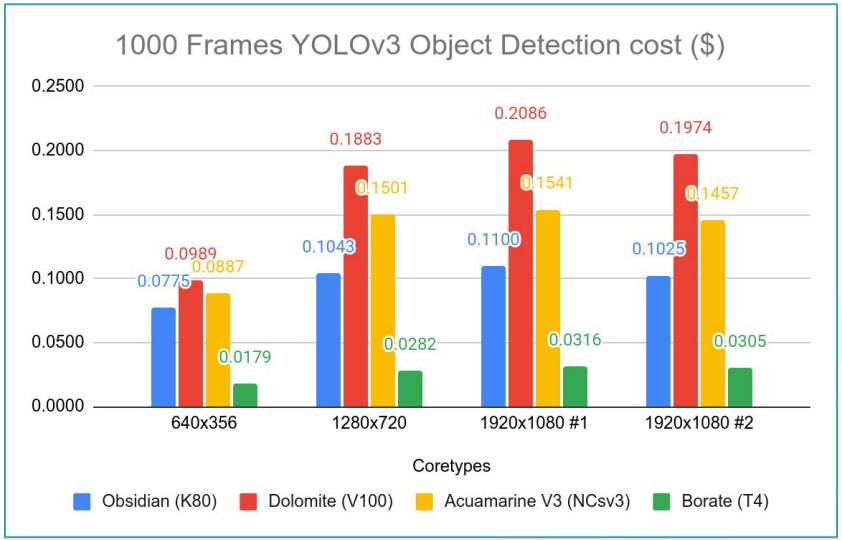
References of DrivAer full car model

[1] DrivAer Model : https://www.epc.ed.tum.de/en/aer/research-groups/automotive/drivaer/, 2021.11
[2] Drivaer validation case, Wolf Dynamics: http://www.wolfdynamics.com/validations/drivAer/tut_drivaer.pdf, 2021.11
[3] R. Yazdani. Steady and Unsteady Numerical Analysis of the DrivAer Model. Chalmers University of Technology, Master Thesis, 2015
[4] Experimental Comparison of the Aerodynamic Behavior of Fastback and Notchback DrivAer Models. SAE 2014-01-0613.

ImageAI object detection for the real time video stream

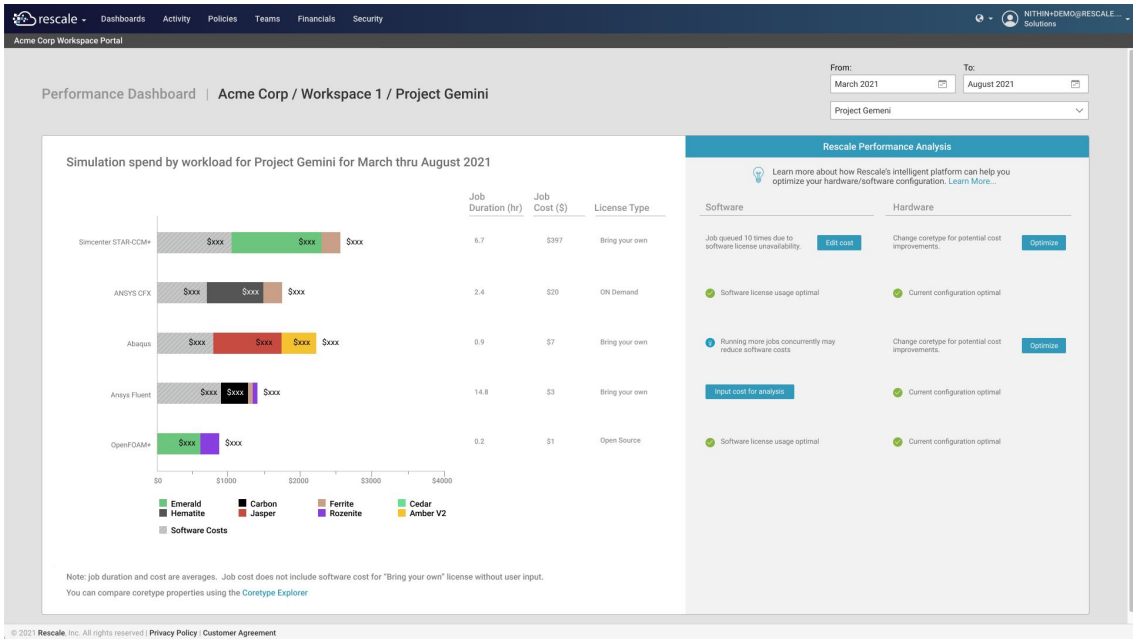
- Open source based AI/ML simulation cost depends on the hardware cost
- The frame rate of the detection in the real time video streams is dependent on the GPU&CPU performance
  - The frame rates of AcuamarineV3 (6 CPU with one V100 GPU) is the highest
  - Detection cost in the Borate (2 CPU with one T4 GPU) is the lowest





Performance dashboard and performance optimization

- Performance Dashboard and Optimization help Rescale customers keep HPC spend optimal and aligned to business objectives



- Get better visibility into the highest spend workloads
- Gain insights on how to optimize the spend by selecting the right infrastructure
  - License type used for the software (either Bring Your Own or On-Demand)
  - Software licensing insights - depends on licensing model
  - Hardware insights

Concluding Remarks

Concluding Remarks

- Simulation cost ($) is composed of the HPC hardware cost, simulation software license cost
  - Simulation software license cost is much higher than HPC hardware cost
  - By selecting the optimized hardware, the simulation cost can be optimized
    - As the simulation time decreases, the license cost decreases also

- According to the license model, the simulation cost is variable
  - StarCCM+ Power-session & Power-on-Demand license have no limitation in the number of cores
  - Ansys license requires the complex calculation depends on the number of cores and simulation time
  - Proper coretype selection for each simulation software should be required

- Open source simulation software requires only hardware cost

- Simulation cost optimization possible through the flexible coretype selection on Rescale HPC cloud