PREPARING THE CFD SOFTWARE CODA FOR EXTREME SCALE

Michael Wagner, German Aerospace Center (DLR), Institute of Software Methods for Product Virtualization

HiPEAC 2025, Tackling software exascale challenges: the Centres of Excellence in High Performance Computing perspective, Jan 22nd 2025

DLR

Background and Motivation

Numerical Simulation – Key Enabler for Future Aircraft Design

Future aircraft

- Goals: drastic reduction CO_2, NO_x and noise emissions
- Step changes in aircraft technology and new designs

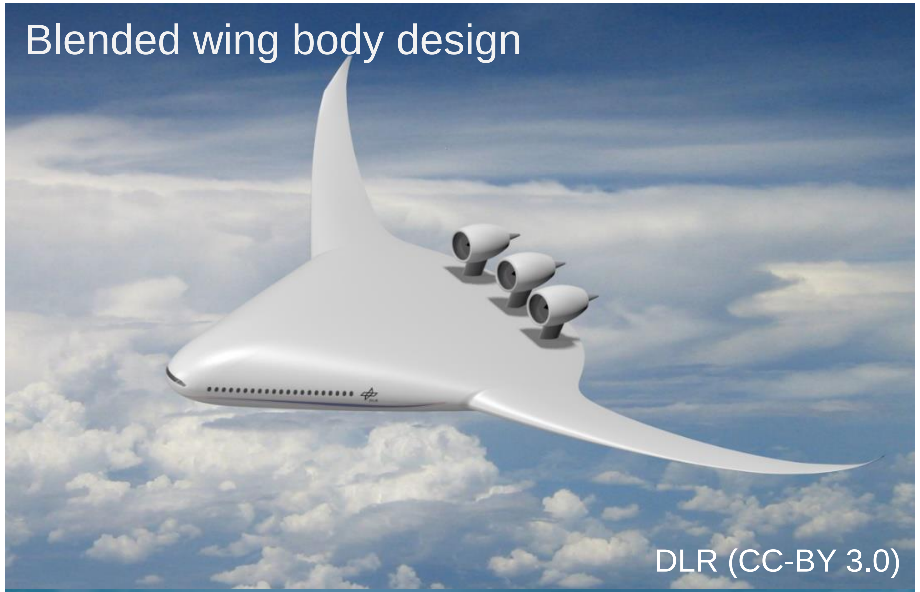
High-fidelity CFD methods indispensable

- Flight characteristics dominated by non-linear effects
- Reliable insight to new aircraft technologies
- High-fidelity CFD simulation of aircraft aerodynamics

Efficient linear system solving important

- CFD requires solving of large linear equation systems
- Linear systems solving makes up majority of time

Further improvement of simulation capabilities, computational efficiency and scalability necessary.



CODA Software Environment

CODA CFD Software

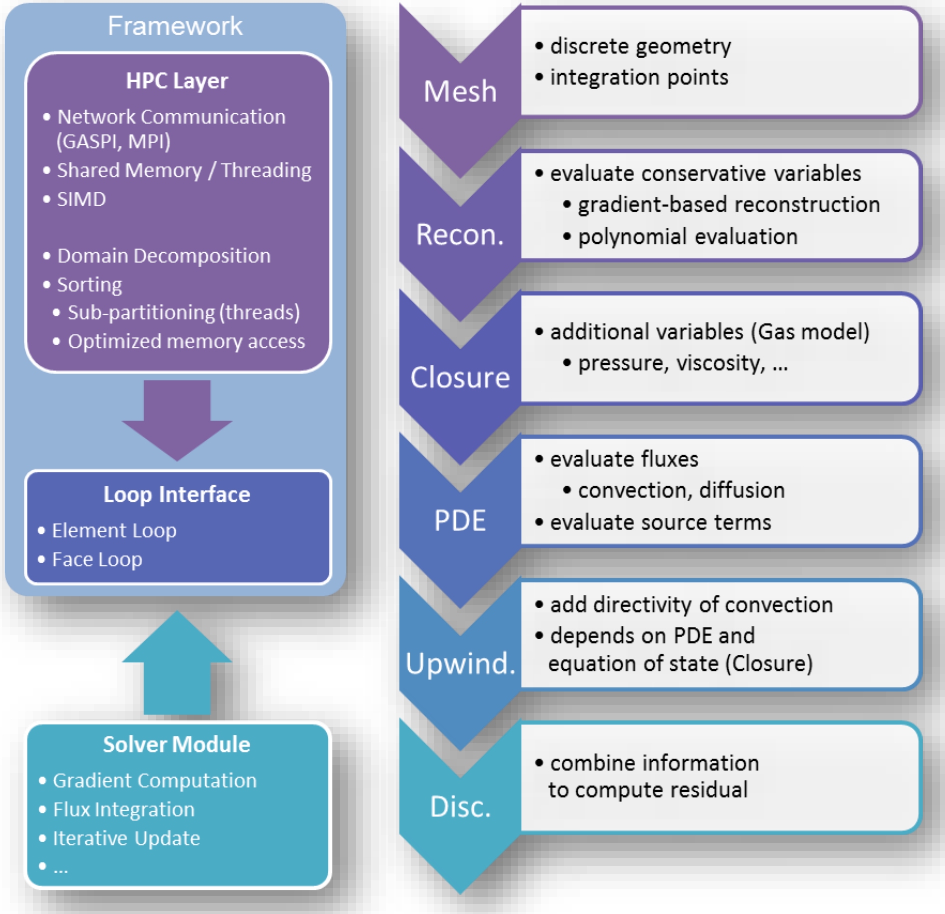
- Collaboration of ONERA, DLR and Airbus*
- 2nd order Finite Volume method and higher-order DG for unstructured grids and compressible flows
- Hybrid parallelization (MPI/GASPI + OpenMP/threads) with overlap of communication & computation
- Seamless integration into multi-disciplinary simulations

FlowSimulator

- Provides plug-ins for all steps of a full aircraft simulation
- FSMesh class for unified data exchange among plug-ins

Spliss: Sparse Linear Systems Solver

- Linear systems solving for implicit methods
- Full HPC support: MPI/GASPI, Threads, SIMD, GPUs



*CODA is the computational fluid dynamics (CFD) software being developed as part of a collaboration between the French Aerospace Lab ONERA, the German Aerospace Center (DLR), Airbus, and their European research partners. CODA is jointly owned by ONERA, DLR and Airbus.

EXCELLERAT (P2 2023 – 2026)

The European Centre of Excellence for Engineering Applications (P1 2019 – 2022)

Preparing European engineering for exascale computing

- 15 partners
- 7 use cases: Alya, AVBP, CODA, m-AIA, Neko, Flew, OpenFoam
- Aerospace & Energy; CFD & Combustion

Cooperation with European engineering and HPC community

- Expertise from other leading-edge engineering codes with similar challenges and problems
- Access to the largest HPC systems in Europe
- Early access and experiences with new hardware and trends

EXCELLERAT Project Targets

Evaluate and demonstrate CODA’s and FlowSim’s readiness for exascale computing

- Continuous evaluation (and analysis) of CODA/FlowSimulator scalability improvements
- Large scale demonstrator: large mesh + large system
- Evaluation of new systems and emerging technologies

Use case: external aircraft aerodynamics

- Airflow for steady forward flight at subsonic speed
- Reynolds-averaged Navier-Stokes equations (RANS) with Spalart-Allmaras turbulence model (SA-neg)


DLR (CC-BY 3.0)

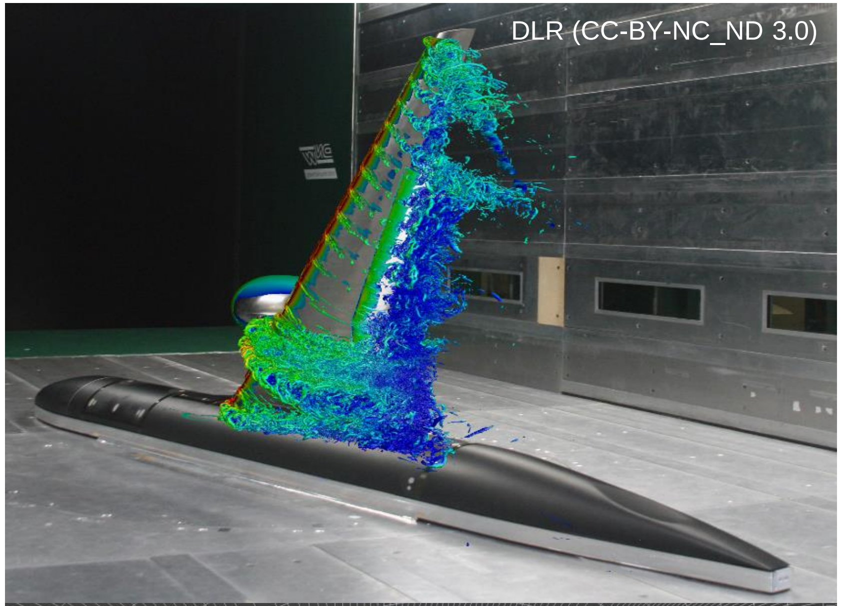
EXCELLERAT Use Case and Inputs

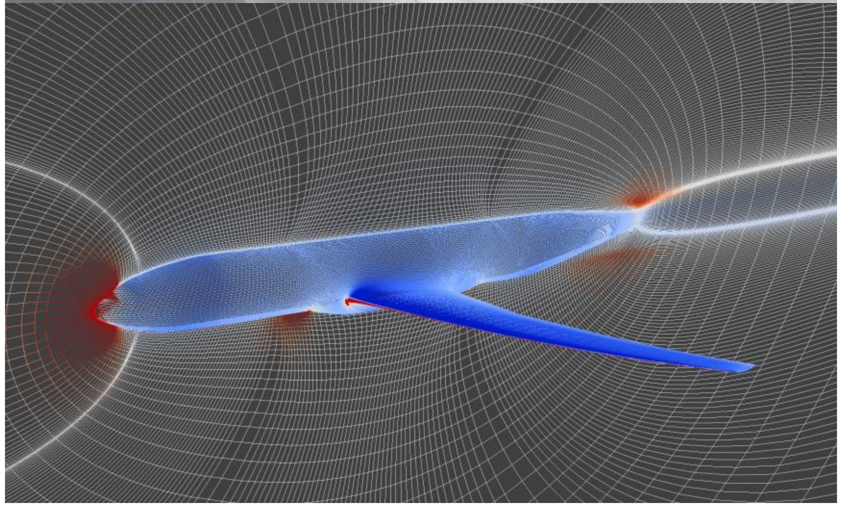
Strong and weak scaling use case

- NASA Common Research Model CRM (wing-body configuration)
- Mesh set with 3, 10, **24**, 81 and 192 million elements
- Practical size to see large scaling effects at smaller scales*
- Public, widely used and well-studied (also experimentally)

Capability demonstrator

- Demonstrate capabilities for big meshes on big systems
- Mesh with about 1 – 5 billion elements
- Upcoming European (pre-)exascale systems





* within the range of available resources at DLR, i.e. up to 32/64k cores

Preparing for Extreme Scale (Exascale)

Three main challenges

- Efficiency
  - Scale as efficiently as possible on current systems
    → Estimate performance for large systems
- Scalability
  - Support large meshes in FlowSimulator framework
  - Support large core counts in the entire workflow
    → Overcome any hard scalability limits
- Heterogeneity
  - Readiness for a variety of different systems (CPU, GPU, …)
    → Support offloading to Nvidia and AMD GPUs

Preparing for Extreme Scale (Exascale)

Three main challenges

- Efficiency
  - Scale as efficiently as possible on current systems
    → Estimate performance for large systems
- Scalability
  - Support large meshes in FlowSimulator framework
  - Support large core counts in the entire workflow
    → Overcome any hard scalability limits
- Heterogeneity
  - Readiness for a variety of different systems (CPU, GPU, …)
    → Support offloading to Nvidia and AMD GPUs

The CARA and CARO HPC Systems at DLR

CARA (AMD Naples architecture):

- 2168 nodes with 128 GB DDR4 (2666 MHz)
- 2x AMD Epyc 7601 (32 cores; 2,2 GHz) per node
- 145.920 cores delivering 1.7 TFLOP/s
- Infiniband HDR network

CARO (AMD Rome architecture)

- 1354 nodes with 256 GB DDR4 (3200 MHz) RAM
- 2x AMD Epyc 7702 (64 cores; 2,0 GHz) per node
- 174592 cores delivering 3.5 TFLOP/s
- Infiniband HDR network

DLR (CC-BY 3.0)

DLR (CC-BY 3.0)

The CARA and CARO HPC Systems – Comparison

Number of cores:

- CARO (AMD Epyc 7702) has 2x cores (128 vs. 64 per node)

Cache:

- CARO has 4x last-level cache (256 MiB vs 64 MiB), i.e. twice as much per core.
- 16 vs. 8 NUMA domains
- 3 NUMA distances (on die, on socket, 2nd socket)
- 4 cores per die share L3

Memory access:

- 8 memory channels and memory controllers
- Memory controllers: 3200MHz (CARO) vs. 2666MHz (CARA)
- CARO has 1.2x memory bandwidth (191 GiB/s vs. 159 GiB/s) for twice the number of cores.

DLR (CC-BY 3.0)

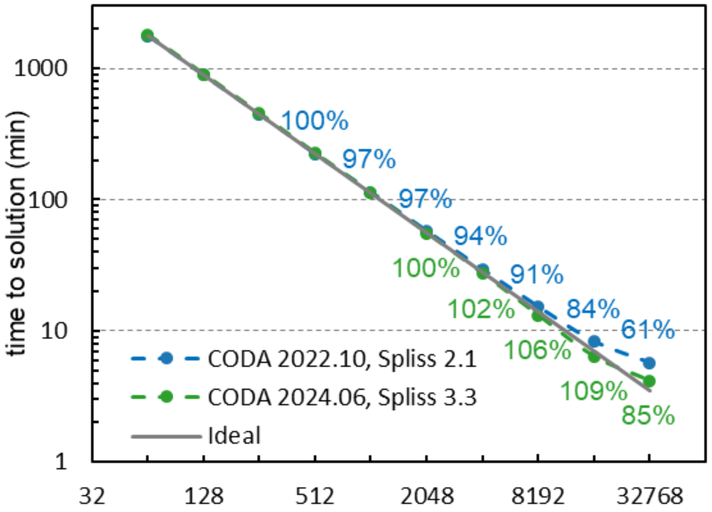
DLR (CC-BY 3.0)

Efficiency: Scalability on Current Systems (CARA)

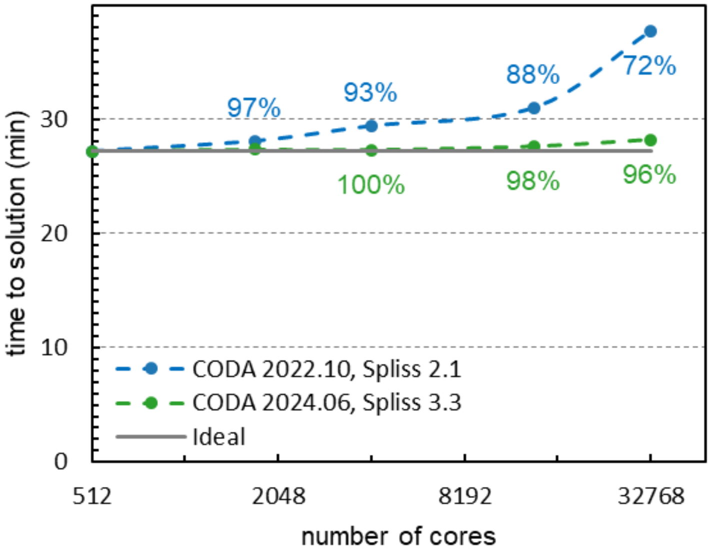
Scalability assessment on DLR's production system CARA

- Strong scaling (CRM, fixed problem size, 24M elements):
  - Scaling from 1 – 512 nodes (largest available partition)
  - Reduce runtime from 1.2 days to 4.2 minutes
  - **Small mesh**: just 730 elements/core @ 32,768 cores
  - Scaling 64 – **32,768 cores: 85% strong scaling efficiency**
  - Small super-linear speedup
- Weak scaling (CRM, fixed workload per core, 3M – 192M elements):
  - Scaling 512 – 32,768 cores: 96% weak scaling efficiency

Runtime | strong scaling 24M elements



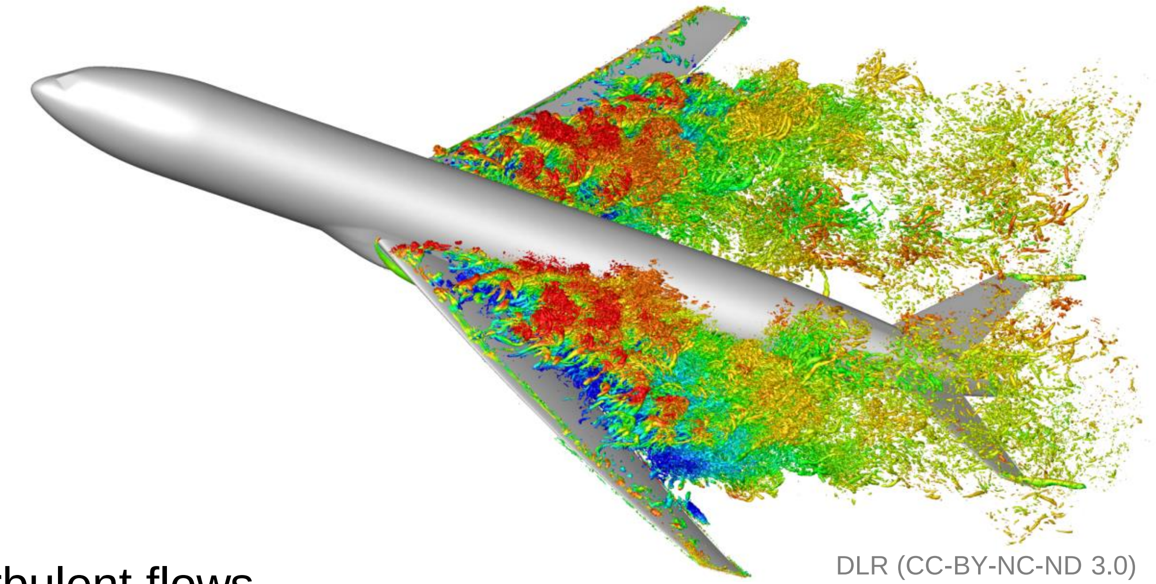
Runtime | weak scaling 3M–192M elements



Efficiency: Extrapolation

Extrapolation (same architecture but more cores)

- ~55% efficiency at 5M cores (~0.1 Exa-FLOP)
- Requires a 13B elements mesh
- Not required for industrial RANS simulations (only marginally better resolution)



DLR (CC-BY-NC-ND 3.0)

Applications for extreme scale

- Scale-resolving simulations (hybrid RANS+LES) for turbulent flows → require much finer meshes
- Unsteady RANS simulations (URANS)
- Parameter exploration: run thousands of simulations in parallel to optimize for different input parameters

Preparing for Extreme Scale (Exascale)

Three main challenges

- Efficiency
  - Scale as efficiently as possible on current systems
    → Estimate performance for large systems
- Scalability
  - Support large meshes in FlowSimulator framework
  - Support large core counts in the entire workflow
    → Overcome any hard scalability limits
- Heterogeneity
  - Readiness for a variety of different systems (CPU, GPU, …)
    → Support offloading to Nvidia and AMD GPUs

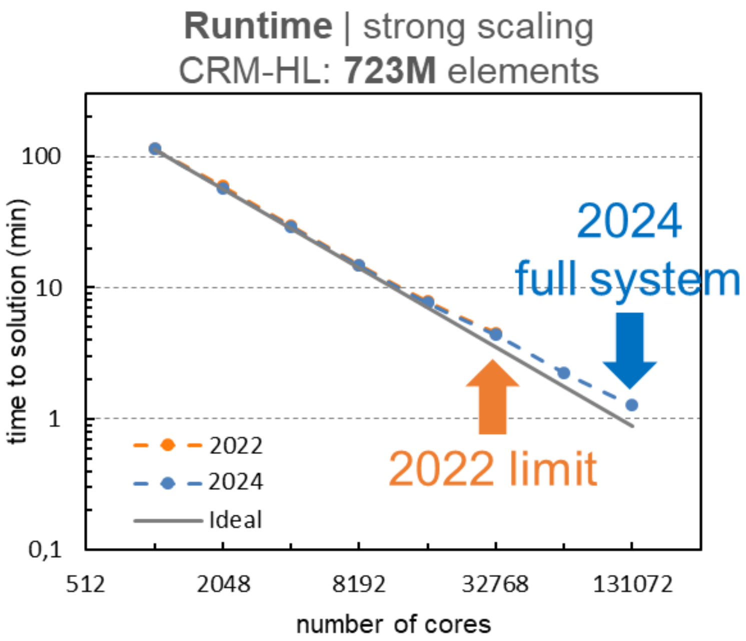
© CSC, https://lumi-supercomputer.eu

© BSC, https://www.bsc.es/marenostrum

Scalability of the Entire Workflow (FlowSimulator)

Scalability assessment on DLR’s production system CARO

- The entire workflow needs to …
  - support large meshes (>1B elements)
  - support large core counts (>1M cores)



Achievements (so far)

- Several improvements in FlowSimulator to scale to full system
- Improved hierarchical graph partitioning
- Support for meshes >1 billion elements tested
- Efficient scaling to 131,072 cores (full system CARO@DLR)

Preparing for Extreme Scale (Exascale)

Three main challenges

- Efficiency
  - Scale as efficiently as possible on current systems
    → Estimate performance for large systems
- Scalability
  - Support large meshes in FlowSimulator framework
  - Support large core counts in the entire workflow
    → Overcome any hard scalability limits
- Heterogeneity
  - Readiness for a variety of different systems (CPU, GPU, …)
    → Support offloading to Nvidia and AMD GPUs

© CSC, https://lumi-supercomputer.eu


© BSC, https://www.bsc.es/marenostrum

Heterogeneity: Support for Nvidia GPUs via Spliss

System – Juwels Booster

- AMD Epyc 7401 (2x 24 cores) per node
- 4x Nvidia Tesla A100 per node

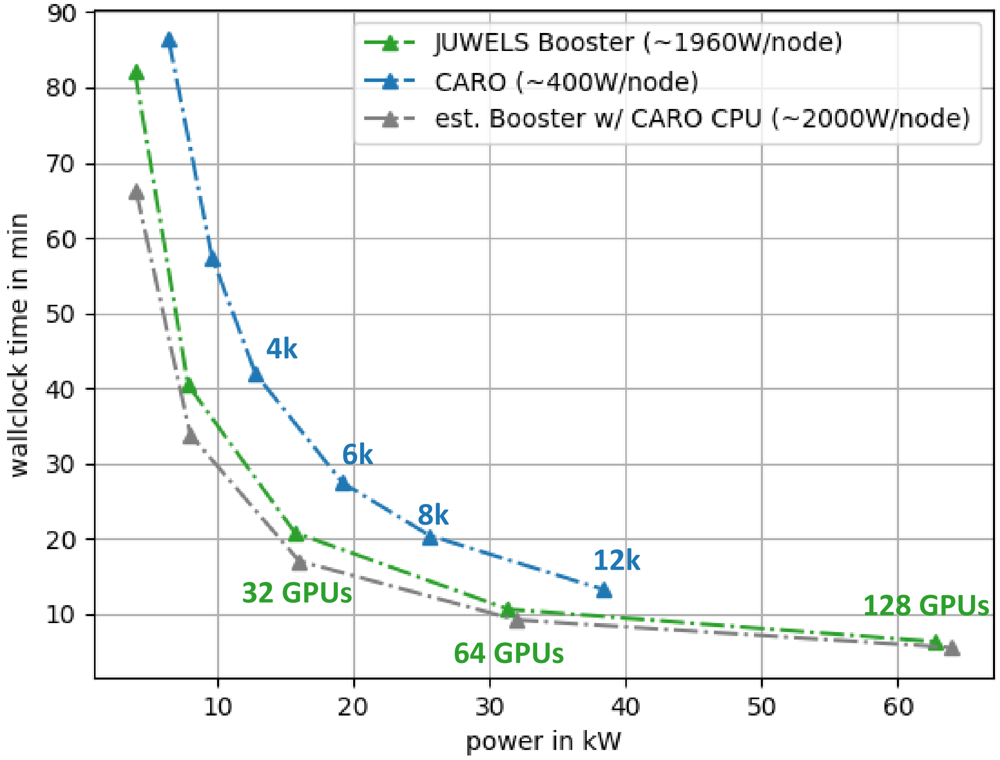
System – CARO

- Nodes with 128 cores (AMD Rome, 2x CARA)
- 8 memory channels @3.2GHz (1.2x CARA)

Observations

- Node-wise comparison (“unfair”): 8-9x speedup
- Energy-wise comparison (“fair”): 1.6-1.9 speedup
- Performance limited by non-linear part on slow CPU
- Hypothetical Juwels Booster node with CARO CPU: 1.8-2.3 speedup (energy-wise)

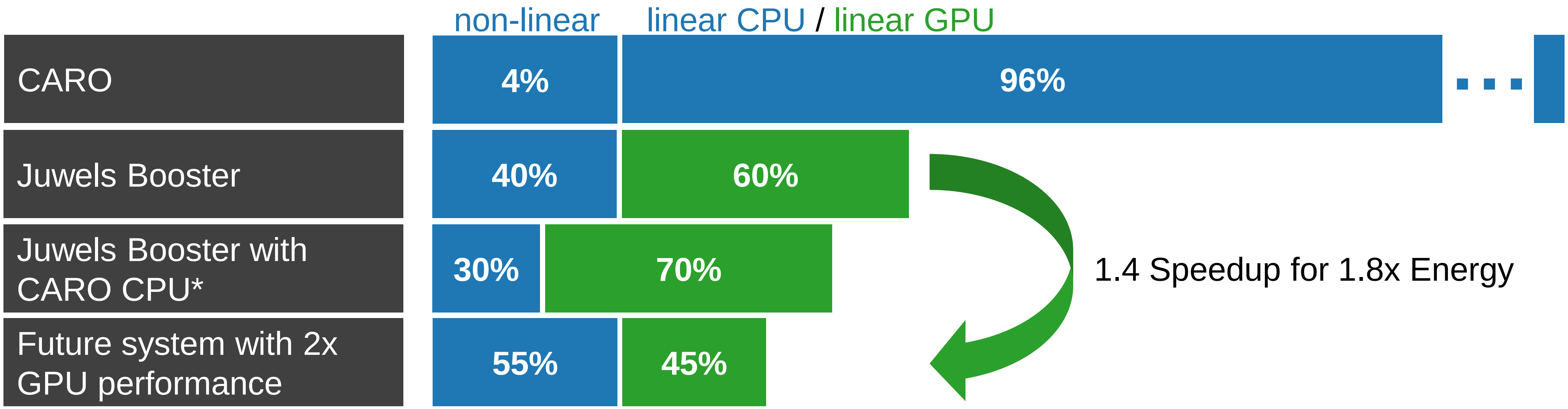
Runtime | CARO (AMD Rome) vs. Juwels (4x Nvidia A100)
M6 wing, 69.2M elements, implicit Euler, Jacobi + Block Inv.



J. Mohnke and M. Wagner: A Look at Performance and Scalability of the GPU Accelerated Sparse Linear System Solver Spliss. In: Euro-Par 2023: Parallel Processing. DOI 10.1007/978-3-031-39698-4_43

Heterogeneity: Linear vs. Non-linear Part

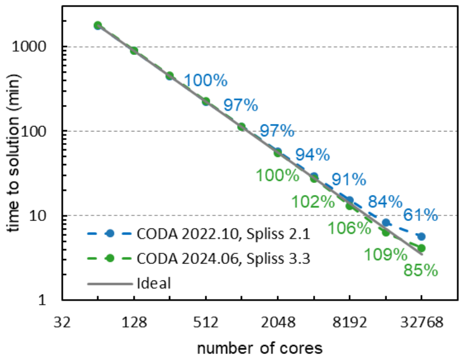
- For implicit methods, linear equation systems are solved via Spliss on GPUs
- Thus, only the linear part benefits from GPUs via Spliss
- For future system with more or more powerful GPUs the non-linear part may become bottleneck
- Port relevant parts of CODA to GPUs

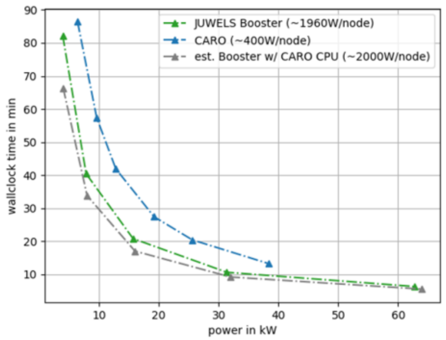


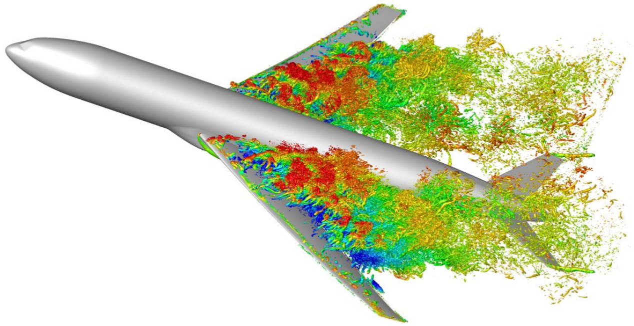
* i.e. CARA GPU partition

Summary

- Is CODA ready for extreme scale?
- Efficiency
  - Efficient scaling on current systems (96% full system)
- Scalability
  - Tested with meshes >1B elements
  - Tested on full DLR system (131.072 cores)
- Heterogeneity
  - Support for Nvidia GPUs, AMD (ongoing)







Acknowledgments

Funded by the European Union. This work has received funding from the European High Performance Computing Joint Undertaking (JU) and Germany, Italy, Slovenia, Spain, Sweden, and France under grant agreement No 101092621.

The authors gratefully acknowledge the scientific support and HPC resources provided by the German Aerospace Center (DLR). The HPC system CARA is partially funded by “Saxon State Ministry for Economic Affairs, Labour and Transport“ and „Federal Ministry for Economic Affairs and Climate Action”. The HPC system CARO is partially funded by “Ministry of Science and Culture of Lower Saxony“ and „Federal Ministry for Economic Affairs and Climate Action”.

CODA is the computational fluid dynamics (CFD) software being developed as part of a collaboration between the French Aerospace Lab ONERA, the German Aerospace Center (DLR), Airbus, and their European research partners. CODA is jointly owned by ONERA, DLR and Airbus.