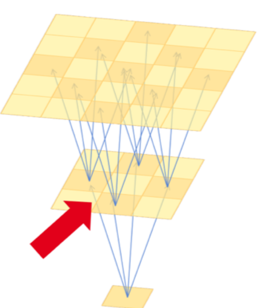
Towards a Python-Based Performance-Portable Finite-Volume Dynamical Core for Numerical Weather Prediction

S. Ubbiali[1], T. Ehrengruber[2], N. Krieger[1], C. Kühnlein[3], L. Papritz[1], H. Wernli[1]

[1] Institute for Atmospheric and Climate Science (IAC), ETH Zürich, Switzerland
[2] Swiss National Supercomputing Center (CSCS), Lugano, Switzerland
[3] European Centre for Medium-Range Weather Forecasts (ECMWF), Bonn, Germany

1. Introduction

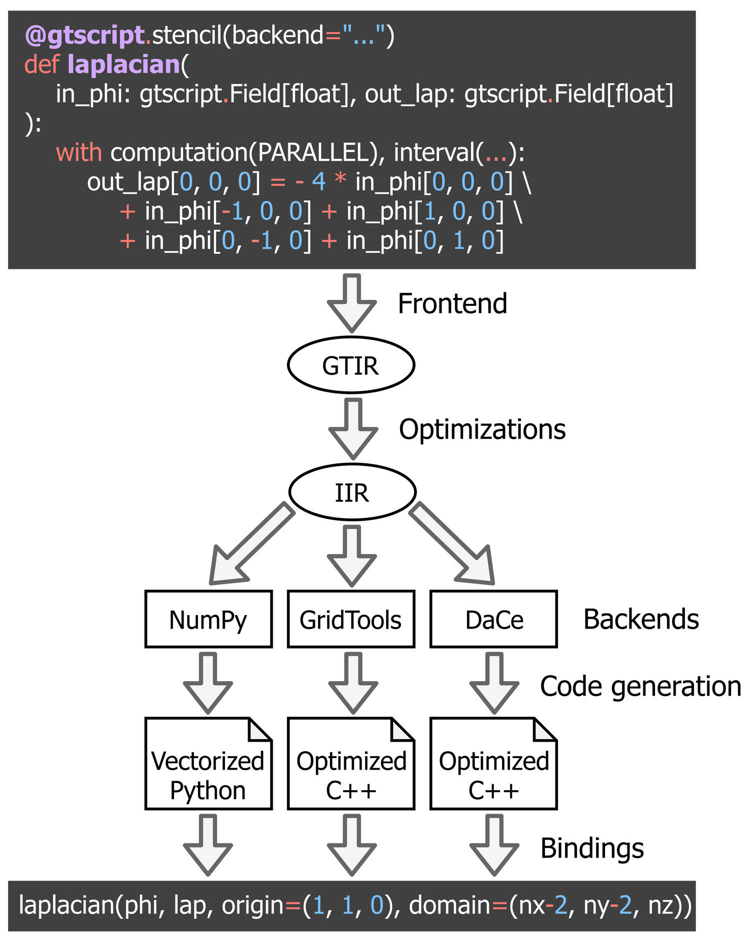
- In the context of the **PASC project KILOS** ("Kilometer-scale non-hydrostatic global weather prediction with IFS-FVM"; pasc-ch.org/projects/2021-2024/kilos), we present recent progress in the development of a **Python** implementation of the **FVM** dynamical core (Kühnlein et al. 2019, Smolarkiewicz et al. 2014).
- The primary goal of the KILOS project is establishing the FVM as a **flexible weather and climate research tool** at the Institute for Atmospheric and Climate Science (IAC) of ETH Zurich.
- FVM solves the fully compressible equations using 3D semi-implicit integration and conservative finite-volume non-oscillatory advection. The model provides very competitive time-to-solution but at the same time maps well onto modern computing architectures.
- The Python implementation of FVM is designed to attain high performance on multiple platforms by encoding stencil computations using the **domain specific library (DSL)** GridTools (**GT4Py**; github.com/GridTools/gt4py).



***Stencils** compute grid point values by accessing a **fixed pattern** of neighbors.*

2. GT4Py: A Framework for Stencil Applications

- Jointly developed by CSCS, MeteoSwiss and the Allen Institute for AI (AI2) as part of the community effort towards **single source code**, highly **maintainable** and **performance-portable** models.
- Produce **high-performance** implementation of stencil kernels starting from a **symbolic** and **hardware-agnostic** definition.
- Frontend and optimization passes translate the definition into hierarchy of tree-like **intermediate representations** (IR), later consumed by **backends** to synthetize optimized and/or purpose-built code.
- Generated code is transparently compiled and seamlessly accessible from Python.



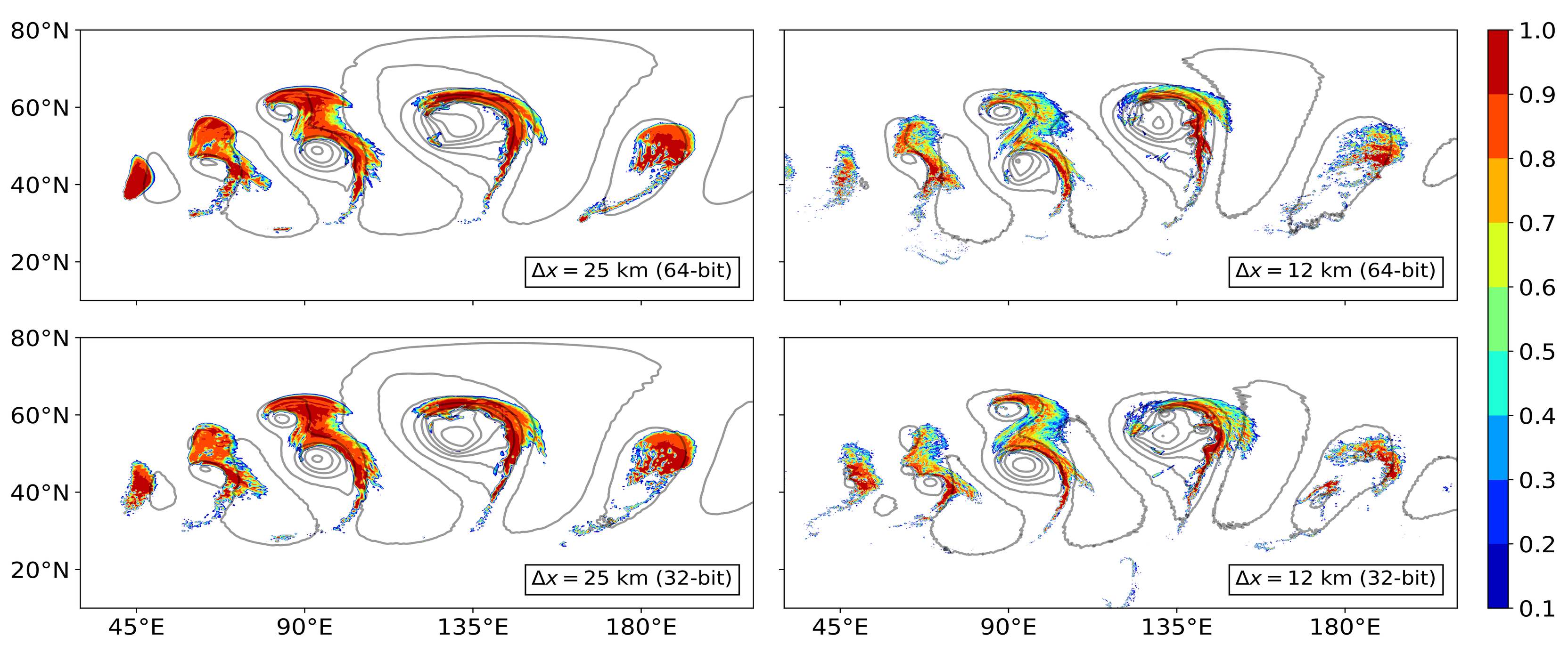
3. GT4Py Porting of FVM

Near-Global Moist Configuration

→ Moist compressible equations coupled to ECMWF cloud scheme (**CLOUDSC**).

→ Regular grid covering the latitude range +/- 80°.

→ **Global model version** with the quasi-uniform ECMWF octahedral grid under development with the new **declarative GT4Py** (cf. poster **"GT4Py: A Python Framework for the Development of High-Performance Weather and Climate Applications"**).



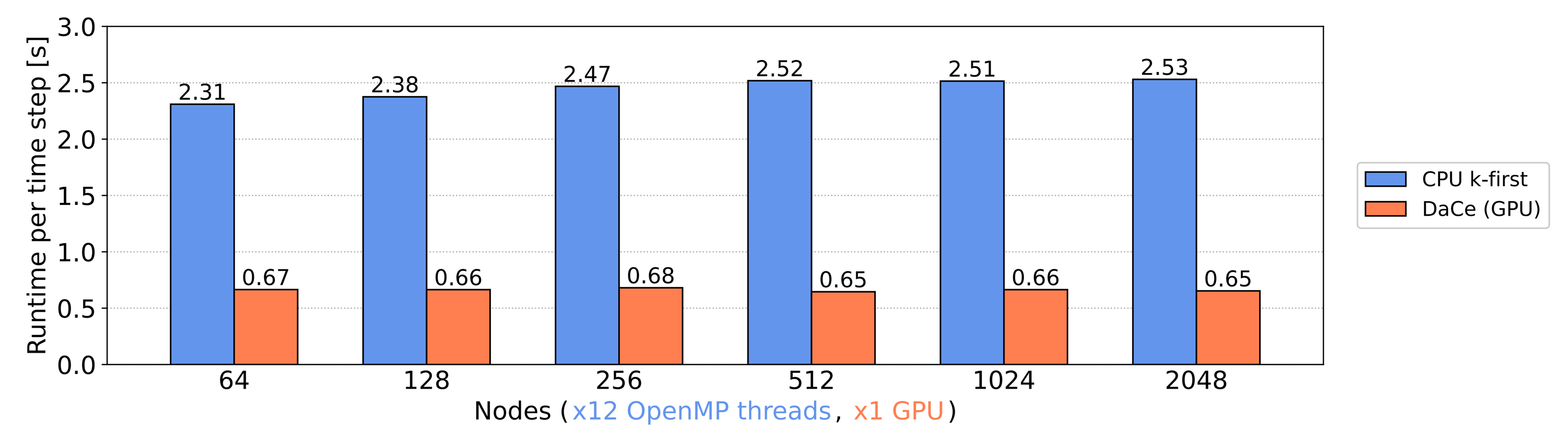
*Solution to the **moist baroclinic wave** test case (Ullrich et al. 2014) at day ten using the FVM in a near-global configuration with a horizontal equatorial grid spacing of 25 km (left) and 12 km (right), using 64-bit (top) and 32-bit (bottom) floating point arithmetic. Shown are cloud fraction at about 2 km above the surface (shading) and surface pressure (contour levels with 10 hPa interval).*

Large-Eddy Simulation (LES) Configuration

→ The model has been extended with LES capabilities for boundary layer research over complex orography (cf. poster **"Investigating the Mechanism of a Local Windstorm in the Swiss Alps Using Large-Eddy Simulations"**).

Distributed-Memory Implementation

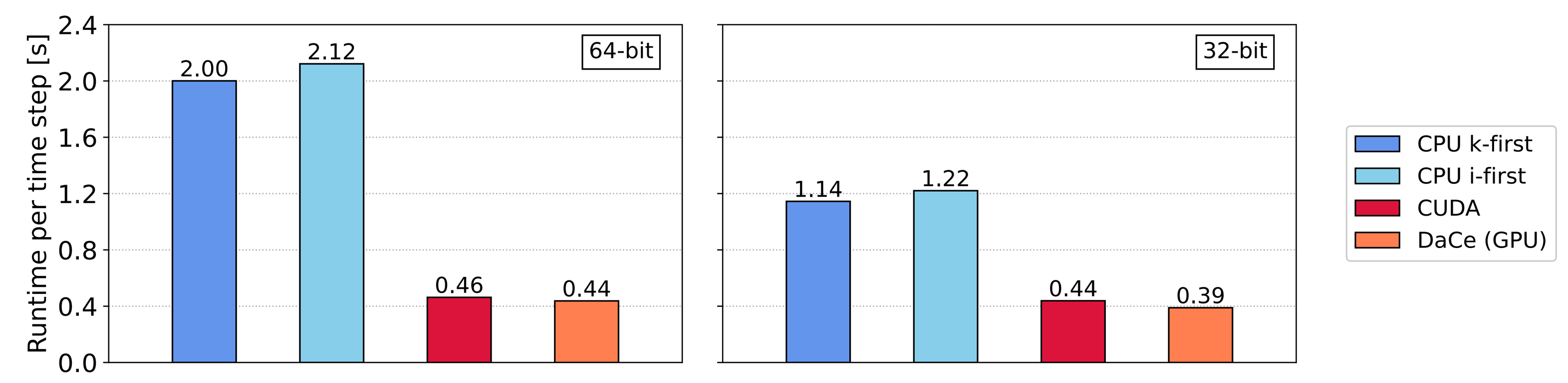
→ The model runs on **multiple CPUs and GPUs** via a prototype distributed-memory version where halo exchanges are performed in a semi-automatic fashion using the Python bindings to the generic exascale-ready library **GHEX** (github.com/ghex-org/GHEX).



***Weak scaling** of FVM from 64 to 2048 nodes of the hybrid partition of the **Piz Daint** supercomputer at CSCS. Runtimes on either CPUs (**GridTools** k-first memory layout backend of GT4Py (Afanasyev et al. 2021)) or GPUs (**DaCe** backend of GT4Py (Ben-Nun et al. 2022)).*

Reduced Precision

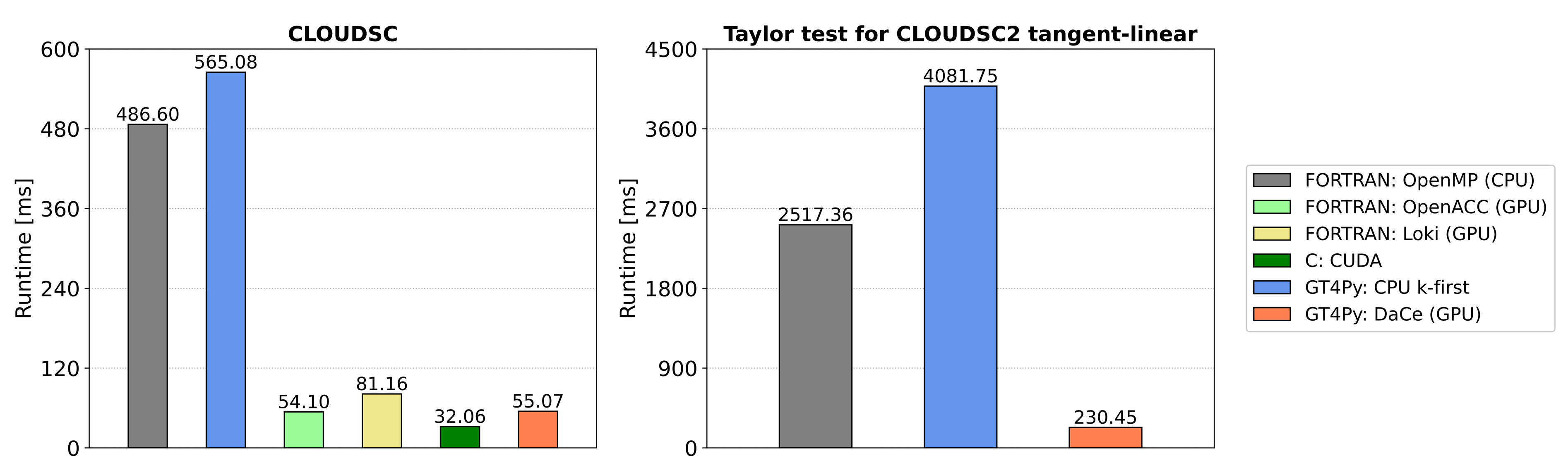
→ The model executes entirely in either double and single floating point precision.



Performance delivered by four backends of GT4Py on a single hybrid node of Piz Daint. The backends run in 64-bit (right) and 32-bit (left) floating point precision either on CPU (k-first and i-first memory layout) or GPU (CUDA and DaCe).

4. GT4Py Porting of ECMWF Physics

The microphysics packages CLOUDSC (Müller et al. 2019; github.com/ecmwf-ifs/dwarf-p-cloudsc) and **CLOUDSC2** (in its **nonlinear**, **tangent-linear** and **adjoint** formulations; github.com/ecmwf-ifs/dwarf-p-cloudsc2-tl-ad) have been **fully rewritten** in Python using GT4Py.



Execution time for the CLOUDSC dwarf (left) and the Taylor test for the tangent-linear formulation of CLOUDSC2 (right) measured on a single hybrid node of Piz Daint. Displayed are three FORTRAN versions (either blended with OpenMP and OpenACC compiler directives, or using the source-to-source translation tool Loki (github.com/ecmwf-ifs/loki)); an optimized CUDA C implementation; and the GT4Py rewriting (CPU k-first and DaCe backends).

5. References

Afanasyev, A., et al. (2021). GridTools: A framework for portable weather and climate applications. *SoftwareX*, *15*, 100707.

Ben-Nun, T., et al. (2022). Productive performance engineering for weather and climate modeling with Python. In *SC22: International Conference for High Performance Computing, Networking, Storage and Analysis* (pp. 1-14). IEEE.

Kühnlein, C., et al. (2019). FVM 1.0: a nonhydrostatic finite-volume dynamical core for the IFS. *Geoscientific Model Development*, *12*(2), 651-676.

Kühnlein, C., et al. (2023). ECMWF collaborates with Swiss partners on GPU porting of FVM dynamical core. *ECMWF Newsletter 175.*

Müller, A., et al. (2019). The ESCAPE project: energy-efficient scalable algorithms for weather prediction at exascale. *Geoscientific Model Development*, *12*(10), 4425-4441.

Smolarkiewicz, P. K., et al. (2014). A consistent framework for discrete integrations of soundproof and compressible PDEs of atmospheric dynamics. *Journal of Computational Physics*, *263*, 185-205.

Ullrich, P. A., et al. (2014). A proposed baroclinic wave test case for deep-and shallow-atmosphere dynamical cores. *Quarterly Journal of the Royal Meteorological Society*, *140*(682), 1590-1602.