



Towards Cloud-Resolving European-Scale Climate Simulations using a fully GPU-enabled Prototype of the COSMO Regional Model

David Leutwyler (1), Oliver Fuhrer (3), Benjamin Cumming (5), Xavier Lapillonne (4), Tobias Gysi (2,6), Daniel Lüthi (1), Carlos Osuna (2), and Christoph Schär (1)

(1) Institute for Atmospheric and Climate Science, ETH Zurich, Switzerland(david.leutwyler@env.ethz.ch), (3) Federal Office of Meteorology and Climatology, MeteoSwiss, Zurich, Switzerland, (5) Swiss National Supercomputing Centre (CSCS), Lugano, Switzerland, (4) Supercomputing Systems AG, Zurich, Switzerland, (2) Center for Climate Systems Modeling (C2SM), ETH Zurich, Switzerland, (6) Swiss Federal Institute of Technology (ETH Zurich), Scalable Parallel Computing Lab, Switzerland

The representation of moist convection is a major shortcoming of current global and regional climate models. State-of-the-art global models usually operate at grid spacings of 10-300 km, and therefore cannot fully resolve the relevant upscale and downscale energy cascades. Therefore parametrization of the relevant sub-grid scale processes is required. Several studies have shown that this approach entails major uncertainties for precipitation processes, which raises concerns about the model's ability to represent precipitation statistics and associated feedback processes, as well as their sensitivities to large-scale conditions. Further refining the model resolution to the kilometer scale allows representing these processes much closer to first principles and thus should yield an improved representation of the water cycle including the drivers of extreme events.

Although cloud-resolving simulations are very useful tools for climate simulations and numerical weather prediction, their high horizontal resolution and consequently the small time steps needed, challenge current supercomputers to model large domains and long time scales. The recent innovations in the domain of hybrid supercomputers have led to mixed node designs with a conventional CPU and an accelerator such as a graphics processing unit (GPU). GPUs relax the necessity for cache coherency and complex memory hierarchies, but have a larger system memory-bandwidth. This is highly beneficial for low compute intensity codes such as atmospheric stencil-based models. However, to efficiently exploit these hybrid architectures, climate models need to be ported and/or redesigned.

Within the framework of the Swiss High Performance High Productivity Computing initiative (HP2C) a project to port the COSMO model to hybrid architectures has recently come to an end. The product of these efforts is a version of COSMO with an improved performance on traditional x86-based clusters as well as hybrid architectures with GPUs.

We present our redesign and porting approach as well as our experience and lessons learned. Furthermore, we discuss relevant performance benchmarks obtained on the new hybrid Cray XC30 system "Piz Daint" installed at the Swiss National Supercomputing Centre (CSCS), both in terms of time-to-solution as well as energy consumption. We will demonstrate a first set of short cloud-resolving climate simulations at the European-scale using the GPU-enabled COSMO prototype and elaborate our future plans on how to exploit this new model capability.