

# PP IMPACT / HPC

## Latest Developments at DWD

## Introduction

- A new colleague, Marek Jacob, is working at DWD on the GPU port of ICON (since 1st of April).
- At DWD we are working on:
  - Nesting feature: Marek
  - Tiedtke-Bechtold convection: Uli
- GPU machine available at DWD: Up to now only our Linux workstations, with PGI compiler 21.2. This is a good platform for development, but not for testing larger domains, especially for the nesting.
  - Device: NVIDIA Quadro P400
  - Memory: 2 GB

## Porting Tiedtke-Bechtold Convection

- ➔ Rather similar code has been ported by Italian colleagues (Riccardo Scatamacchia et al.) for COSMO, so most of the OpenACC commands could be taken from COSMO with little adaptations.
- ➔ But ported code did not work „out-of-the-box“:
  - ➔ Module mo\_cuparameters.f90 already contained an acc statement:  
„acc declare copyin(rd)“ after REAL declaration of rd (gas constant),  
but no „update device“ after computation of rd: So ICON worked with rd=0
  - ➔ On the workstation, the CPU does not work with FMA, but GPU does. This lead to slightly different results.
- ➔ Thanks to the work done by Riccardo and his team, porting the ICON version of Tiedtke-Bechtold to GPUs was rather easy. BUT....

## Testing Tiedtke-Bechtold on GPUs

→ For the small domain I use on my workstation (1788 cells). Timings from a 1-hour forecast:

	CPU nproma=16	GPU nproma=1788
Total	12.390	22.489
nh solve	6.237	3.191
NWP Turbulence	0.566	1.450
NWP Microphysics	0.488	0.367
NWP Convection	0.699	3.372

## Testing Tiedtke-Bechtold on GPUs

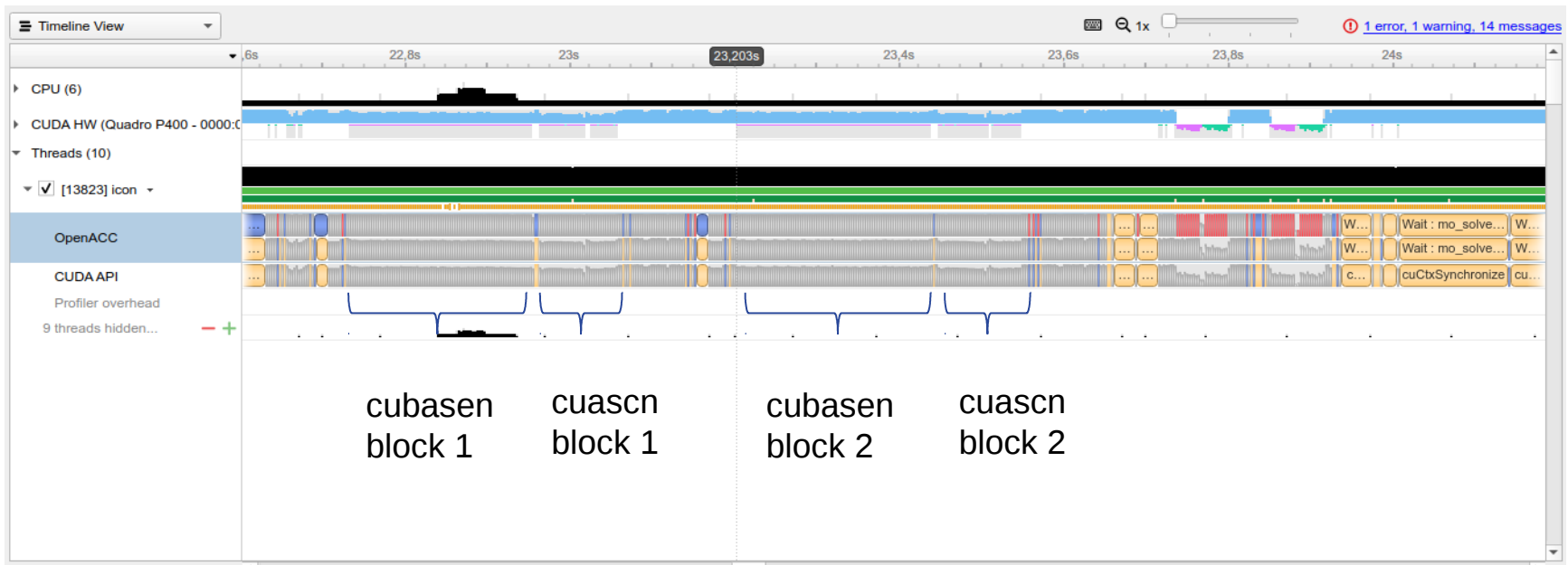
→ ICON-D2 on Daint using 32 nodes. Timings from a 1-hour forecast:

	CPU nproma=16	GPU nproma=8192
Total	1117.193	
nh solve	564.026	
NWP Turbulence	74.506	
NWP Microphysics	52.561	
NWP Convection	43.795	

The run on Daint GPUs crashed after a few timesteps (exceeding CFL)!

# Profiling

- ➔ Started to learn about profiling tools (nsys, nvprof), but not yet matured enough
- ➔ An nsys-example from my workstation:

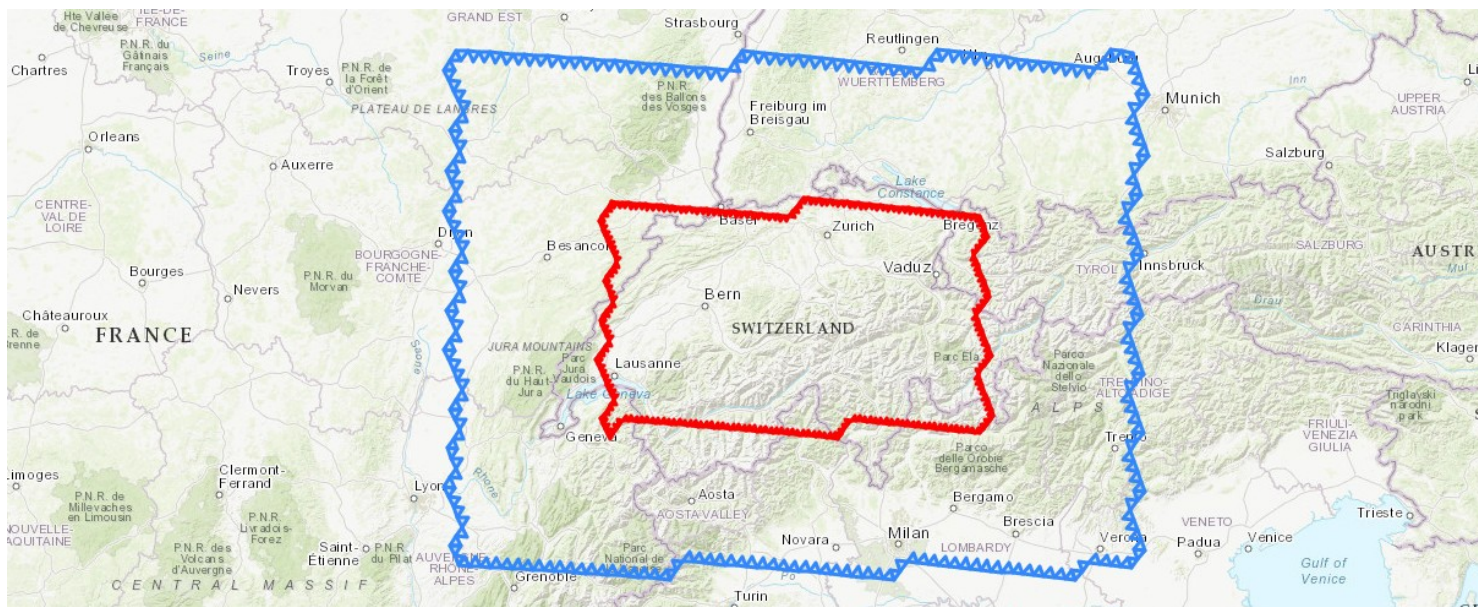


But this was done with optimization level `-O0`!

## Short Summary

- Convection has several nested vertical loops (ascent / descent), which result in a lot of kernels.
- On my workstation, convection on the GPU runs slower than on the CPU.
- On Daint, no big tests could be performed up to now (with the latest icon-nwp/master)
  
- Learning and understanding is „ongoing work“.

- ➔ Lat-lon sampled output in post-processing (output\_nml:remap = .TRUE.)
- ➔ Serialization (for verification of ACC ports): Any (global) scalar or array can be considered in `serialize_all now`
- ➔ Main work: **two-way nesting on GPU**



Icon Grid Generator;  
Map Tiles © Esri – Esri, DeLorme, NAVTEQ, TomTom, Intermap, IPC, USGS, FAO, NPS, NRCAN, GeoBase, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), and the GIS User Community



## → Working features

- two nested local area domains (LAM)
- multiple compute processes (MPI)
- initialization with initial and boundary data
- initialization of both domains at  $t=0$  or delayed (e.g. `grid_nml:start_time = 0., 10`)
  - Delayed, i.e. online, initialization is done on CPU (as is the regular initialization. uses device-to-host copy before and host-to-device copy afterwards.)
- output writing
- Options:
  - Nest boundary nudging (`gridref_nml:grf_intmethod_e=4 (6)`)
  - Density boundary nudging (`gridref_nml:l_density_nudging=.TRUE. (.FALSE.); grid_nml:lfeedback=.FALSE. (.TRUE.)`)

→ Changes include:

- continuation of Günther's work
- Improve serialization tests
  - Make serialization code aware that there can be more than one domain
- Find and fix openACC pitfalls in relation to nested derived types, e.g:
  - !\$ACC ENTER DATA CREATE(turbdiff\_config(jg)) !(not wanted)
  - !\$ACC ENTER DATA CREATE(turbdiff\_config(jg:jg))
- Fix communication for gridrefinement (data locality)
- Added ``IF acc CALL finish()`` to non-ACC functions that branch out from GPU regions.
- Avoid uninitialized memory in global arrays.
- Fixed a synchronization bug with Terra (missing WAIT after ASYNC)
- New buildbot test case for Daint (mch\_opr\_r04b07\_nest)