Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Home Affairs FDHA
Federal Office of Meteorology and Climatology MeteoSwiss

PP IMPACT project and HPC developments

xavier.lapillonne@meteoswiss.ch

Operational Models at COMET

- COSMO ME CPU (version 5,06) 24 nodes, 576 cores in total. ***Tiedke*** convection scheme. Simulation time about 1.5 h (+72h)

- COSMO IT CPU (version 5,06) 24 nodes, 576 cores in total. ***Bechtold*** shallow convection scheme 24 nodes, 576 cores in total. Simulation time about 2.3 h (+48h)

- COSMO ME EPS (pompa) hybrid CPU-GPU 40 members one node per each member. ***Tiedke*** convection scheme. Simulation time about 1,3 h (+72h)

- COSMO IT EPS (pompa) hybrid CPU-GPU 20 members two nodes per each member. ***Tiedke*** shallow convection scheme. Simulation time about 1.9 h.

Bechtold Convection Scheme in COSMO

Port to GPU using OpenACC

Main module: *it computes the physical tendecies of the prognostic variables T, Q, U, V and tracers due to convective processes*

conv_cuinit.f90
Inizialization of physical fields

conv_cuascn.f90
Calculations for cloud ascents for cumulus

conv_cudescn.f90
Calculates cumulus downdraft descent

conv_cufluxtends.f90
Calculates convective fluxes in cloud and in subcloud layers. Update of T, Q, U, V and computation of transport of chemical tracers

12 subroutines ported on GPU – about 10000 code lines

Performance

	CPU (Tiedke)	CPU (Bechtold)	GPU (Tiedke)	GPU (Bechtold)
Total Time	9679 sec	10968 sec	1079 sec	1179 sec
Convection Time	23 sec	121 sec	1,88 sec	58 sec

Convection Speed up with respect to CPU code:

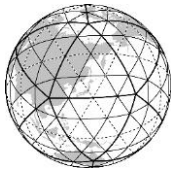
Bechtold x 2
Tiedke x 12

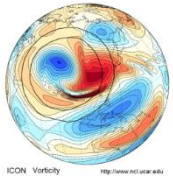
Next step : optimization

Results for +6h COSMO v5.06 with lgsp_first=true using 8 GPU sockets (4 K80 NVIDIA cards) or 8 CPU sockets (8 Intel Hasswell CPUs with 12 cores each) – Measured on HAL System at COMET

ENIAC : baseline OpenACC GPU implementation of ICON for climate

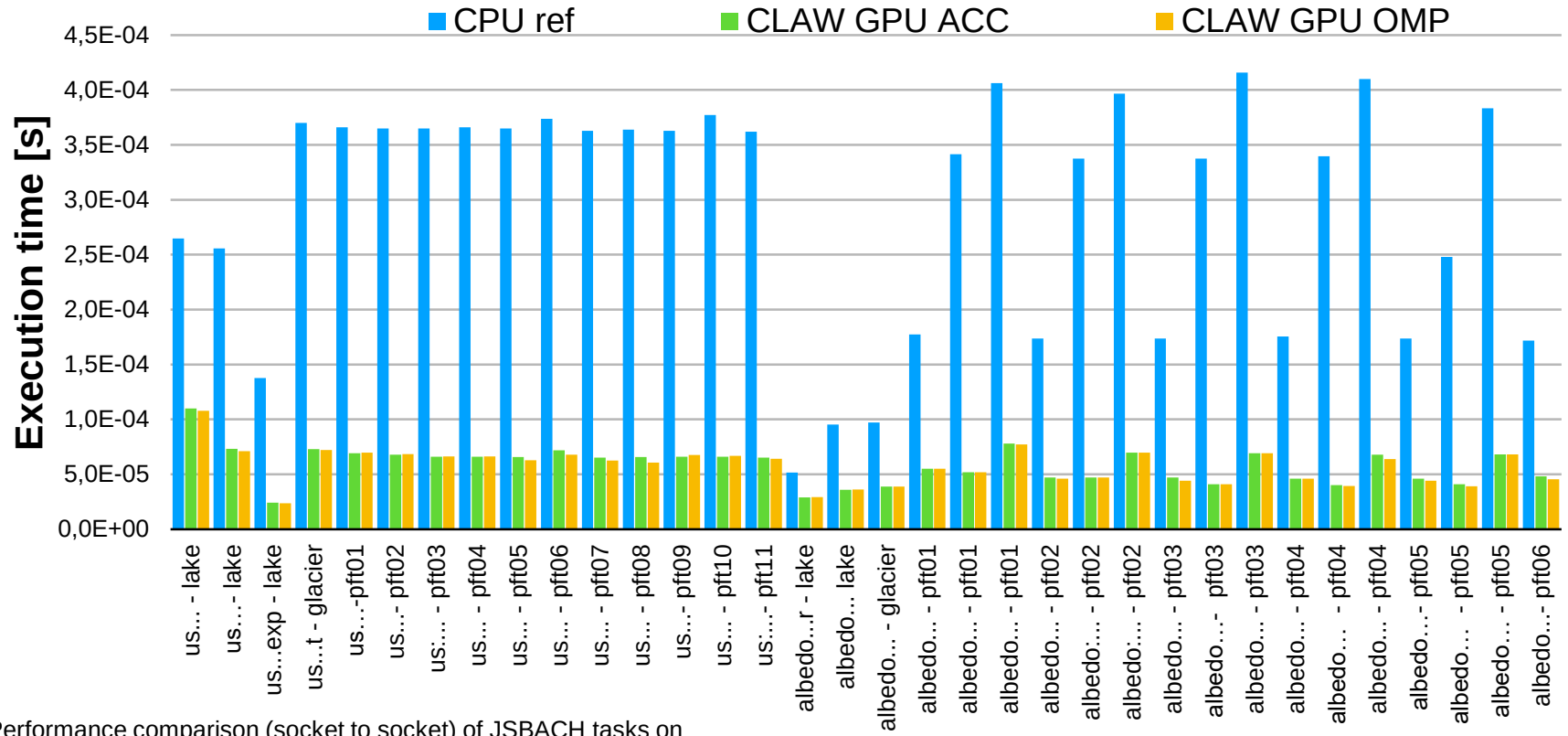
- Participants : ETHZ, C2SM, MeteoSwiss, MPI-M, DWD, DKRZ
- First base version OpenACC:
  - Dynamics+Climate physics+interface code : OpenACC
  - JSBACH (soil model) : claw-dsl (Fortran based DSL for physics)
- ICON is VERY large : ~800 Fortran module files, ~1mio LOC, rapidly evolving code : more difficult to port to GPU than COSMO
- First (reduced) GPU version for Global Climate applications should be ready by Q4 2019





Application: "Quasi-Biennial Oscillation in a Changing Climate (QUBICC)“

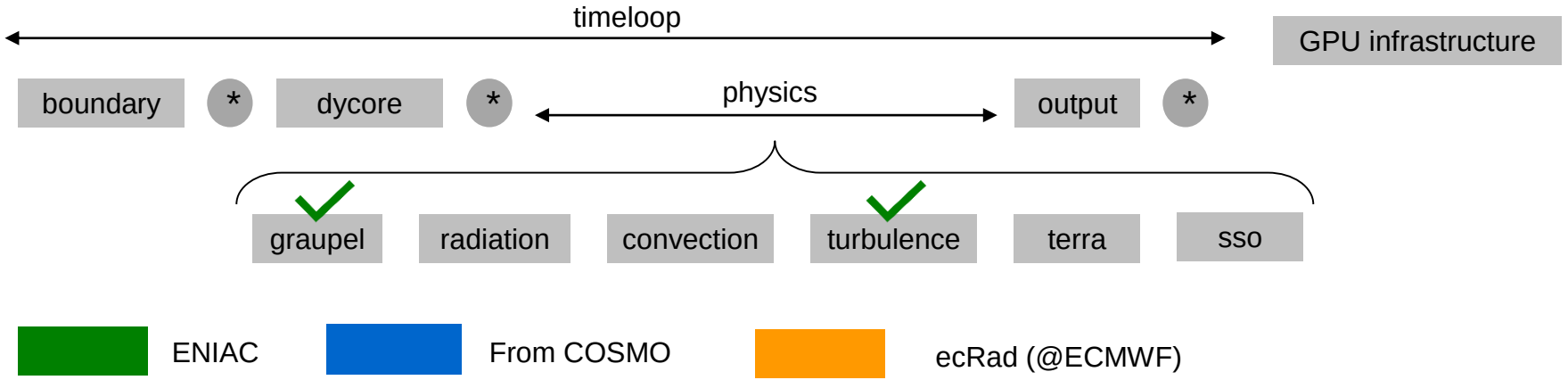
Performance results JSBACH with CLAW



Performance comparison (socket to socket) of JSBACH tasks on Intel Haswell E5-2690v3 and NVIDIA P100. Domain size (horizontal grid points x vertical levels) = 20480 x 47

COSMO PP IMPACT

- GPU baseline version of ICON using **OpenACC** for NWP and LAM application :
  - Initial version of the dycore is available
  - Physical parametrizations require a full port
- Investigate appicability of **OpenMP for accelerators**
- Investigate performance portability using Domain Specific Langage



* 2-way nesting infrastructure, physics interfaces, output diagnostics, interpolation, …

Testing is Key : need to improve validation and performance testing

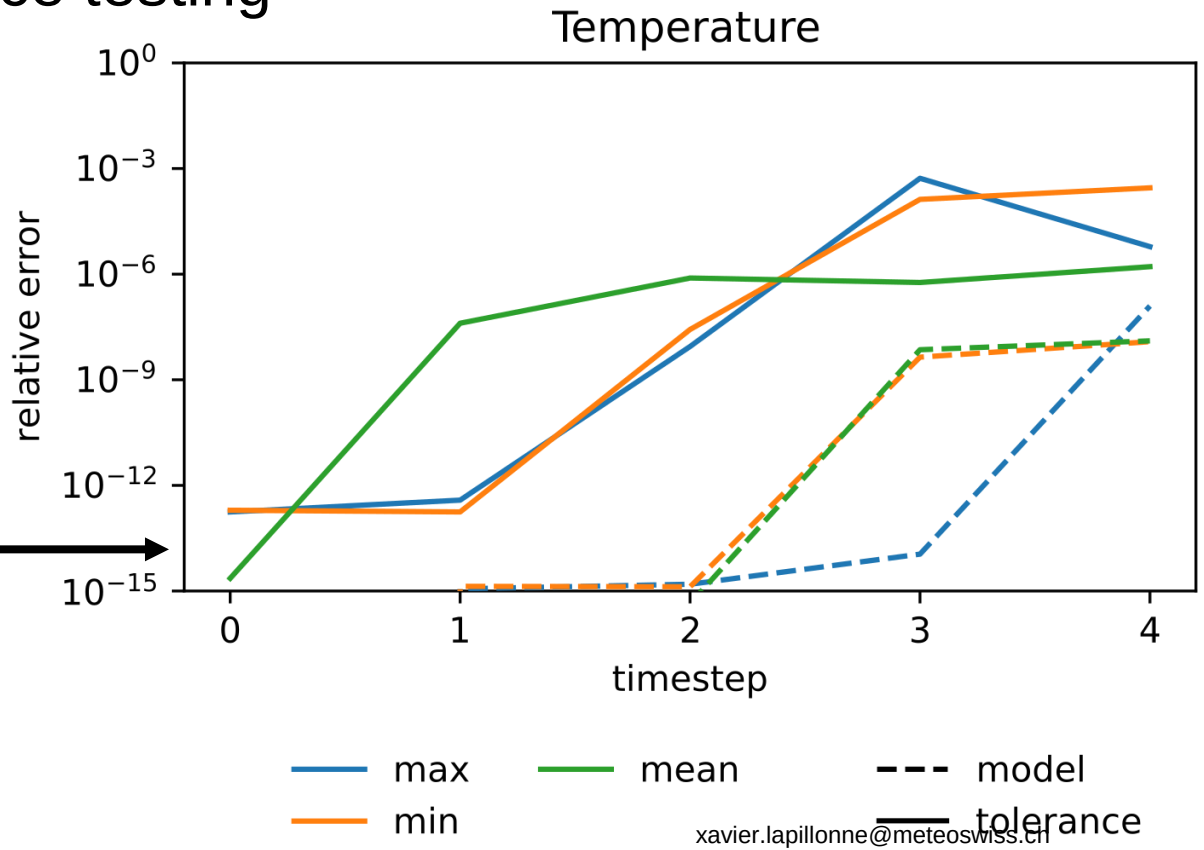
These are results for a global test (atm_amip)

We need a reference output and tolerances for each test

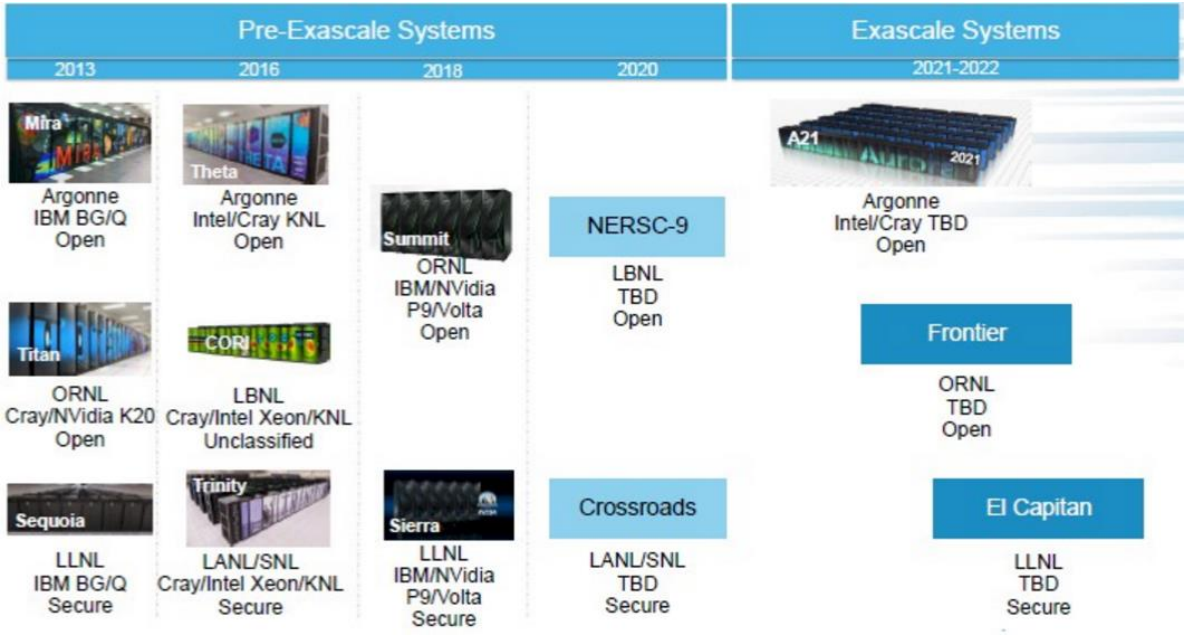
References and tolerances need to be updated for each change in model physics

There is a script to do that, needs to be incorporated in ICON workflow

Perturb initial conditions by 10^-14 (relative)



Why using abstraction



Intel Xe GPUs
OneAPI (incorporates C++ SYCL)

AMD GPUs

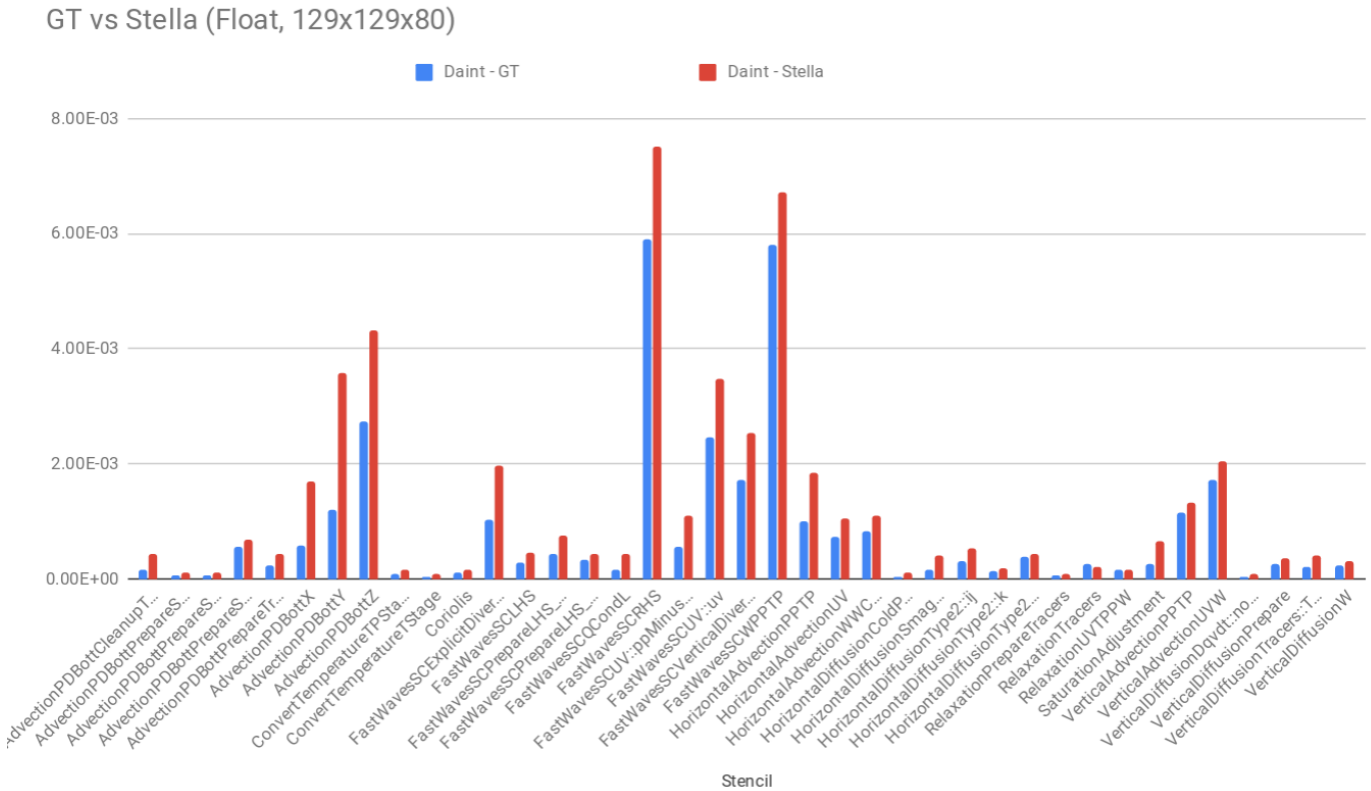
OPENMP, Openacc ?

??

GridTools dycore

- Almost ready for integration in official code

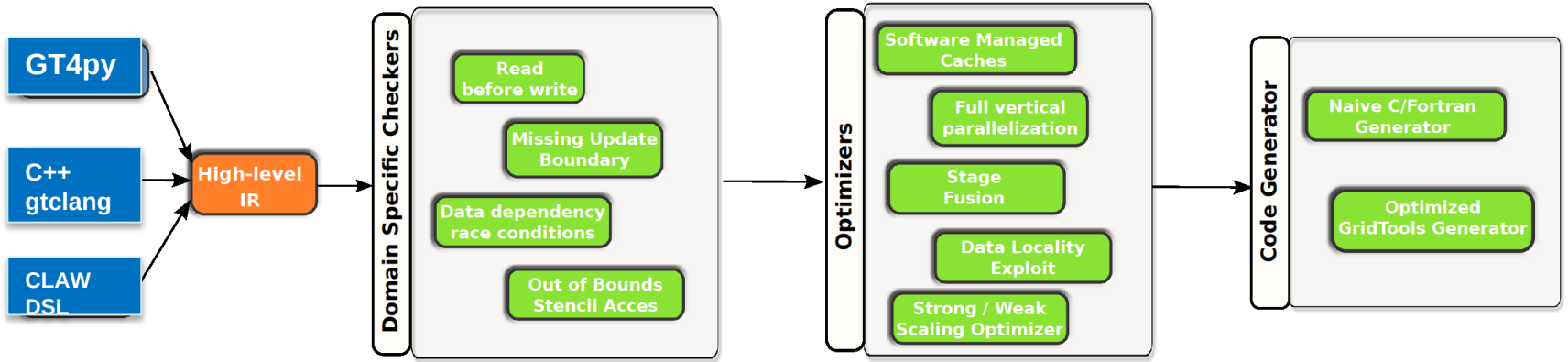
- Will be used on the next production system at MeteoSwiss (Arolla, V100 GPUs)



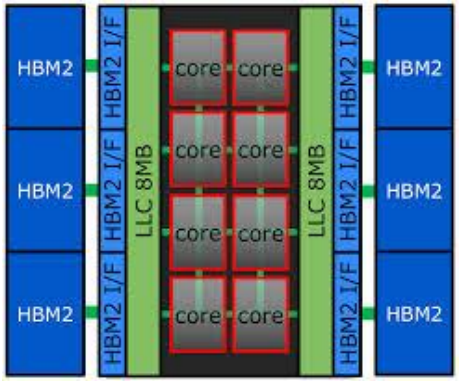
Next Generation High-level DSL

- Easier to use (matlab like)
- Naive code generation for scientific debugging
- No template metaprogramming (faster compilation)
- HIR (High Level Intermediate Representation) decouple frontend and toolchain
- Large collaboration (~10 dev), OpenSource project
  MeteoSwiss (ICON), CSCS (FVM), Vulcan (FV3)

DSL Frontends



New DWD system : NEC vector system



- NEC system nodes:
  - a vector host: 24-core AMD Rome
    (2.8 GHz; 2.048 GByte memory)
  - 8 vector engines: SX-Aurora 1 TSUBASA Typ 10AE
  - Liquid Cooling

- 3 phases, until 2023 with incresing capacity
- Peak Performance: 0: 433 TF/s; 1: 545 TF/s; 2: 545 + X TF/s

- Support SINFONY project:
  Seamless Integrated Forecasting System for Nowcasting and very short range forecasts : 3-4.5 growth in computation

Phase	Operations	Experiments
0	178 / 1424	232 / 1856
1	224 / 1792	292 / 2336
2	224+101/3408	292+132 / 4448

Challenges ahead !

- Many architectures (x86, NEC vector, GPUs, ?)
- ICON, DSL, …
- Can we succeed as a consortium ?
  - we don’t agree on everything
  - we need to find ways to collaborate in the most optimal way
  - Compromise, changes in the way we work, coordination will be critical