Reproducible climate and weather simulations: an application to the COSMO model

C. Charpilloz*, A. Arteaga**, O. Fuhrer*, T. Monkte1, and C. Harrop1

* Federal Office of Meteorology and Climatology MeteoSwiss, 1 Swiss Federal Institute of Technology Zurich,
1 Cooperative Institute for Environmental Sciences

Introduction

Environment with changing climate requires the analysis of convection-resolving simulations at continental scale. However, the huge amount of produced data makes the analysis of the simulation impractical as it cannot be stored.

We propose a new framework where data are transparently traded for computational effort.

Material: the COSMO model

COSMO is a non-hydrostatic atmospheric model based on finite difference solvers and stencils computations. Problems: (i) solution is highly non-linear, (ii) different architectures → different compilers → different computations (iii) climate simulations are highly non-linear → small initial differences in re-runs will rapidly grow into larger differences.

Solution: adapt COSMO to produce bit reproducible results on these architectures [1].

Source of non-reproducibility

Example of reassociation applied by the compiler:

(a) azg = a + b
(b) azg = a + b
(c) azg = a + b

Method: preprocessing the code and optimization flags

The dynamics (C++/CUDA)

C++ standard is restrictive regarding the code optimizations. Two steps are needed:

(i) deactivating the FMA:

with GNU: -fp-contract=off and with NVCC: -nofma

(ii) providing portable transcendental functions:

The physics (Fortran/OpenAcc)

Fortran compilers have much more freedom in reorganizing the expressions or instructions in order to optimize the execution speed. Hence the code must be preprocessed.

(i) The order of evaluation must be unique to ensure the generation of a unique AST. We add parenthesis: ztu8 = ((pa2c * pcb1) + (ztd6 * ztu2) + (ztd7 * ztu4))

(ii) the transcendental functions' calls must be shadowed by a portable implementation:

(iii) as one cannot shadow intrinsic operators in Fortran we need to replace them. In COSMO it's the exponentiation function exp, which can be replaced by two calls to log and texp(x) = exp(x) − 1.

Results

Full bit-reproducibility has been achieved for COSMO in some setups. The following matrix shows which setups are reproducible:

<table>
<thead>
<tr>
<th>Setup</th>
<th>GNU</th>
<th>PGI</th>
<th>Cray</th>
<th>NVCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>marchmatl-arch, hip-contrb, if, od</td>
<td>Eisee, concorval, codc,nmaa, od, nmaa, gd, solir</td>
<td>equp, env, msgam, fluv,mprgd, enst</td>
<td>flud-flags, fluv-flags, prec-divir, prec-sqrt</td>
<td>imp-flags, fluv-flags, prec-divir, prec-sqrt</td>
</tr>
<tr>
<td>marchmatl-arch, hip-contrb, if, od</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

References

