ESiWACE

A Center of Excellence for HPC applications to support cloud resolving earth system modelling

Joachim Biercamp, Panagiotis Adamidis, Philipp Neumann
Deutsches Klimarechenzentrum (DKRZ)
Motivation: Cloud-Resolving Schemes

- High level of detail → 100-1000m horizontal resolution
- Less parametrisation → In the limit, we know the equations!
- Challenges: very compute/memory/data intensive!
Overview

1. Motivation: Cloud Resolving Schemes

2. ESiWACE
   1. Overview and Goals
   2. High-resolution Demonstrators
   3. Scalability Results

3. Work in Progress: I/O and Online Diagnostics

4. Summary
ESiWACE: Overview

- **ESiWACE** = Centre of Excellence in Simulation of Weather and Climate in Europe

- Funded by H2020, e-Infrastructures „Centres of Excellence for computing applications“

- ESiWACE leverages two European networks:
  - European Network for Earth System Modelling (ENES)
  - European Centre for Medium-Range Weather Forecasts (ECMWF)
ESiWACE substantially improves efficiency and productivity of numerical weather and climate simulation on high-performance computing platforms by supporting the end-to-end workflow of global Earth system modelling.
ESiWACE substantially improves efficiency and productivity of numerical weather and climate simulation on high-performance computing platforms by supporting the end-to-end workflow of global Earth system modelling.
ESiWACE substantially improves efficiency and productivity of numerical weather and climate simulation on high-performance computing platforms by supporting the end-to-end workflow of global Earth system modelling.
ESiWACE substantially improves efficiency and productivity of numerical weather and climate simulation on high-performance computing platforms by supporting the end-to-end workflow of global Earth system modelling.

WP1 Governance and engagement

WP2 Scalability

Global high resolution model demonstrators

WP3 Usability

WP4 Exploitability

WP5 Management and dissemination
ESiWACE: The Demonstrator Approach

- Establishment of global high-resolution simulations to demonstrate the computability of weather/climate predictions with (at least)
  - 1km resolution (atmosphere only)
  - 10km resolution (atmosphere-ocean)
  -> enable the simulation of convective clouds, small-scale ocean eddies
  -> higher fidelity of high-impact regional events
- Implementation and operation of required infrastructures
- Long-term: extreme-scale robust high-resolution simulations in 50 member ensemble at 100-1000 forecast days per day
- Codes: IFS, NEMO, ICON, EC-EARTH, MPI-ESM2

Courtesy by Matthias Brück and Daniel Klocke
High-resolution Predictions with ICON

- ICON: Icosahedral Non-hydrostatic
- Global, unstructured grid: created via successive refinement of icosahedron

10km

Courtesy by Michael Böttinger
High-resolution Predictions with ICON

- ICON: Icosahedral Non-hydrostatic
- Global, unstructured grid: created via successive refinement of icosahedron

Courtesy by Michael Böttinger

5km
High-resolution Predictions with ICON

- ICON: Icosahedral Non-hydrostatic
- Global, unstructured grid: created via successive refinement of icosahedron

2.5km

Courtesy by Michael Böttinger
High-resolution Predictions with ICON

- ICON: Icosahedral Non-hydrostatic
- Global, unstructured grid: created via successive refinement of icosahedron

1.2km

Courtesy by Michael Böttinger
High-resolution Predictions with ICON

- ICON: Icosahedral Non-hydrostatic
- Global, unstructured grid: created via successive refinement of icosahedron

1.2km

Courtesy by Michael Böttinger
High-resolution Predictions with ICON

- ICON: Icosahedral Non-hydrostatic
- Global, unstructured grid: created via successive refinement of icosahedron

1.2km
Challenge: Scalability

- Toy problem „Burger‘s equation“
  \[ u_{i}^{n+1} = u_{i}^{n} + \frac{dt}{dx} u_{i}^{n} (u_{i}^{n} - u_{i-1}^{n}) + \nu \frac{dt}{dx^2} (u_{i-1}^{n} - 2u_{i}^{n} + u_{i+1}^{n}) \]

- Domain size: \( N=2\,097\,152 \)
- Platform: Mistral@DKRZ
  1 node=2x18 Broadwell cores (E5-2695v4)
- Implementation: C++, OpenMP/MPI with communication hiding
- Single-node perf.: ca. 80% peak (mem-bound)
Challenge: Scalability

- Toy problem „Burger‘s equation“
  \[ u_i^{n+1} = u_i^n + \frac{\Delta t}{\Delta x} u_i^n (u_i^n - u_{i-1}^n) + \nu \frac{\Delta t}{\Delta x^2} (u_{i-1}^n - 2u_i^n + u_{i+1}^n) \]

- Domain size: \( N = 2 \, 167,384 \)
- Platform: Mistral@DKRZ
  1 node=2x18 Broadwell cores (E5-2695v4)
- Implementation: C++, OpenMP/MPI with communication hiding
- Single-node perf.: ca. 80% peak (mem-bound)

Issue: \( \Delta x \downarrow \Rightarrow \Delta t \downarrow \downarrow \downarrow \downarrow \)
Scalability of ICON: Local High-resolution

Experiment hdcp2_lam_240m on JUQUEEN

Excerpt refactoring list, HD(CP)^2
- compute decomposition (fixed by using distributed algorithm)
- compute local halo information (fixed by rewriting algorithm)
- generate local grid partition (fixed by using distributed data structures; based on shared mem.)
- store decomposition information (fixed by rewriting data structures)
- store gather communication pattern (fixed by using two-phase gather algorithm)
Setup: Germany at 120m resolution
NWP-LEM + 42 270 720 horizontal cells, 160 levels
Parallel efficiency of 71-80% at 131 072 cores,
corresponding to 322.5 x 160 = 51 600 process-local (volumetric) cells
Scalability of ICON: Global High-resolution

- IFS: Hydrostatic; ICON: Non-hydrostatic
- Desire for exascale: If I had 10M cores,...
  - ...I could solve 1.25km global simulations at 440 days/day
  - ...if we can retain scalability with 2080 local cells (33 horizontal cells)
- ...I’d have trouble with big data: 20 TB/forecast day or 8800TB/compute day
  → neither storing nor (brute-force) writing this amount of data is a good idea!
Scalability of ICON: Global High-resolution

- IFS: Hydrostatic; ICON: Non-hydrostatic
- Desire for exascale: If I had 10M cores, ...
  - ...I could solve 1.25km global simulations at 440 days/day
  - ...if we can retain scalability with 2080 local cells (33 horizontal cells)
  - ...I’d have trouble with big data: 20 TB/forecast day or 8800TB/compute day
  - → neither storing nor (brute-force) writing this amount of data is a good idea!
Work in Progress: Asynchronous I/O

Worker Processes

File System

I/O Processes

1 GB/s
Work in Progress: Asynchronous I/O

Worker Processes

I/O Processes

Parallel File I/O

File System

5 GB/s for NetCDF
7 GB/s for GRIB
Work in Progress: Online Diagnostics

- Online-Diagnostics via the Modular Earth Submodel System (MESSy): VIsual Satellite Operator (VISOP)

Offline, full 3D
26.04.2013, MODIS

Offline, columns

Online, columns (VISOP)
Leonhard Scheck, LMU
Summary

- **ESiWACE** – Joining forces to explore computability of extreme-scale weather and climate simulations
  - ISC, June 2017, Frankfurt: BoF session on cloud resolving models
    - speakers from USA, Japan, China, Europe
  - HPC Summit week, May 2017, Barcelona: ½-day ESiWACE workshop
  - PASC, June 2017, Lugano: Minisymposia in weather & climate tracks
- Current models suggest $O(1-10)$ SYPD to be doable at extreme scale (not counting in I/O...)
  - ICON: NWP 2.5km global
    - ICON-ECHAM 1.2km global (aqua-planet experiments)
- ICON: Scalability for local and global high-resolution simulations

Contacts: [neumann@dkrz.de](mailto:neumann@dkrz.de), [www.esiwace.eu](http://www.esiwace.eu)

Acknowledgement:
ESiWACE has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 675191. Parts of the work were funded by the Federal Ministry of Education and Research in Germany through the research programme High Definition Clouds and Precipitation for Advancing Climate Prediction (HD(2)CP). The authors gratefully acknowledge the computing time granted by the German Climate Computing Centre and Jülich Supercomputing Centre.