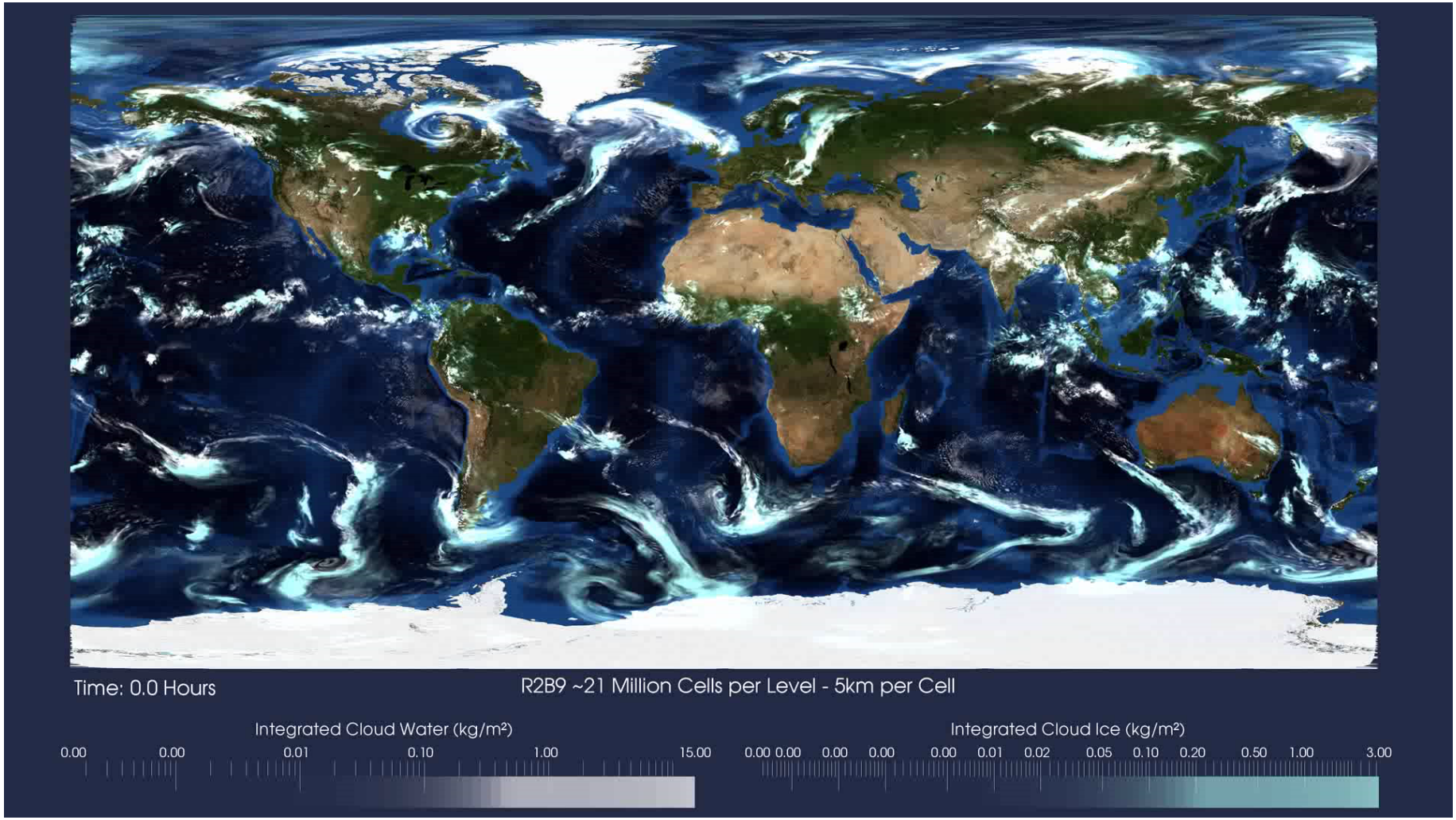
The ESiWACE Demonstrators: Scalability, Performance Prediction, Evaluation

Philipp Neumann
Deutsches Klimarechenzentrum (DKRZ)

&

The ESiWACE Team

Global High-Resolution Simulations



Top 10 Reasons…Why 1km Would Be A Great Leap

Bjorn Stevens. ExtremeEarth. Presentation @EGU 2018, Vienna:

1. Convection is resolved (rainbands).
2. Surface orographic effects and gravity waves are resolved (storm tracks).
3. Shallow circulations (clouds and convection, feedbacks and forcing).
4. Ocean eddies are resolved (southern ocean stratification).
5. Tropopause dynamics are resolved (storm tracks and stratospheric water vapor).
6. Bathymetric effects on water mass formation are resolved (variability).
7. Allows a native representation of land surface (land use changes).
8. Remaining problems, such as microphysics & turbulence become tractable (parameterization).
9. Simulates observables (brings different science communities to the same table).
10. Direct link to impacts (connects directly to application communities).

Overview

1. ESiWACE: Overview and Goals
2. Scalability
3. Performance Prediction
4. Evaluation: The DYAMOND project
5. Summary

ESiWACE: Overview and Goals

- ESiWACE = Centre of **E**xcellence in **Si**mulation of **W**eather **a**nd **C**limate in **E**urope
- 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)

Coordinator: DKRZ Deutsches Klimarechenzentrum

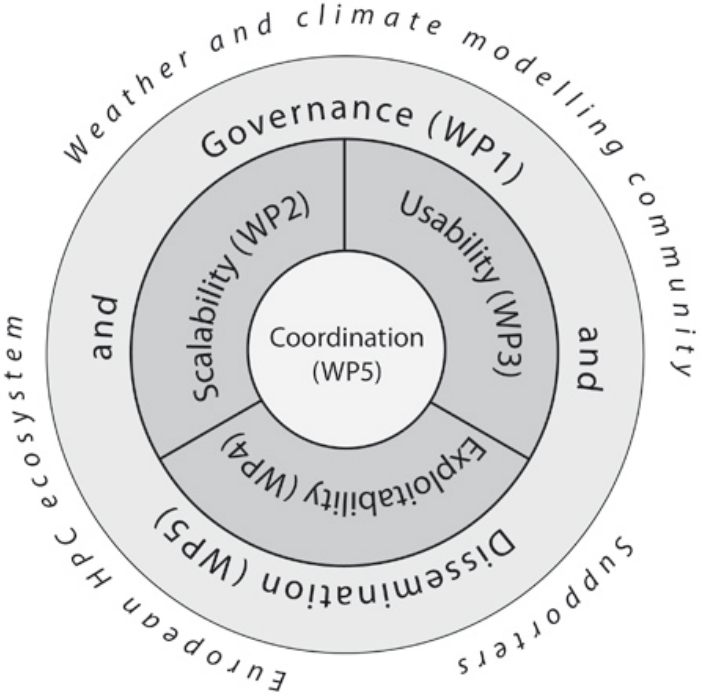
WEATHER CLIMATE HPC

ECMWF | Max-Planck-Institut für Meteorologie | BSC | Bull atos technologies | SEAGATE

DWD Deutscher Wetterdienst Wetter und Klima aus einer Hand | Sciences de l'environnement Institut Pierre Simon Laplace | cmcc | CERFACS | allinea

Met Office | SMHI | National Centre for Atmospheric Science Natural Environment Research Council | Science & Technology Facilities Council | ICHEC Irish Centre for High-End Computing

ESiWACE: Overview and Goals



WP1 Governance and engagement

WP2 Scalability

Global high resolution model demonstrators

→ **ICON, IFS, EC-Earth, NEMO**

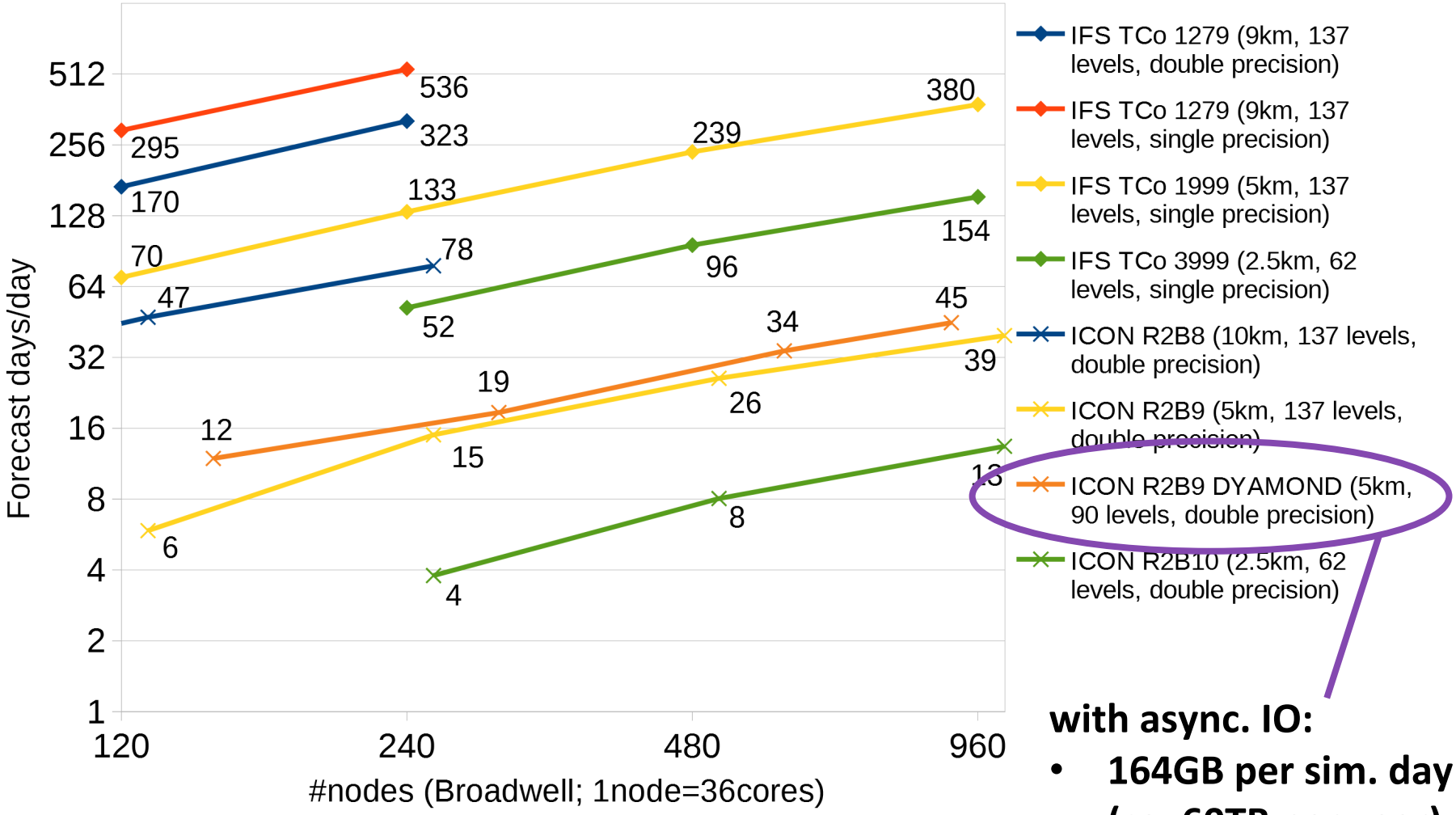
WP3 Usability

WP4 Exploitability

WP5 Management and dissemination

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.

Scalability of ICON and IFS



with async. IO:

- **164GB per sim. day (ca. 60TB per year)**
- **682GB checkpoint**

I/O in Numbers: Outtakes from ICON-DYAMOND run

Nodes	No I/O procs	wrt_output (s)
150	6	1091
300	6	1332
600	6	1661
600	11	863
900	15	749

→ How to determine optimal splitting?

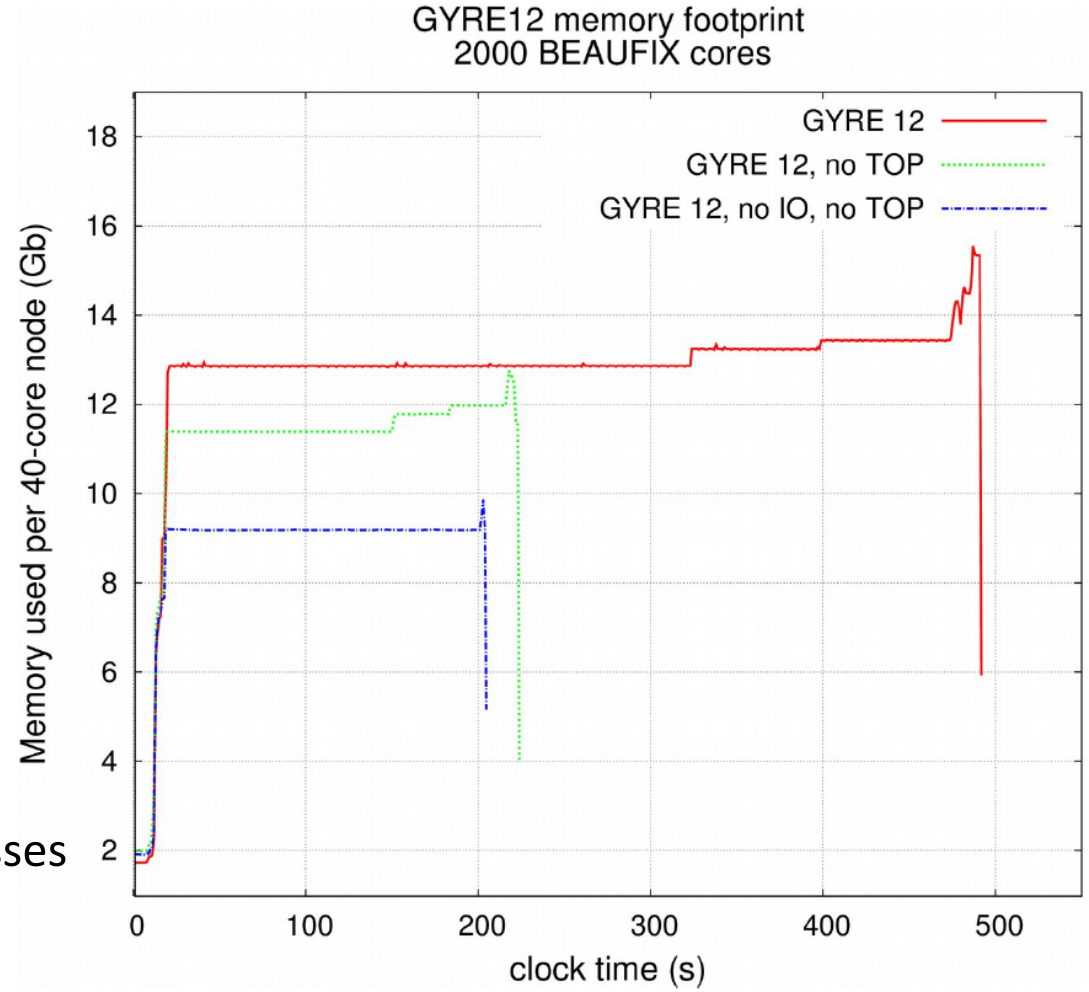
I/O in Numbers: Grib vs Netcdf (ICON-DYAMOND 5km)

900 nodes (Mistral, compute2), 15 IO procs,
1 simulated day, 2D/3D/rh,omega output every 15min/3h/15min

filename	variables	grb (GB)	nc (GB)	ratio nc/grb
atm1_2d_ml_20160801T000000Z	tqv_dia, tqc_dia, tqi_dia, tqg, tqs	5.7	38	6.7
atm2_2d_ml_20160801T000000Z	clct, lhfl_s, shfl_s, pres_sfc, tot_prec, cape_ml	9.9	46	4.6
atm_2d_avg_ml_20160801T000000Z	asob_s, athb_s, asob_t, athb_t, asou_t, asodifu_s, athd_s, athu_s	16	61	3.8
atm3_2d_ml_20160801T000000Z	u_10m, v_10m, t_2m, qv_2m, tqr	9.1	38	4.2
atm_3d_pres_ml_20160801T000000Z	pres	9.1	49	5.4
atm_3d_qv_ml_20160801T000000Z	qv	13	49	3.8
atm_3d_t_ml_20160801T000000Z	temp	13	49	3.8
atm_3d_tot_qc_dia_ml_20160801T000000Z	tot_qc_dia	1.4	49	35.0
atm_3d_tot_qi_dia_ml_20160801T000000Z	tot_qi_dia	0.96	49	51.0
atm_3d_u_ml_20160801T000000Z	u	14	49	3.5
atm_3d_v_ml_20160801T000000Z	v	14	49	3.5
atm_3d_w_ml_20160801T000000Z	w	12	49	4.1
atm4_2d_ml_20160801T000000Z	cin_ml, t_g, qv_s, umfl_s, vmfl_s	8	38	4.8
atm_omega_3d_pl_20160801T000000Z	omega	9	38	4.2
atm_rh_3d_pl_20160801T000000Z	rh	14	38	2.7

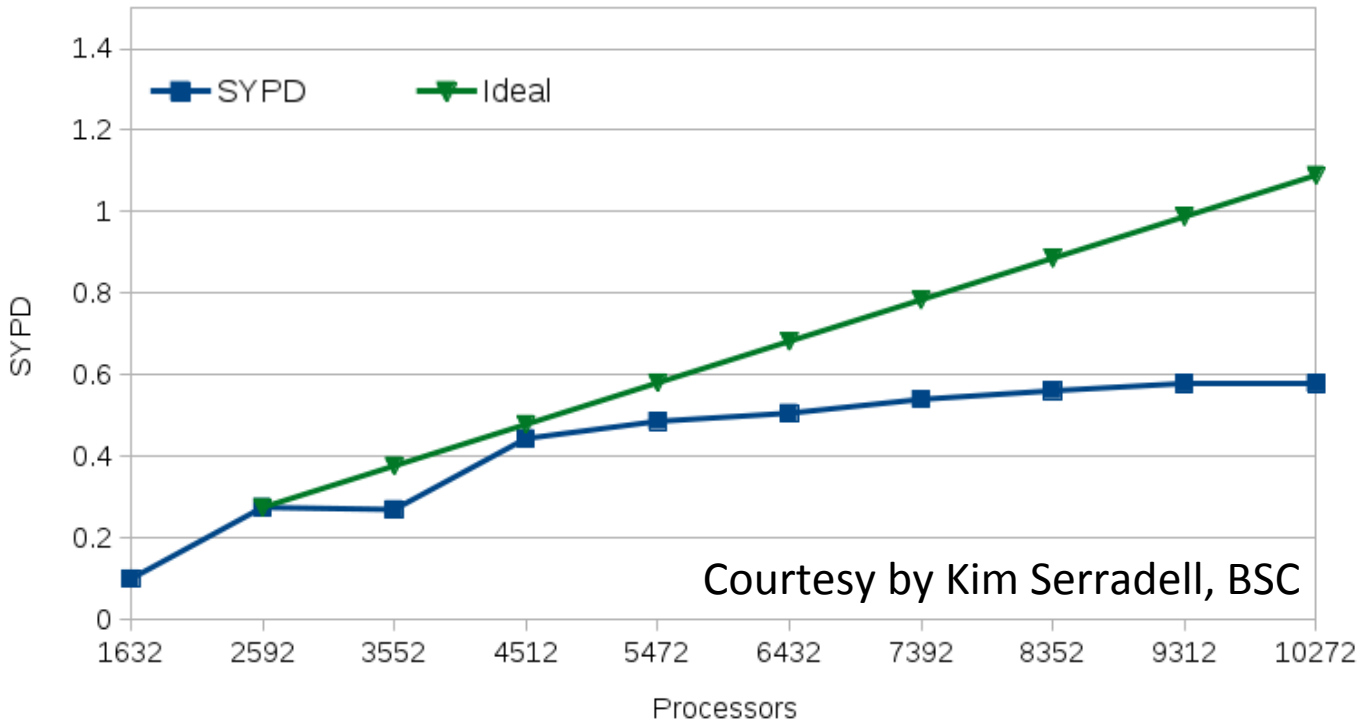
NEMO: High-Resolution Simulations and Exascale

- Estimate GYRE-KM memory requirements via GYRE12 (1/12°, ca. 8km)
  - → removal of global arrays
  - → **preliminary conclusion: memory is the limit (29TB estimate for GYRE-KM)**
- Estimates of scalability limits and power consumption
  - → **on up to 24 000 MPI** ranks on BEAUFIX@Météo-France
- Code optimisation
  - **8% gains** from vectorisation, communication, memory accesses
  - **5% gains** from hybrid parallelisation
  - Single precision experimentation



Courtesy by Eric Maisonnave, Cerfacs

EC-Earth Demonstrator: Scalability on MareNostrum4



- Configuration T1279 - ORCA12 (ca 16km res.; IFS+NEMO)
- Technical issues with OPA network and large number of cores
- Run with real production output (3.5 TB per year)
- Low amount of SYPD (360s timestep):
  - In IFS, I/O is done by a single process
  - IFS-36r4 is an old release (Nov. 2010)

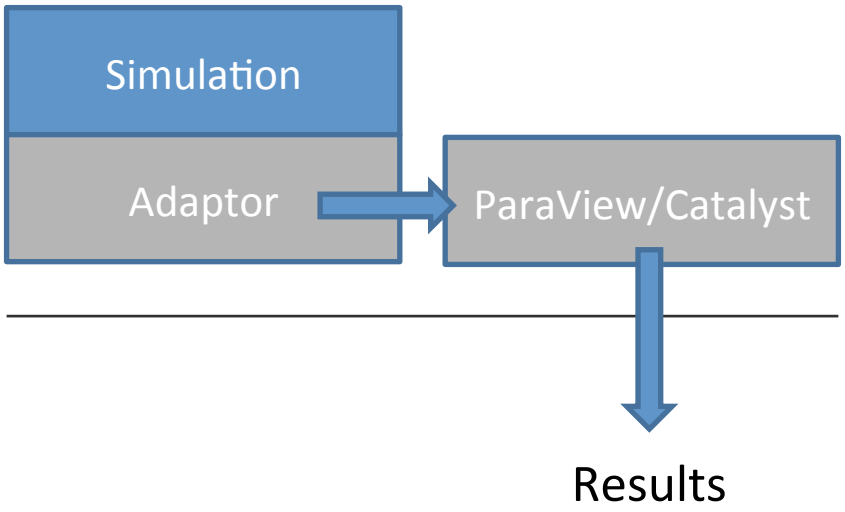
Scalability: Perspectives

- **0.1-0.3 SYPD at 5km, atmosphere-only**
- Long-term goal:
  1km atmosphere-ocean simulations at 1 SYPD
  3km ensemble atmosphere-ocean simulations
  → **this will require at least exascale computing** and corresponding **data capacity/handling capability**
  → challenge: model=long-term development
- **ESiWACE: The evolutionary path**
  → optimise current (production) models
- **ESCAPE/ESCAPE-2: The revolutionary path**
  - DSLs to enhance programmability/portability
  - Mixed precision arithmetics
  - Increasing the levels of concurrency (dwarf concept)

Big Data Visualisation: Towards Exascale…

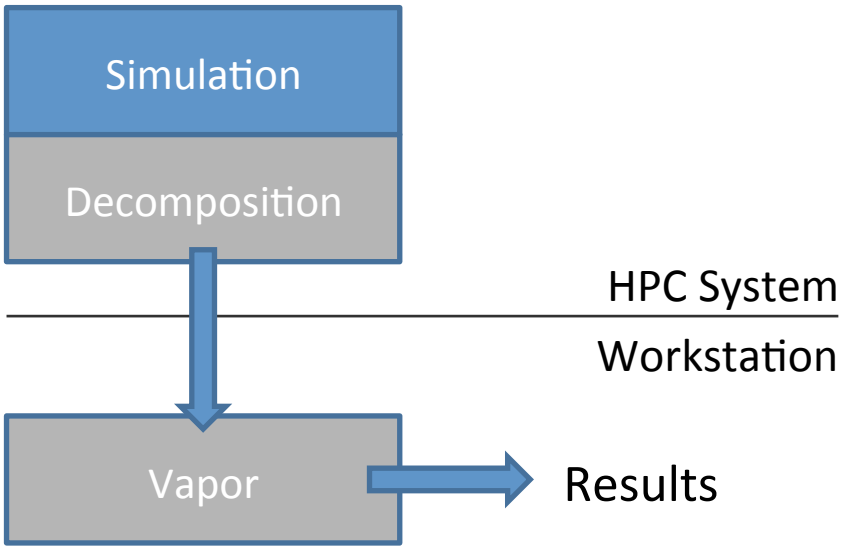
in-situ Visualisation

(ParaView/Catalyst/Cinema)



in-situ Compression

(Vapor)



17/05/18

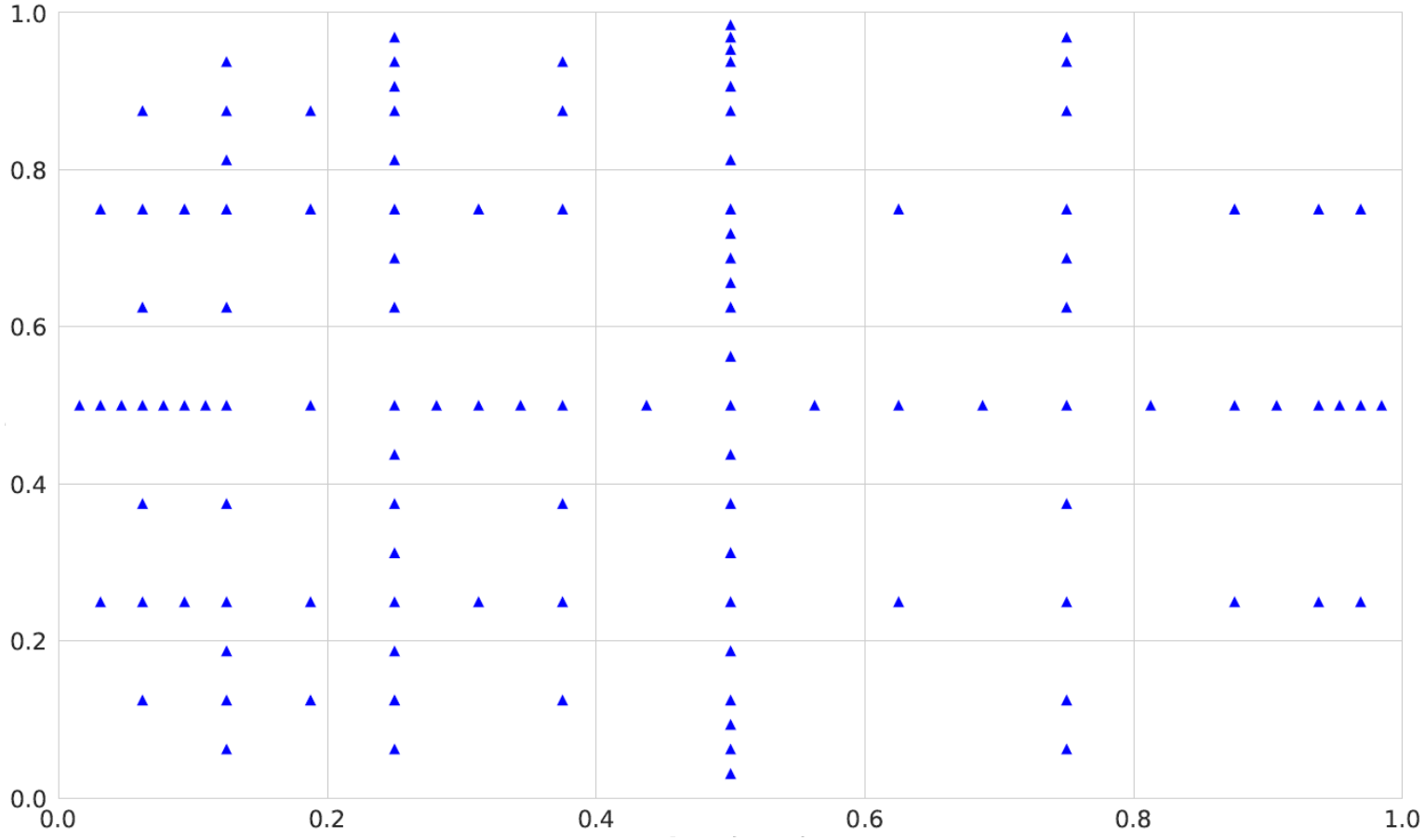
Performance Prediction: Objective

- **Multi-parameter influence on computational performance**
  - → **computational:** OpenMP/MPI decomposition, loop-blocking, vector lengths, …
  - → **algorithmic:** time step, number of iterations, error control/tolerance,…
  - → **all aforementioned categories for every model subcomponent**
    - → **high-dimensional parameter space**
- **Objective: performance estimate for complex ESMs…**

  …to gain insight into (wanted or unwanted) hotspots

  …to improve scheduling (relevant to workflows?)
- **<u>Approach:</u> Regression on high-dimensional parameter space via adaptive sparse grids**

Sparse Grids



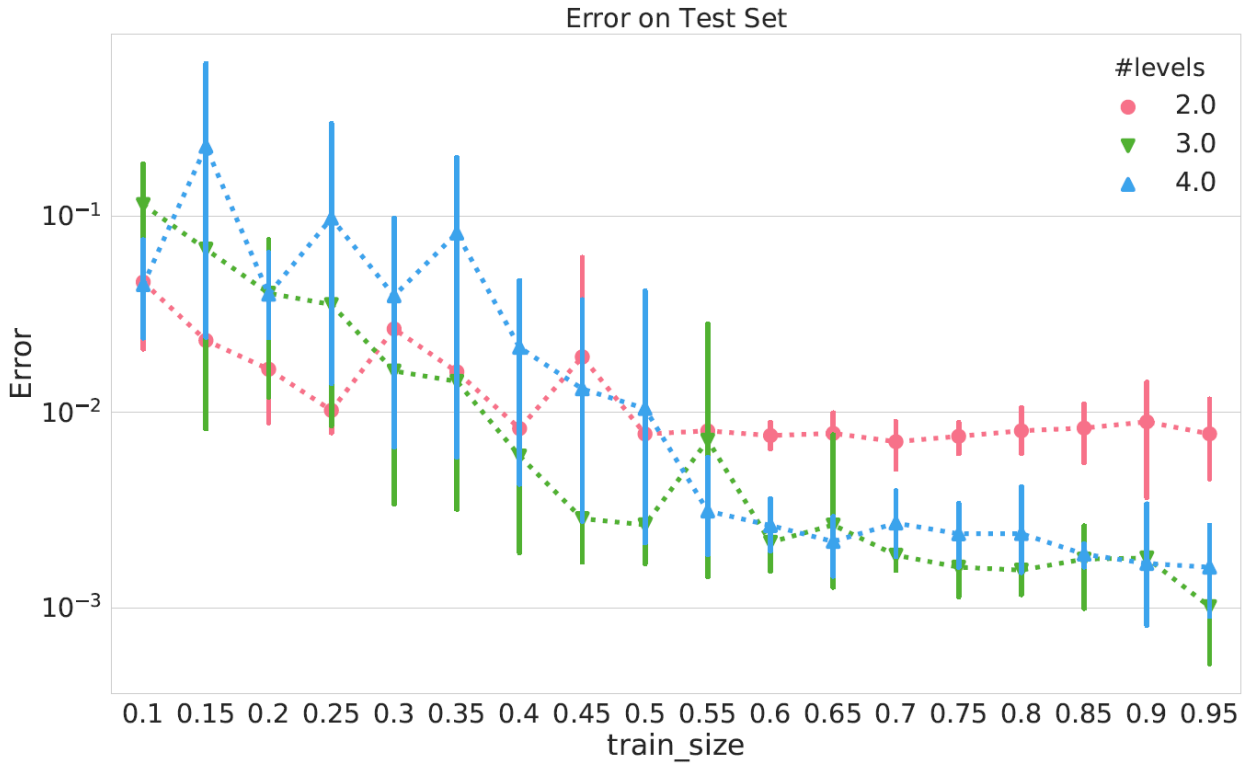
Theorem 1 *For the interpolation error of a function* $f \in H^2_{0,mix}$ *in the sparse grid space* $V^s_{0,n}$ *holds*

$$||f - f_n^s||_2 = \mathcal{O}(h_n^2 \log(h_n^{-1})^{d-1}).$$

SG: O(N(log N)$^{d-1}$) points

Full grid: O(N^d) points

Performance Prediction: Sparse Grid Regression



- Configuration: ICON-DYAMOND R2B4 (160km global res.), no I/O
- Run times on single-node (dual-socket Broadwell)
- Parameters: number OpenMP threads/MPI tasks
  loop-blocking (nproma)
- Acknowledged: Paula Harder, DKRZ

Evaluation/Perspectives: The DYAMOND Project

- DYAMOND= **DY**namics of the **A**tmospheric general circulation **M**odeled **O**n **N**on-hydrostatic **D**omains
- Goal: Intercomparison of global high-resolution models
- Participation list:
  - ICON/ Luis Kornblueh
  - NICAM/ Ryosuke Shibuya, Chihiro Kodama
  - MPAS/ Falko Judt
  - nu-FV3/ Shiann-Jiann Linn
  - SAM/ Marat Khairoutdinov
  - NASA GEOS5/ William Putman
  - UM/ Pier Luigi Vidale

- Data management and provision via DKRZ
- More information: www.esiwace.eu/services/dyamond

Summary

- ESiWACE – Joining forces to explore computability of extreme-scale weather and climate simulations
  - European Seminar on Computing, June 3-8 2018, Pilsen/Czech Rep.
  - Teratec Forum, June 19-20 2018, Palaiseau/France
  - ISC, June 24-28 2018, Frankfurt/Germany
  - PASC, July 2-4 2018, Basel/Switzerland
- 5km global resolution simulations incl. IO at O(0.1-0.3) SYPD
  → still some way to 1 SYPD…
- Performance predictions with sparse grids deliver accurate run time estimates in various applications
- DYAMOND: A project to intercompare global high-resolution models

Contacts: neumann@dkrz.de, www.esiwace.eu

Acknowledgement:
ESiWACE has received funding from the **European Union's Horizon 2020 research and innovation programme** under grant agreement No 675191. The authors gratefully acknowledge the computing time granted by the **German Climate Computing Centre**.

Backup: The Purpose of Simulation is Insight…

See, understand, learn, communicate …

- Confirmatory visualisation
- Interactive visualisation
- Animations & stills for communication