Deutscher Wetterdienst
Wetter und Klima aus einer Hand

The Enhanced DWD-RAPS Suite

Testing Computers, Compilers and More?

Ulrich Schättler, Florian Prill, Harald Anlauf

Deutscher Wetterdienst

Research and Development

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

The Enhanced DWD-RAPS Suite

Testing Compilers, Computers and More?

Ulrich Schättler, Florian Prill, Harald Anlauf

Deutscher Wetterdienst

Research and Development

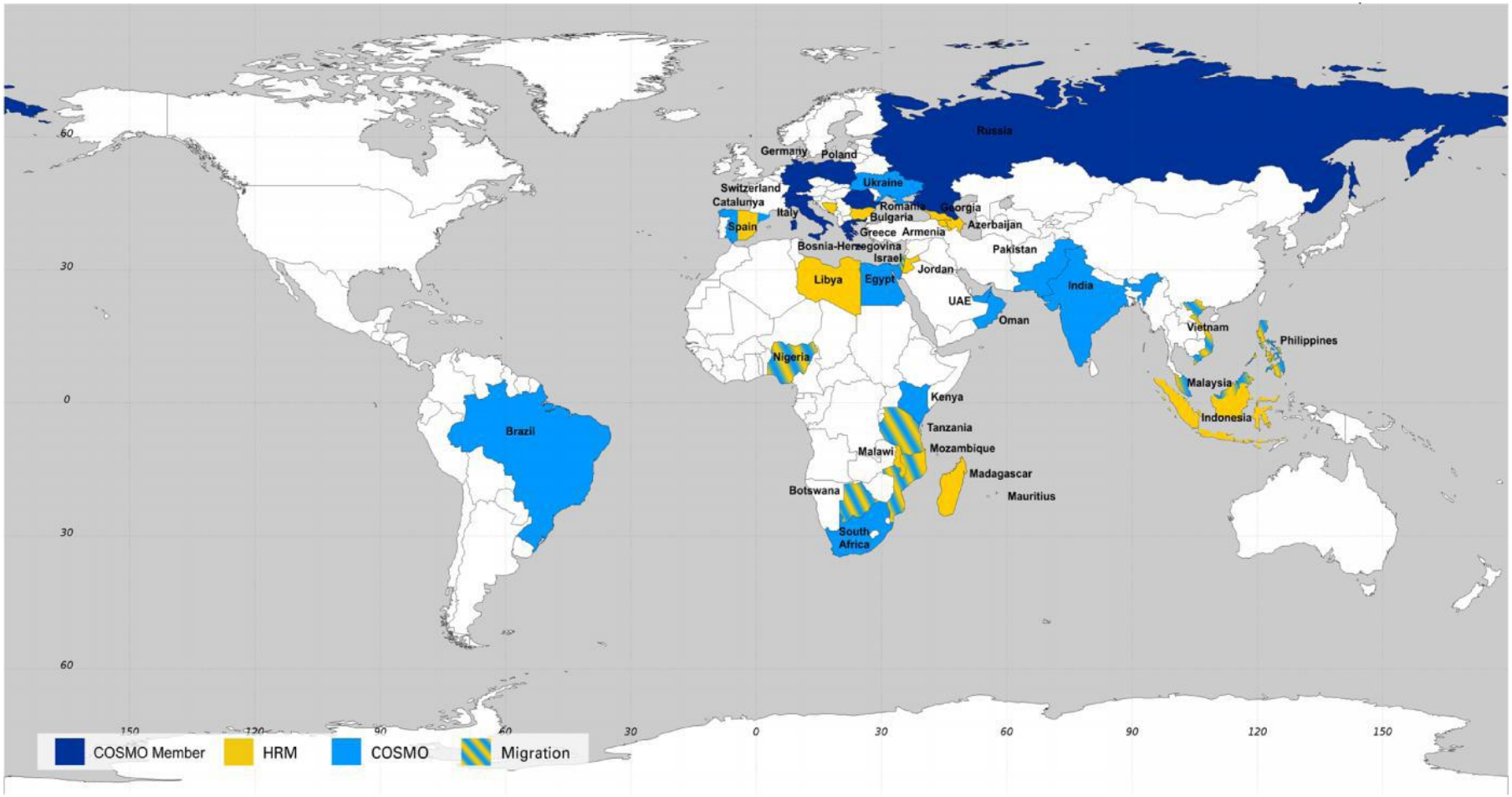
Contents

- Why an enhanced RAPS Suite?
- Contents of the DWD-RAPS Suite
- Testing the RAPS Suite
  - Compilers
  - Computers
  - What else?
- Future Computational Requirements at DWD

Why an enhanced RAPS Suite?

- In the last decade we only used COSMO_RAPS
  - The COSMO-Model is a non-hydrostatic, limited-area model used at DWD for deterministic forecasts over Europe and Germany
  - Since May 2012 also used operationally at DWD for convection-permitting ensemble (COSMO-DE-EPS) over Germany (development and pre-operational use started few years earlier)
  - The applications of the COSMO-Model took most of the available CPU time ($\geq$ 80%) on our HPC facilities. Therefore it was sufficient to estimate the necessary operational resources just with the COSMO-Model.
- The COSMO-Model is also used by the national weather services of the COSMO Consortium, the CCLM (COSMO CLimate Mode) and the COSMO-ART (Aerosols and Reactive Tracers) Communities
- Maintaining a RAPS version of the COSMO-Model is also beneficial for our partners.

COSMO-Model Distribution in 2012



Why an enhanced RAPS Suite?

- The situation at DWD is about to change
  - Development of a non-hydrostatic global model: ICON
    - It is planned to start operational runs with a resolution of 20 km soon
  - Development of ensemble-based data assimilation systems for ICON and the COSMO-Model
  - Consequence: The amount of CPU time used by the different models will change: ICON and the assimilation codes will have a bigger share.
- The colleagues at DWD realized:  IS COOL!
- Because when porting the models to new computers, there almost never have been problems with the COSMO-Model.
- Therefore we now have additional models in the DWD-RAPS Suite (first published in December 2011)
- And: The new suite represents typical workloads of the forthcoming operational suite (compute and I/O)

Contents of the DWD-RAPS Suite

- COSMO_RAPS 5.1:
  - non-hydrostatic limited-area model
  - written in Fortran90/95
  - pure MPI parallelization using domain decomposition
  - memory intensive; performance of actual implementation mainly limited by band-width
- ICON_RAPS 2.0:
  - non-hydrostatic multi-resolution model, targeted at O(10^4+) scalar or O(10^2+) vector cores
  - written in Fortran 95 + selected Fortran2003 features
  - hybrid MPI/OpenMP parallelization using domain decomposition and blocking
  - implementation of unstructured grid by using indirect addressing

Contents of the DWD-RAPS Suite

- 3DVAR_RAPS_1.1
  - Common code base for object oriented approach to data assimilation
    - 3D-Var: Variational data assimilation
    - LETKF: Local Ensemble Transform Kalman Filter
  - Data assimilation is very heterogenous (in particular observation processing)
  - written in Fortran 95 + selected Fortran2003 features (development unfortunately limited by least common denominator of compilers for production use)
  - pure MPI parallelization

Testing the RAPS Suite

- The suite has been developed on NEC SX-9 at DWD (of course) and parts of the suite could also be tested on a (SUN) Linux Cluster
- We thank the colleagues at the following institutions for the possibility to perform tests on their systems:
  - ECMWF: IBM pwr6
  - Deutsches Klimarechenzentrum (DKRZ): IBM pwr6 and Linux Clusters
  - Swiss National Supercomputing Centre (CSCS): Cray XE6
- These companies gave us feedback from testing the first version of the DWD-RAPS suite:
  - Cray, Fujitsu, IBM, SGI, BULL / T-Systems
  - Many thanks to all benchmarkers for their work

Testing Compilers (Tests done in March 2012)

		COSMO	ICON	3DVAR
NEC	sxf90	rev430	rev451	rev451
IBM	xlf	12.1.0.0	13.1.0.7 necessary	14.0.0x necessary
Cray	native	8.0.1 and 8.0.2.101	8.0.2.101 necessary	1 Routine did not compile
	Intel	12.0.3.174	12.0.3.174	12.0.3.174
Linux	Sun Studio	---	---	12.1
	GNU Fortran	4.5	4.5	≥ 4.3

„necessary“ means: older versions of the compiler failed

Testing Compilers: Lessons Learned

- Why has the COSMO-Model the fewest problems with compilers?
  - ☐ because of the programming skills of the developers
  - ☑ it mainly uses Fortran90/95 (the code is aging slowly)
  - ☑ it has the longest RAPS history
- ICON and 3DVAR make use of Fortran2003 features. Even if many compilers do support these features now, they are not yet fully matured.
- Always ask for the latest compiler available!

- By the way: The actual standard is Fortran 2008!

Testing Computers

- All benchmarks were tested by Dr. Rainer Johanni (external consultant) on the different platforms.
- Some problems could be identified and resolved, e.g.:
  - COSMO: problems when using many (!) asynchronous I/O processes
  - ICON: hybrid runs on Cray possible only with reduced compiler optimization (-O1)
  - 3DVAR / LETKF: required prerelease of Cray Fortran Compiler for successful compilation
- Results from the tests:
  - Sorry: Because of the ongoing procurement we do not want to go into detail
  - Some numbers from earlier tests given later

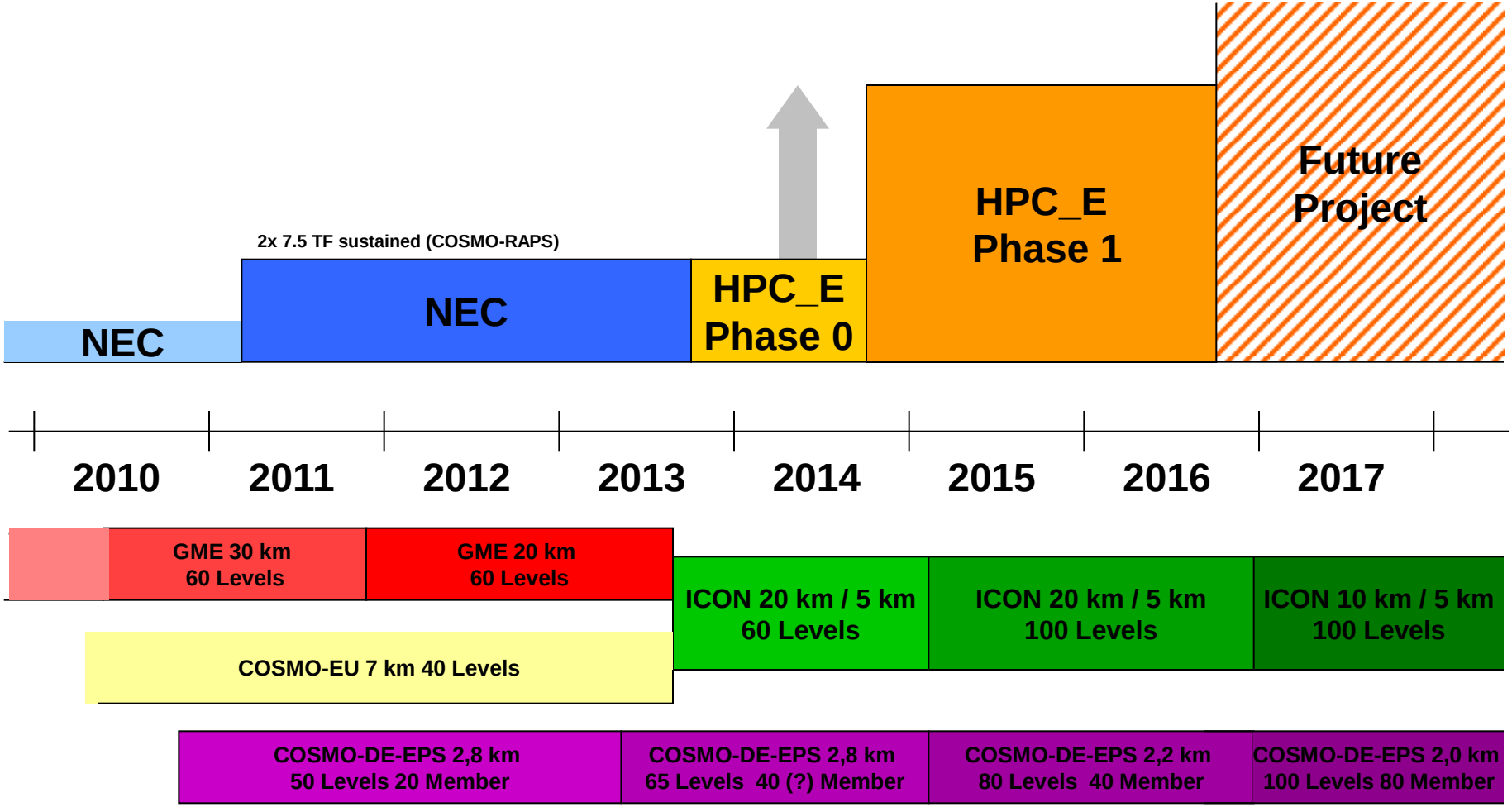
What else can be tested?

- Running models in (DWD) operational mode: Preemption
  - High priority jobs can overlay experimental jobs by just stopping them. The jobs are not killed or swapped out.
  - This is possible, because all jobs do not use more than half of the memory.
  - This puts some restriction on the amount of memory that must be available
- Running models in (DWD) experimental suites: Bit-Reproducibility
  - If the same binary runs with a different processor configuration, the results must be the same.
  - From the algorithmic side we take some penalties, what do we have to expect from the compiler side?

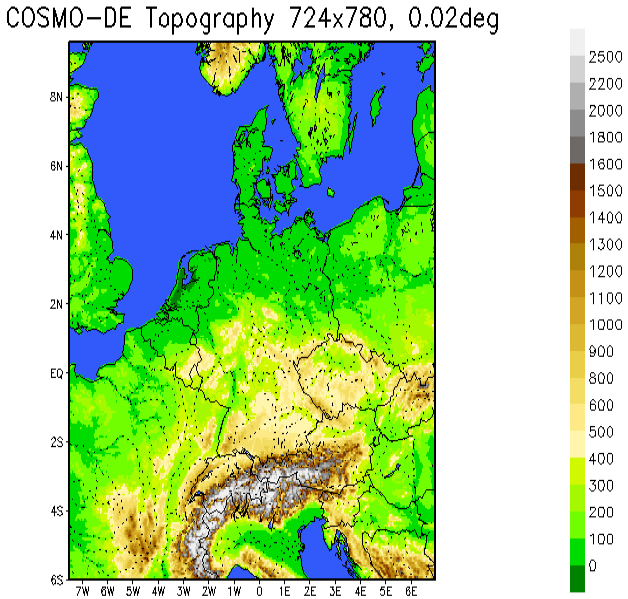
Can you test everything?

- Many problems only show up, when the full system is up and running:
  - HPC + database + archive
  - System integration always was and still is the hard job.
  - How to test these issues in advance?

Future Computational Requirements at DWD



COSMO-Model



- COSMO-DE Ensemble defines size of machine
  - 15 members
  - 724×780×65 grid points
  - 2.2 km resolution;
  - 12 h forecast in 1200 seconds
- SX-9 Tests for 1 member with early configuration
  - 1×31+1 CPUs: 1119.79 seconds; 0.470 TFlop/s
  - 1×63+1 CPUs: 675.05 seconds; 0.814 TFlop/s
  - 2×63+2 CPUs: 483.72 seconds; 1.149 TFlop/s
- IBM pwr6 Tests for 1 member with early configuration:
  - 08×16+0: 10040.77
  - 16×16+0: 4692.22
  - 16×32+0: 2412.64
  - 32×32+0: 1297.70
  - 31×32+32: 1200.27
  - 32×63+32: 692.10
  - 40×51+ 8: 662.48

Tests on DWD's NEC SX-9

# Cores (Times of NEC SX-9)	1×15+1 16	1×31+1 32	1×63+1 64	2×63+2 128
Total Time		1119.79	675.05	483.72
Dyn. Computations		565.21	307.93	196.12
Dyn. Communic. FW		21.86	22.14	23.96
Dyn. Communic. Rest		49.53	47.03	50.18
Phy. Computations		280.28	144.28	92.52
Phy. Communications		3.77	4.68	4.94
Add. Computations		91.26	55.88	40.82
Input		20.46	21.09	22.64
Output		81.43	65.58	45.88
gather_data		25.64	26.03	21.10
write		16.39	10.97	1.27

Tests on ECMWF's IBM pwr6

# Cores (Times of IBM pwr6)	16×32+00 512	31×32+32 1024	32×63+32 2048	40×51+08 2048
Total Time	2412.64	1200.27	692.10	662.48
Dyn. Computations	1077.05	486.79	228.39	227.65
Dyn. Communic. FW	268.23	84.34	38.54	34.62
Dyn. Communic. Rest	232.52	112.11	75.05	60.48
Phy. Computations	365.82	196.49	97.43	96.72
Phy. Communications	47.79	33.53	14.12	22.18
Add. Computations	183.17	91.26	35.17	24.13
Input	49.18	53.35	59.24	59.50
Output	238.94	136.84	128.77	122.10
gather_data	93.81	85.02	97.43	91.74
write	130.36	8.82	0.05	0.02


Thank you
very much
for your
attention