CONSORTIUM FOR SMALL SCALE MODELING

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Future Trends in High Performance Computing at DWD

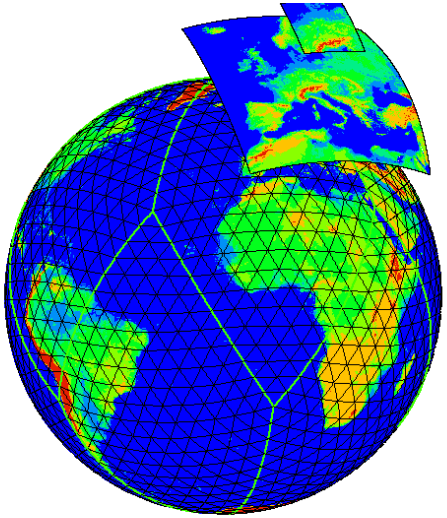
Ulrich Schättler
Deutscher Wetterdienst
Research and Development

Elisabeth Krenzien
Deutscher Wetterdienst
Technical Infrastructure

Contents

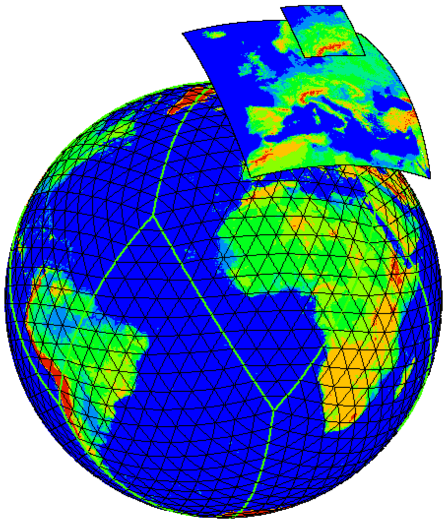
HPC at DWD

Main Parallel Models currently in use at DWD

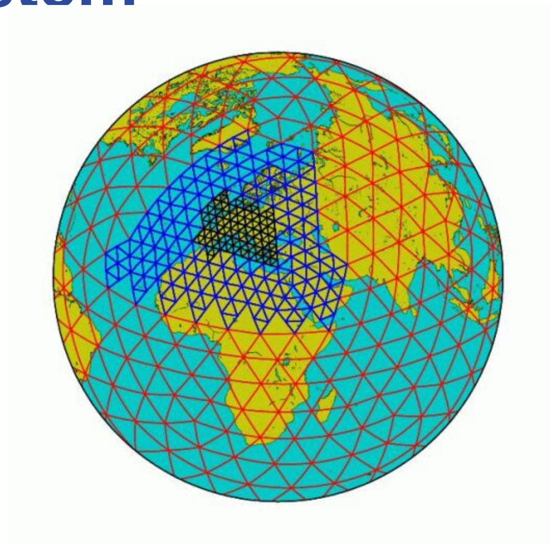


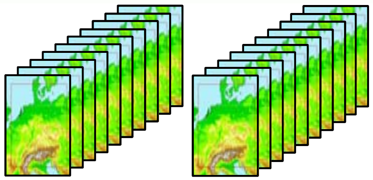
- COSMO-Model
  - This is the regional model used by different communities for a variety of applications (COSMO for NWP, COSMO-CLM, COSMO-ART)
  - DWD runs 2 applications: COSMO-EU and COSMO-DE
  - In the past, the COSMO-Model accounted for >80% of the total computing time
- GME
  - First operational global model based on a triangular grid
  - Runs now with about 30 km resolution and needs a higher percentage of the available computing time
- 3DVar
  - Replaced the OI in 2008 to allow direct assimilation of remote sensing data

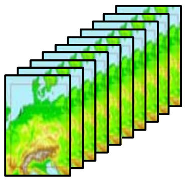
Upcoming Changes to the NWP System

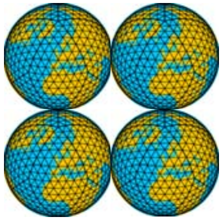


< 2011	2011 >
GME COSMO-EU	ICON with local refinement over Europe
COSMO-DE	COSMO-DE COSMO-DE-EPS
3DVar	EnKF / LETKF









CONSORTIUM FOR SMALL SCALE MODELING
COSMO

Deutscher Wetterdienst
Wetter und Klima aus einer Hand
DWD

The new Global Model
ICON

Common Development of the
Max-Planck-Institut für Meteorologie, Hamburg,
and the
Deutscher Wetterdienst

Material provided by G. Zängl and the ICON colleagues
from DWD and MPI Hamburg

Main Goals of the ICON Project

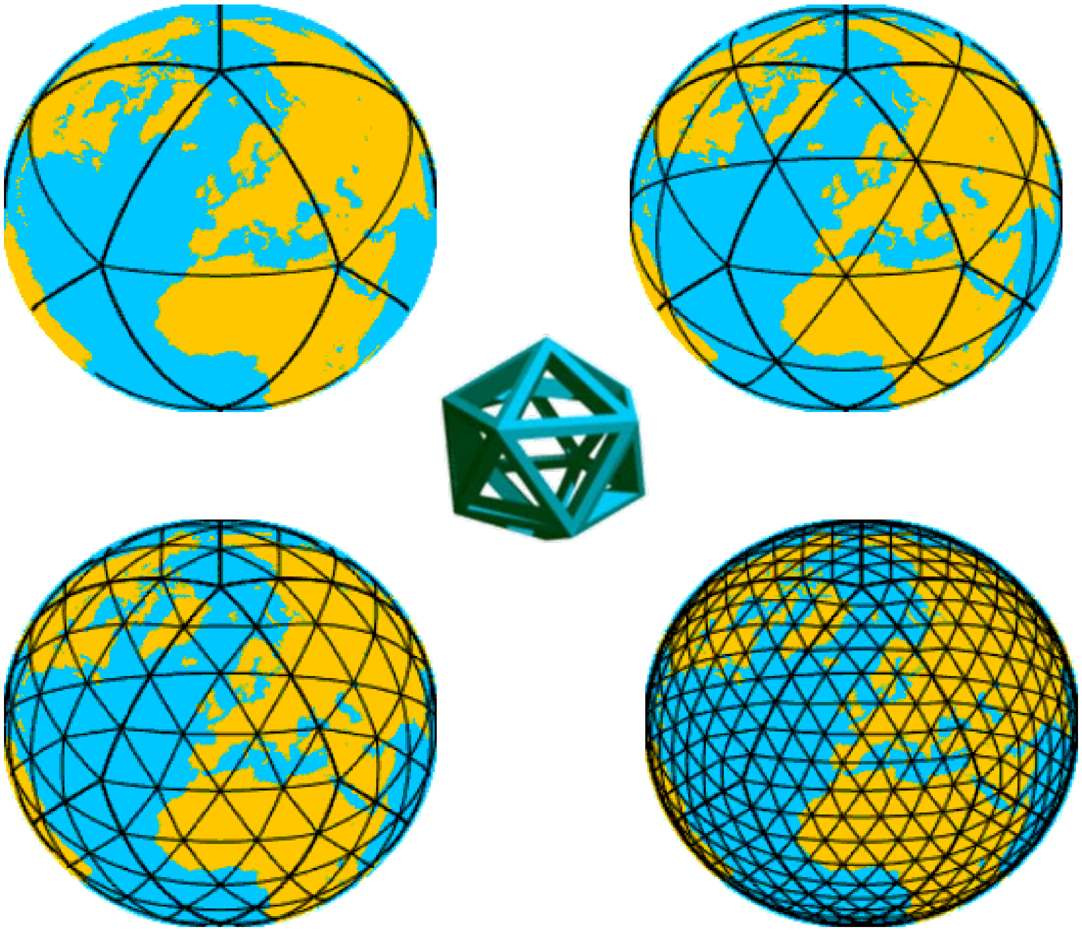
- Centralize Know-how in the field of *global modelling* at DWD and the Max-Planck-Institute (MPI-M) in Hamburg.
- Develop a *non-hydrostatic global model with static local zooming option* (ICON: ICOsahedral Non-hydrostatic; http://www.icon.enes.org/).
- At DWD: Replace global model GME and regional model COSMO-EU by ICON with a high-resolution window over Europe. Establish a library of scale-adaptive physical parameterization schemes (to be used in ICON and COSMO-DE).
- At MPI-M: Use ICON as dynamical core of an Earth System Model (COSMOS); replace regional climate model REMO. Develop an ocean model based on ICON grid structures and operators.
- DWD and MPI-M: Contribute to operational seasonal prediction in the framework of the Multi-Model Seasonal Prediction System EURO-SIP at ECMWF).
- First investigations for that project date back to 2003/04.

Requirements

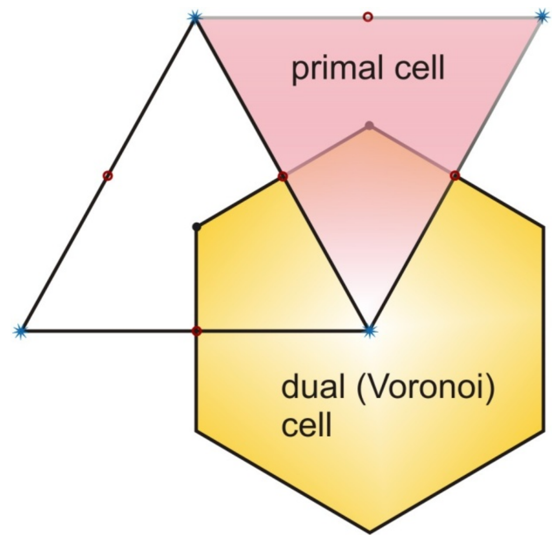
- Applicability on a wide range of scales in space and time
  → „seamless prediction“
- (Static) mesh refinement and limited area model (LAM) option
- Scale adaptive physical parameterizations
- Conservation of mass (chemistry, convection resolving), energy?
- Scalability and efficiency on massively parallel computer systems with more than 10,000 cores
- Operators of at least 2nd order accuracy

- Planned pre-operational start in 2H2011 with a resolution of about 20/10/5 km

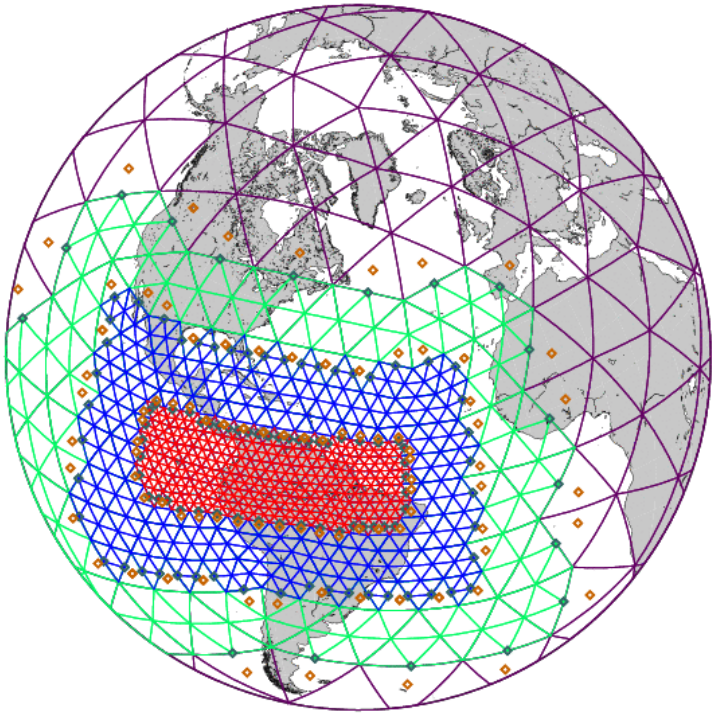
The Horizontal Grid



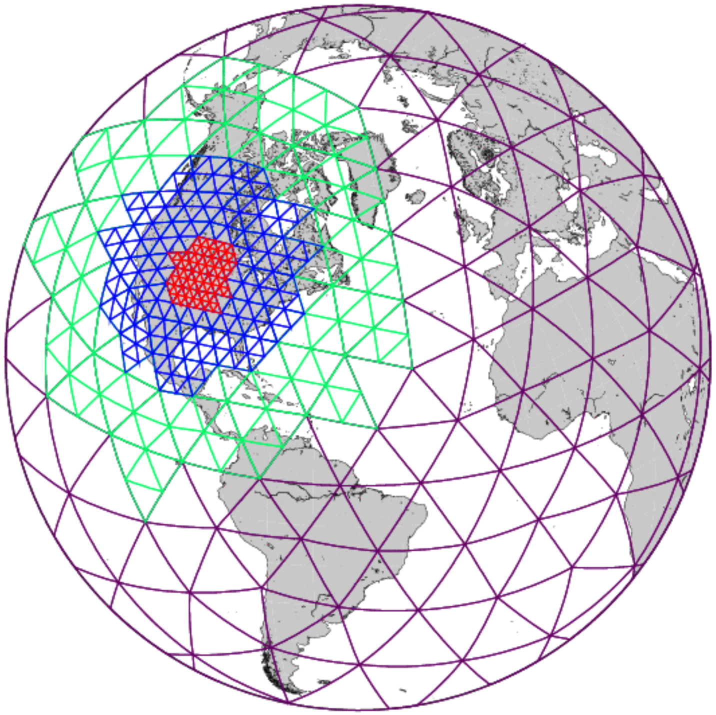
The horizontal grid is similar to the GME grid, but additional grid optimization is applied, to get nearly centered differences for gradient computations.



Static Mesh Refinement



latitude-longitude windows

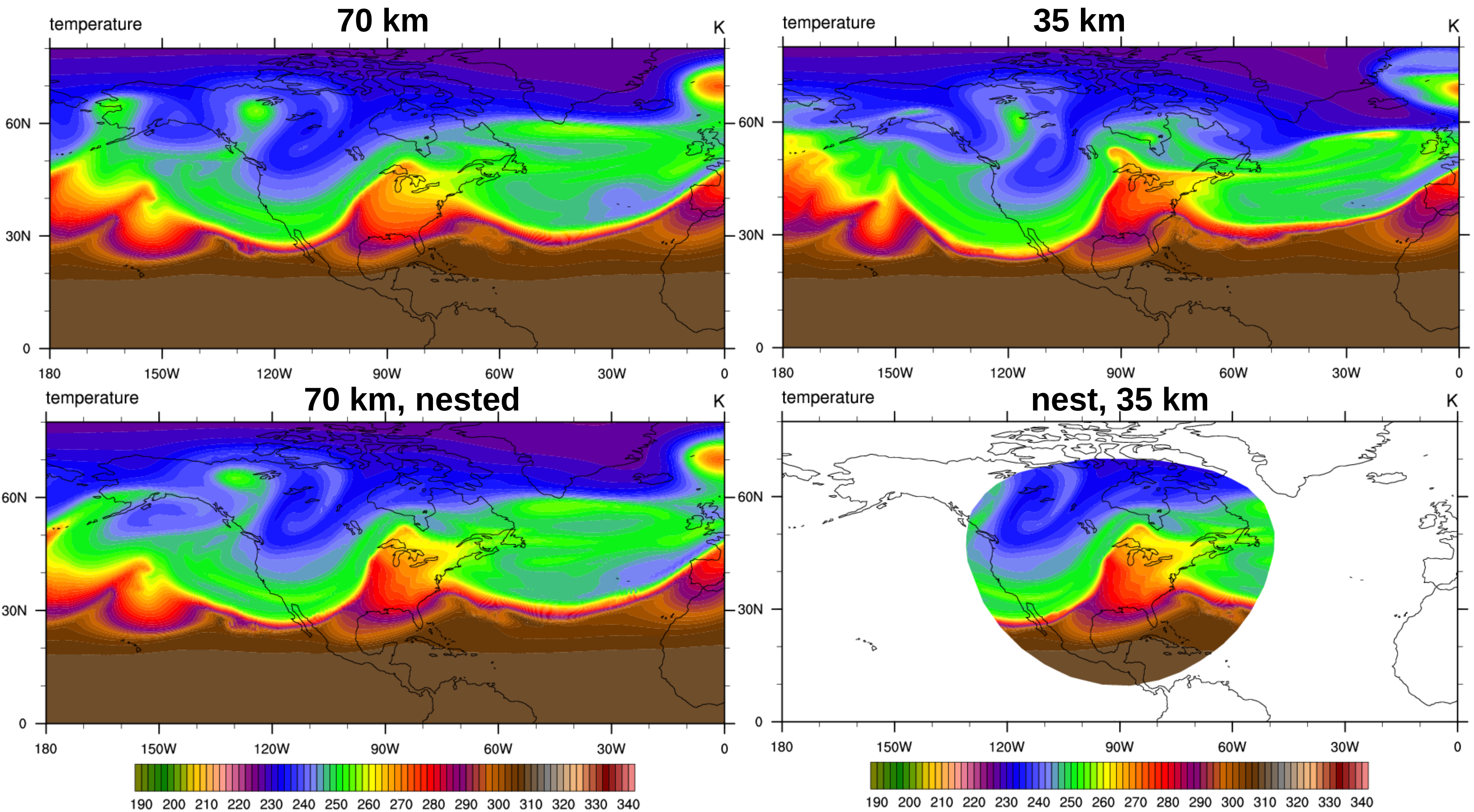


circular windows

Baroclinic Wave Test with Moisture

- Modified baroclinic wave case of Jablonowski-Williamson (2008) test suite with moisture and Seifert-Beheng (2001) cloud microphysics parameterization (one-moment version; QC, QI, QR, QS)
- Initial moisture field: RH=70% below 700 hPa, 60% between 500 and 700 hPa, 25% above 500 hPa; QV max. 17.5 g/kg to limit convective instability in tropics
- Transport schemes for moisture variables:
  - Horizontal: Miura 2nd order with flux limiter
  - Vertical: 3rd-order PPM with slope limiter
- Grid resolutions 70 km and 35 km, 35 vertical levels

- The next picture shows the temperature at the lowest model level after 14 days


temperature
70 km
K
35 km
70 km, nested
nest, 35 km
60N
30N
0
180
150W
120W
90W
60W
30W
190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340

Interesting Computational Issues

- The implementation uses the hybrid programming model MPI + OpenMP
- ICON is implemented using indirect addressing
- Times for a 10 day forecast: 70 km resolution (81920 grid points); 60 vertical levels; non-hydrostatic dynamical core with cloud microphysics and convection scheme; Δt = 150s for dynamical core and 600s for humidity advection and physics.

Time (in seconds)	NEC SX-9 8 procs (only OpenMP)	IBM pwr6 32 procs (only MPI)
without radiation	442	2286
with radiation (Δt=30 min)	2035	7080

NOTE: The numbers are just for demonstrating that something is running. They are not meant for a comparison between machines.

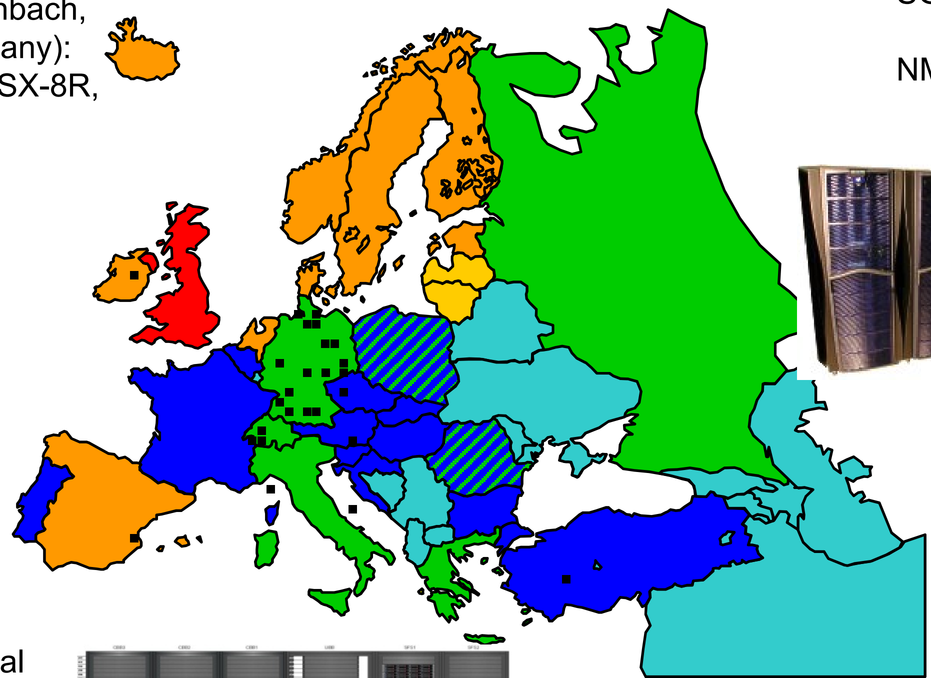
Developments for the COSMO-Model

Past Developments

- ➔ There were major changes to the COSMO-Model in the last years
  - ➔ Going to the convection-permitting scale with (2-3 km resolution) affected nearly all components of the model (dynamics, physics, assimilation)
  - ➔ The EPS based on COSMO-DE is about to start operational this month! COSMO-DE-EPS was the application for which a new computer system was purchased in 2008/09
  - ➔ Because of the growing user community, there is now a much wider range of applications: climate applications, an online-coupled aerosols and reactive tracers module (COSMO-ART)
  - ➔ And therefore much more resources are necessary for User Support
- ➔ Due to lack of time there was no re-consideration of the parallel implementation (Fortran + MPI)

COSMO NWP-Applications

DWD (Offenbach, Germany): NEC SX-8R, SX-9



Roshydromet (Moscow, Russia), SGI

NMA (Bucharest, Romania):

IMGW (Warsawa, Poland):

MeteoSwiss:
Cray XT4: COSMO-7 and COSMO-2 use 800+4 MPI-Tasks on 402 out of 448 dual core AMD nodes

USAM (Rome, Italy):
HP Linux Cluster
XEON biproc quadcore
System in preparation

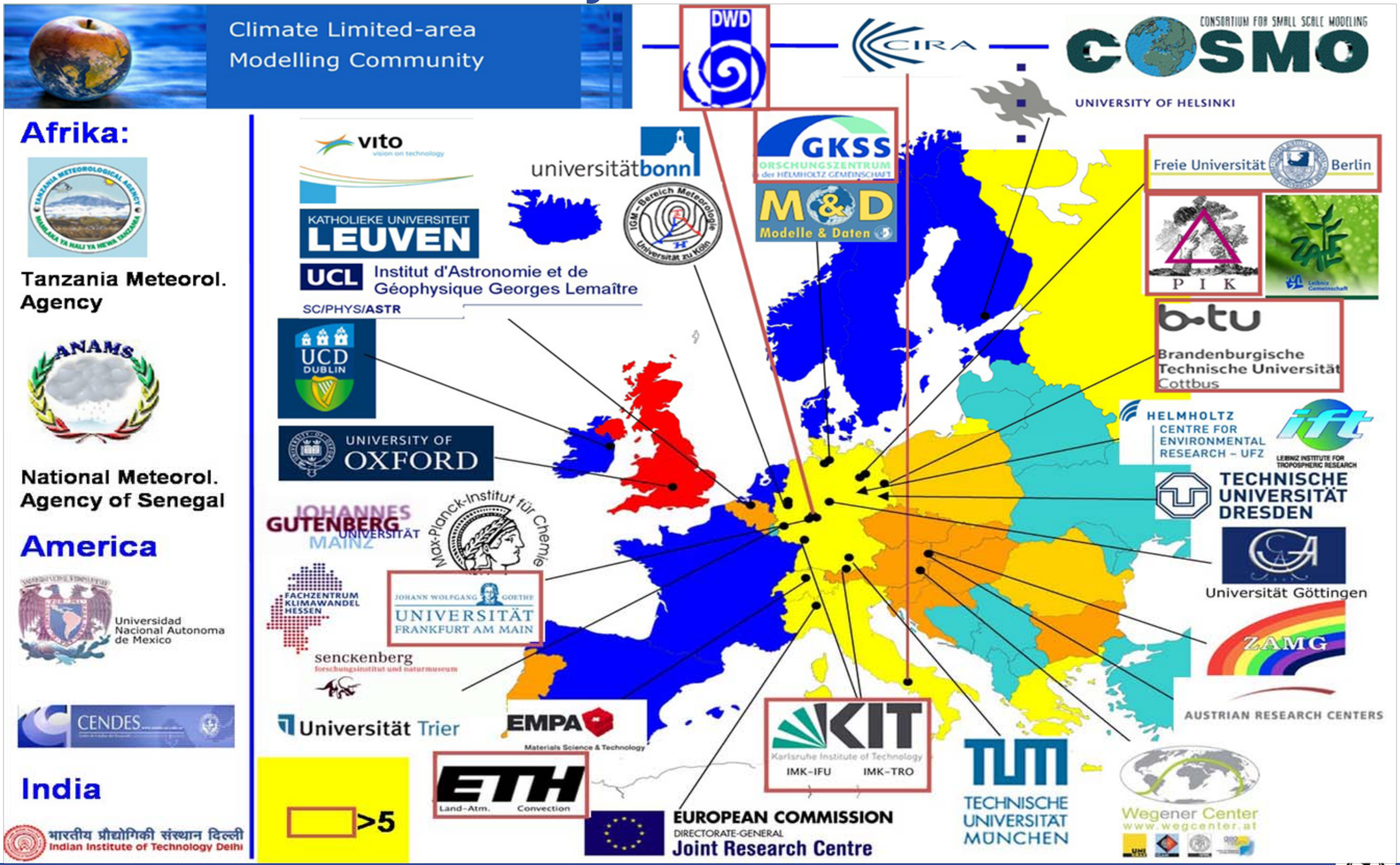
ARPA-SIM (Bologna, Italy):
IBM pwr5: up to 160 of 512 nodes at CINECA

COSMO-LEPS (at ECMWF):
running on ECMWF pwr6 as member-state time-critical application

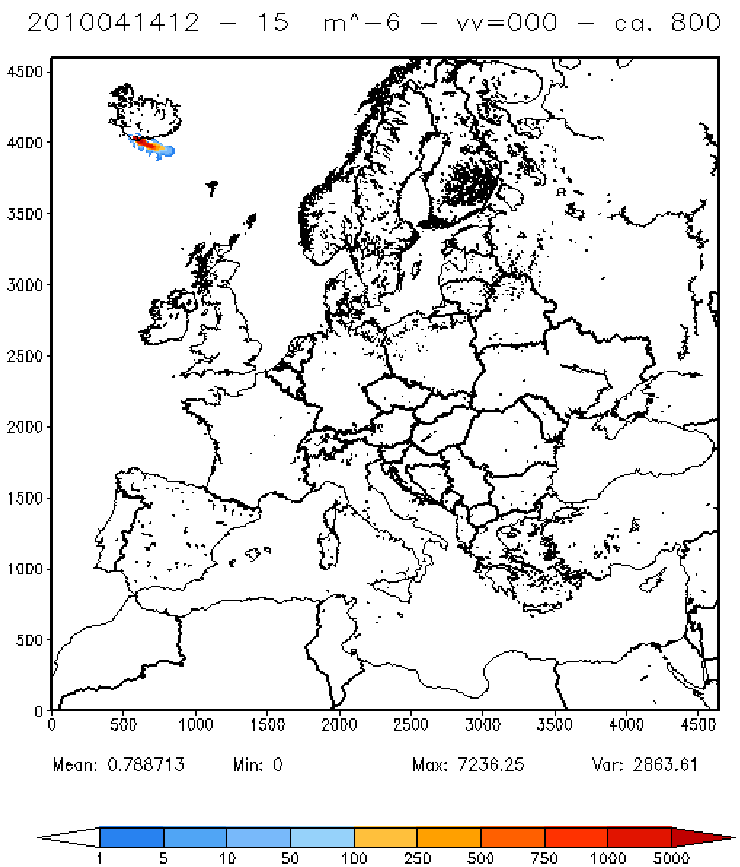
HNMS (Athens, Greece):
IBM pwr4: 120 of 256 nodes

ARPA-SIM (Bologna, Italy):
Linux-Intel x86-64 Cluster for testing (uses 56 of 120 cores)

CLM Community 2010



Eruption of Volcano Eyjafjalla



- ➔ Collaboration of
  - ➔ KIT: providing COSMO-ART
  - ➔ DLR: Flights of „Falcon“ to measure ash concentrations
  - ➔ DWD: provided forecasts where significant concentrations could be found
- ➔ DWD provided the „framework“ that these simulations could be done at all
- ➔ Setting up regular runs of COSMO-EU including COSMO-ART took only 2 days.

But what about…

- ➔ High Performance Computing?
- ➔ Scalability?
- ➔ Efficiency?

The COSMO-Model on existing Computers

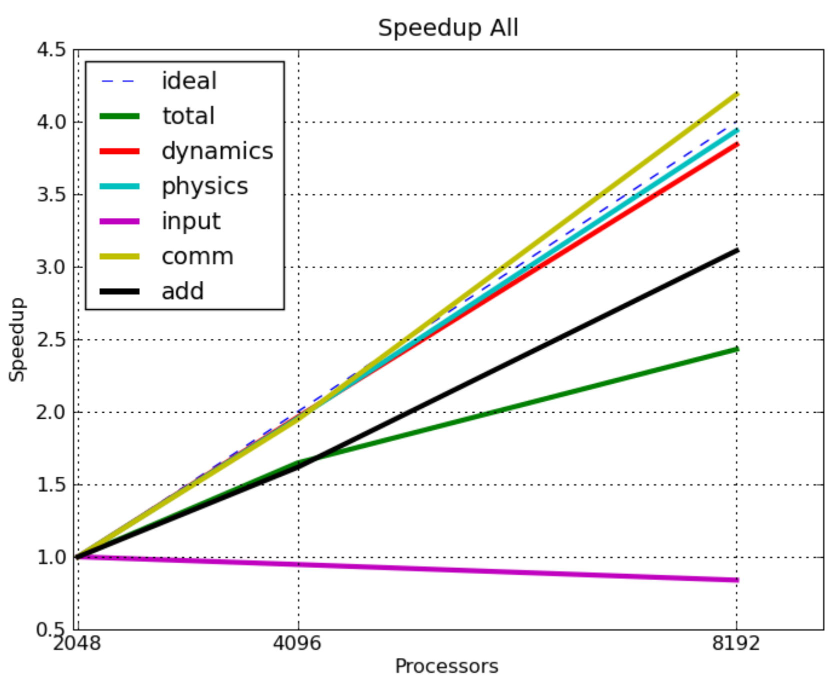
- 21 hours COSMO-DE forecast

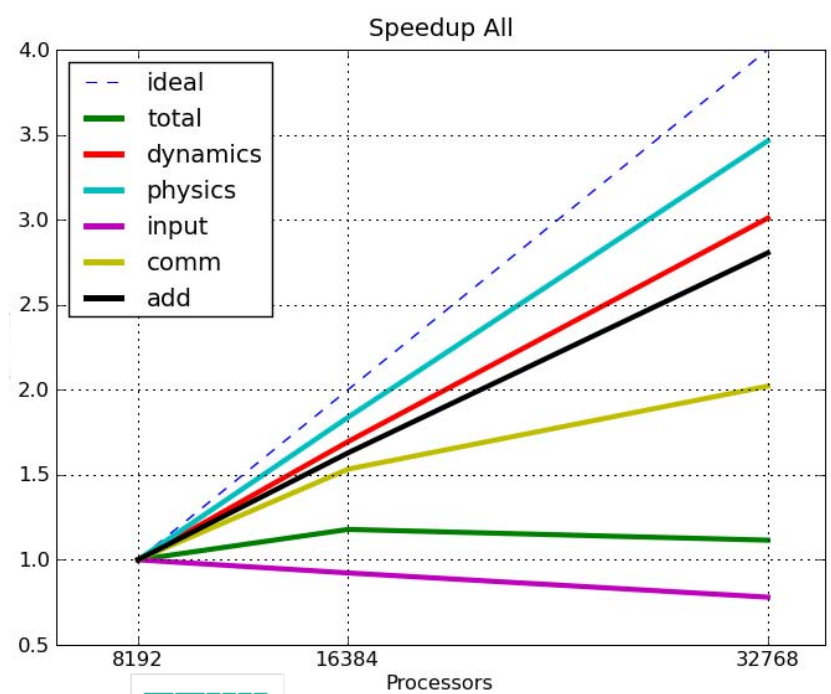
	NEC SX-9 8 Procs	IBM pwr6 256 Procs
Computations Dynamics	729.59	570.44
Computations Physics	506.18	220.45
Communications	115.61	207.69
I/O	124.43	108.40
% of I/O and Comm.	15	25

- Code efficiency
  - NEC SX-9: 13 % of peak
  - IBM pwr6: about 5-6 % of peak
  - Cray XT4: about 2-3 % of peak

Scalability of COSMO-Europe on IBM BlueGene: 1500 × 1500 × 50, 2.8 km, 3 h, no output

Today this is an „artificial" application





Results are Courtesy of Fraunhofer SCAI

Problems of the COSMO-Code

- I/O: Accessing the disks and the global communication involved heavily disturb scalability
- The message passing within a time step includes one global communication. This might be the reason of the degradation on more than 16384 processors
- The efficiency of the code is not satisfying, due to the memory boundedness of the code
  - NEC SX-9: 13 % of peak
  - IBM pwr6: about 5-6 % of peak
  - Cray XT4: about 2-3 % of peak

Tackling the Problems

- In Switzerland, a national HP2C initiative has been started, to prepare scientific software for emerging massively parallel architectures. One subproject deals with the COSMO-CLM

  Regional Climate and Weather Modeling on the Next Generations High-Performance Computers: Towards Cloud-Resolving Simulations

- COSMO accompanies this project with a *Priority Project*

  Performance on Massively Parallel Architectures

- Until end of 2012 we hope to improve not only the scalability of the model but also work on the memory bandwidth problem

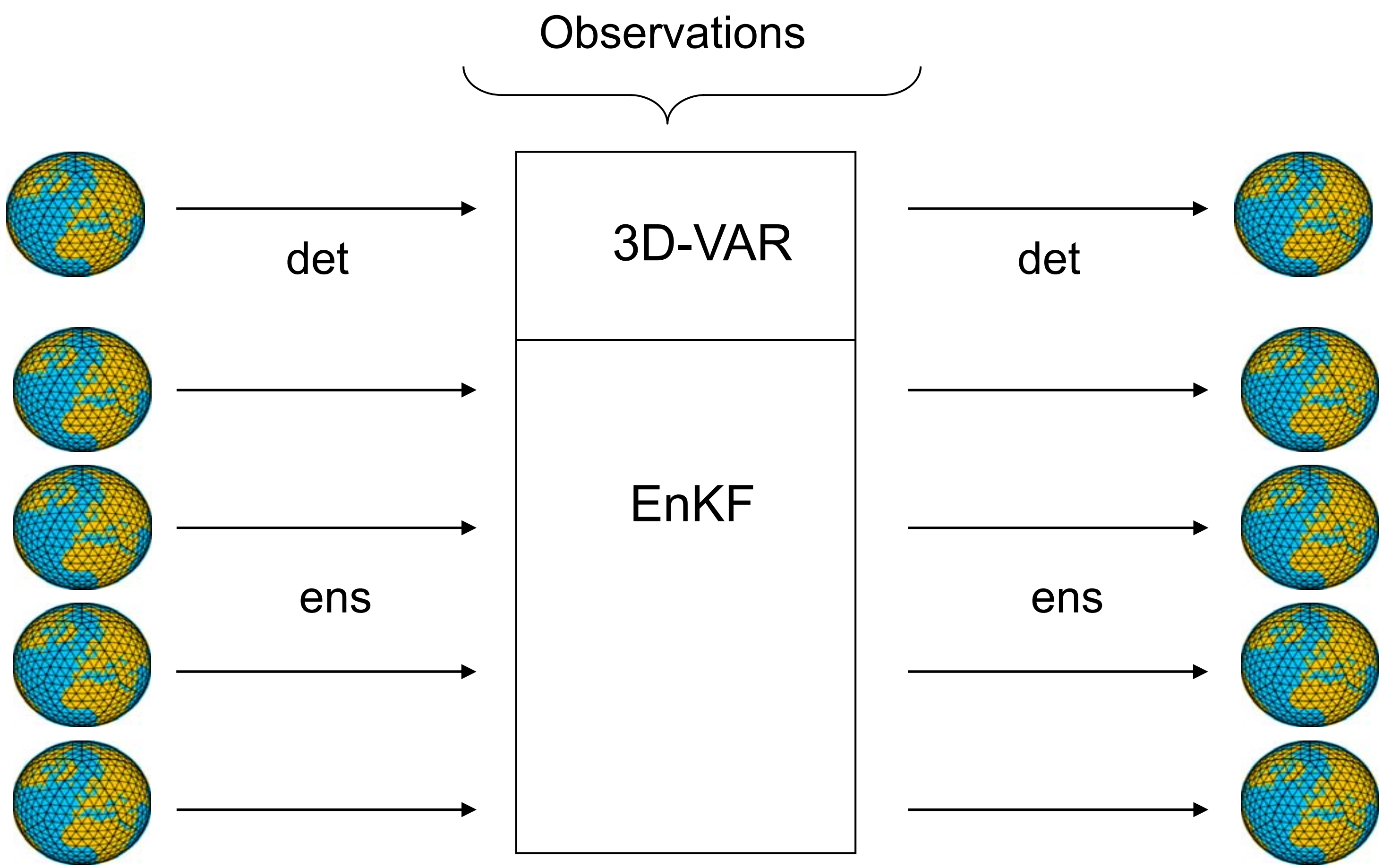
  With traditional optimizations? Also with using GPUs?

More information about HP2C will be given later this morning by Philippe Steiner (MeteoSwiss) and Angelo Mangili, Michele de Lorenzi (CSCS)

Data Assimilation

Thanks to H. Anlauf and the Assimilation-Colleagues

An Ensemble DA Framework / EnKF



3D-Var and EnKF

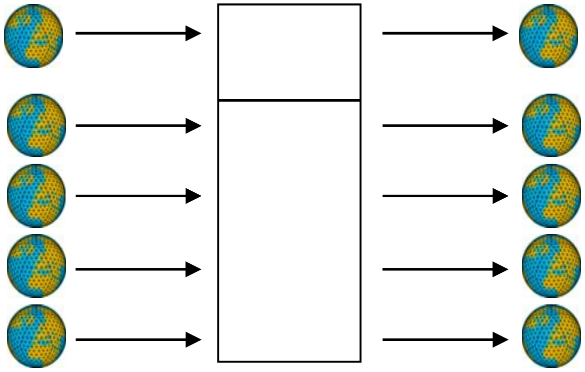
- 3D-Var was introduced to directly assimilate remote sensing data
- It is now extended to include an ensemble component with the Ensemble Kalman Filter
- The COSMO-Model will use a „pure“ version, the LETKF (Local Ensemble Transform Kalman Filter)
- The LETKF and the 3D-Var code share a common code basis
- Due to NEC compiler constraints, only a Fortran95 code can be used, although the code is rather object oriented (use of derived types)
- The developers have a strong interest to use Fortran2003/08 features (at least: allocatable components in derived types)

3D-Var and EnKF: Performance Issues

- Today, the 3D-Var has 8 minutes (at most) to produce an analysis
- On the NEC SX-9, it uses 24 processors (with a degree of vectorization of about 90 %)
- It is a pure MPI implementation

- To run an EnKF system, a lot of I/O is necessary (for the current implementation). *This has to be avoided somehow*

- To run the EPS system, 8 minutes will most probably not be enough



Computing Resources

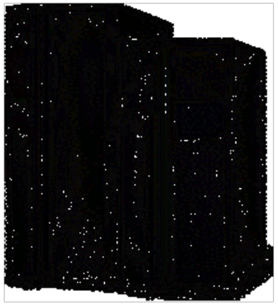
Future Computing Resources

Evolving Computer Architectures

?

Is Moore‘s law still true?
The number of transistors per square inch on integrated circuits doubles every
12 – 18 – 24 (?) months.

NEC SX-9



IBM pwr5

Cray T3E

Cray YMP

What is increasing, is the number of cores and the theoretical peak performance:

But this does not mean, that scientists can run the same code once and forever on all available computer platforms and get an ever increasing performance!

Conclusion

- ➔ The new models / applications are requiring more computer resources
- ➔ Right now DWD takes first steps to prepare the next procurement
- ➔ Up to now, the COSMO-Model is the main application of DWD
- ➔ With the introduction of ensemble systems for data assimilation and with the introduction of the new nonhydrostatic global model ICON, the computing time will be distributed more evenly between the applications
- ➔ Therefore we are about to build benchmarks also for the other models
- ➔ Today a new COSMO_RAPS_5.0 benchmark is already available
- ➔ Benchmarks for ICON and data assimilation will follow in due time
- ➔ Be aware: These benchmarks might use Fortran2003/08 concepts

Model systems are changing

Computer systems / architectures are changing

The real challenge always was and still will be to bring both together

Thank you

very much

for your

attention