TECHNISCHE UNIVERSITÄT DRESDEN

Center for Information Services and High Performance Computing (ZIH)

Highly Scalable Dynamic Load Balancing in the Atmospheric Modeling System COSMO-SPECS+FD4

PARA 2010, June 9, Reykjavík, Iceland

Matthias Lieber [a], Verena Grützun [b], Ralf Wolke [c], Matthias S. Müller [a], Wolfgang E. Nagel [a]

[a] Center for Information Services and High Performance Computing (ZIH), TU Dresden, Germany

[b] Max Planck Institute for Meteorology (MPI-M), Hamburg, Germany

[c] Leibniz Institute for Tropospheric Research (IfT), Leipzig, Germany

LEIBNIZ INSTITUTE FOR TROPOSPHERIC RESEARCH

Max-Planck-Institut für Meteorologie

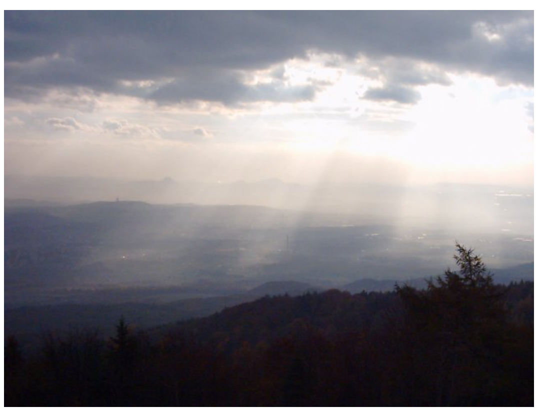
Center for Information Services & High Performance Computing

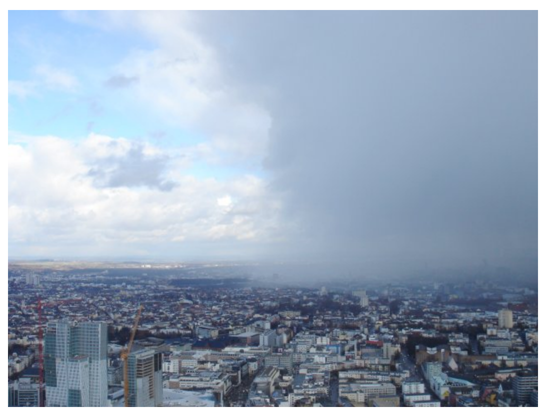
Outline

- Very short introduction to cloud modeling
- COSMO-SPECS
- Framework FD4
- COSMO-SPECS+FD4 benchmarks
- Conclusion & outlook

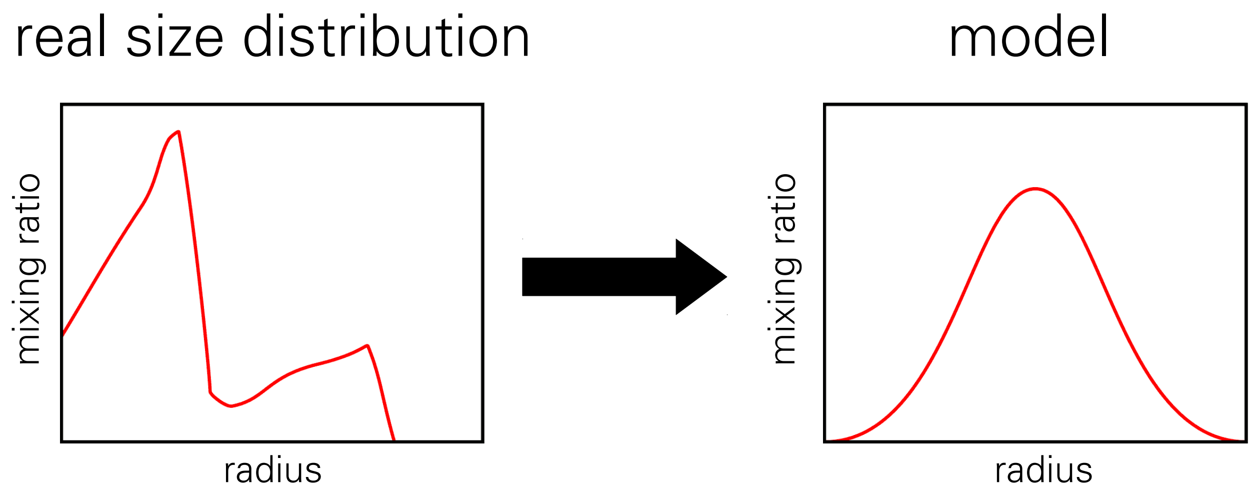
Introduction: Cloud Modeling

- Clouds play a major role for climate and weather: They
  - influence the radiation budget of the planet,
  - are part of the hydrological cycle,
  - interact with aerosol particles and pollution
- Clouds represent one of the major uncertainties in climate and weather models [IPCC07, CCSP09]
- Unsatisfying improvements in precipitation forecast during the last decades





Introduction: Bulk Parameterization Schemes

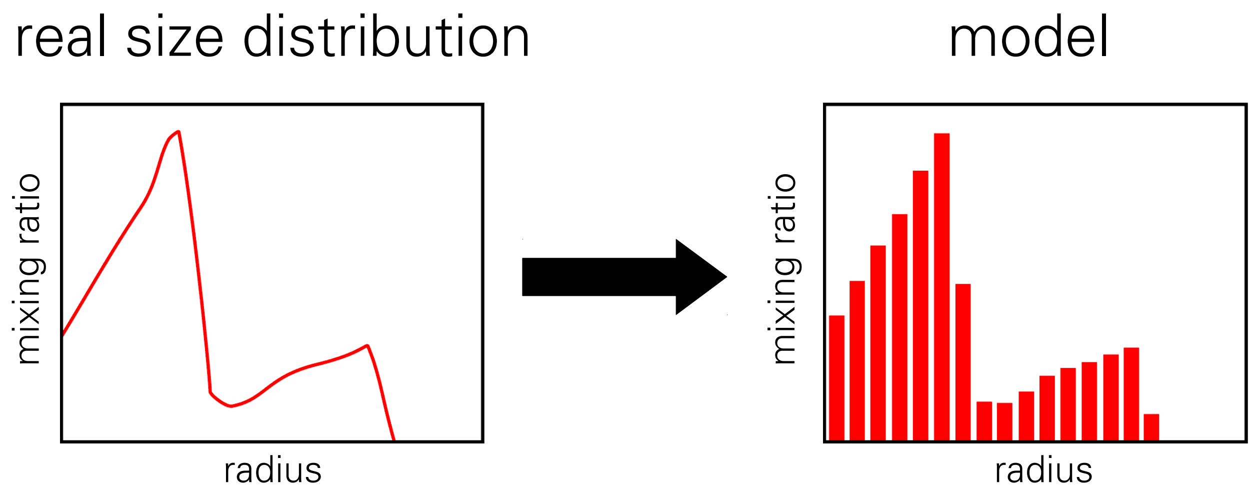


- Cloud droplets are described by their bulk mass only
- Actual size distribution is neglected
- Computationally fast
- Used in most weather models

TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

Introduction: Spectral Bin Microphysics Schemes



- Bin discretization of size distribution
- Allows more detailed modeling of interaction between aerosol particles, clouds, and precipitation
- Computationally very expensive (runtime & memory)
- Only used for process studies up to now (MM5, WRF, COSMO)

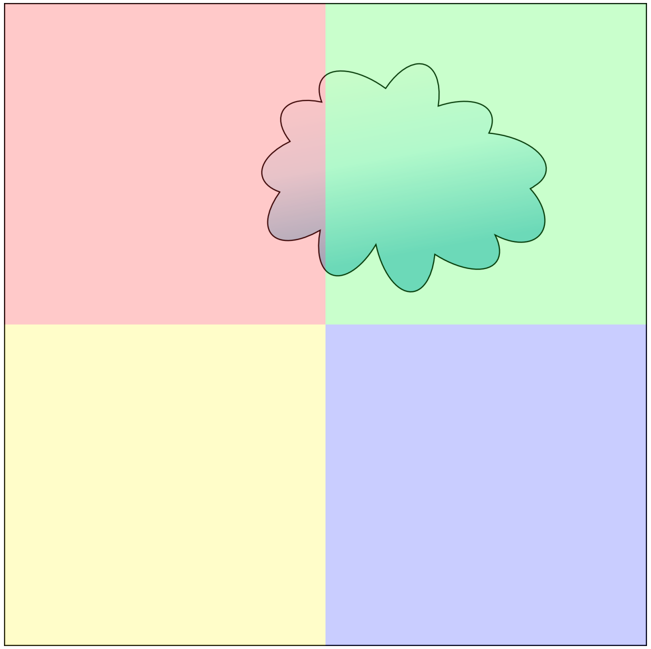
COSMO-SPECS: COSMO with SPECtral bin microphysicS

- COSMO Model: non-hydrostatic limited-area atmospheric model, formerly known as "Lokal-Modell" of DWD
  - http://www.cosmo-model.org

- COSMO-SPECS: Cloud parameterization scheme of COSMO replaced by the spectral bin microphysics model SPECS [Simmel06, Grützun08]
  - 11 new variables to describe 3 types of hydrometeors (water droplets, frozen particles, insoluble particles)
  - Discretized into a predefined number of size bins, e.g. 66 bins x 11 variables = 726 values!

COSMO-SPECS: Performance

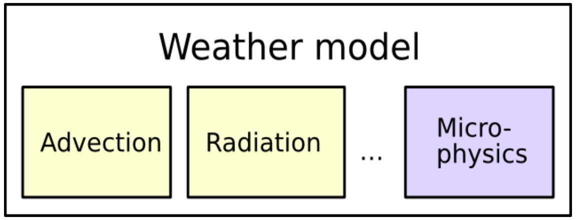
- SPECS is very costly
  - > 90% of total runtime
- SPECS runtime varies strongly
  - Depending on range of droplet size distribution and the presence of frozen particles
- SPECS induces severe load imbalance
  - COSMO's parallelization is based on static 2D partitioning



Dynamic load balancing needed to run realistic cases on large HPC systems

Framework FD4: Basic Idea

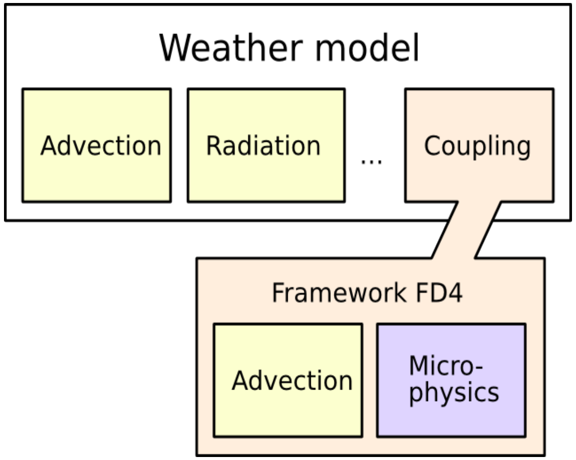
- Present approaches:
  - Cloud model is implemented as a sub-module within the weather model
  - Uses (static) data structures of the weather model



- Our idea:

Functionality provided by FD4

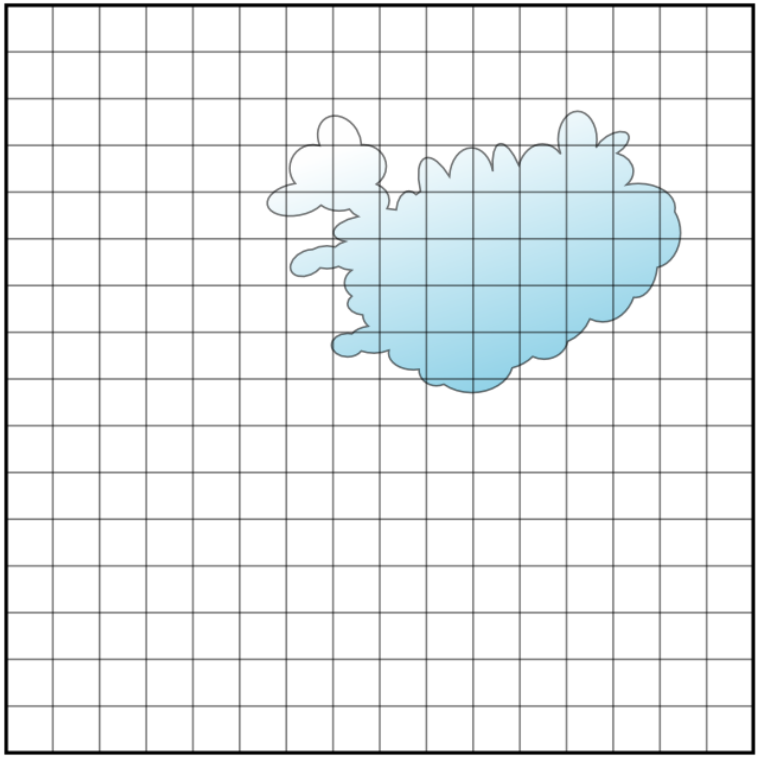
  - Separate cloud model data from weather model data structures
  - Independent domain decompositions
  - Dynamic load balancing for the cloud model
  - (Re)couple weather and cloud model



Framework FD4: Features

FD^4 = Four-Dimensional Distributed Dynamic Data structures

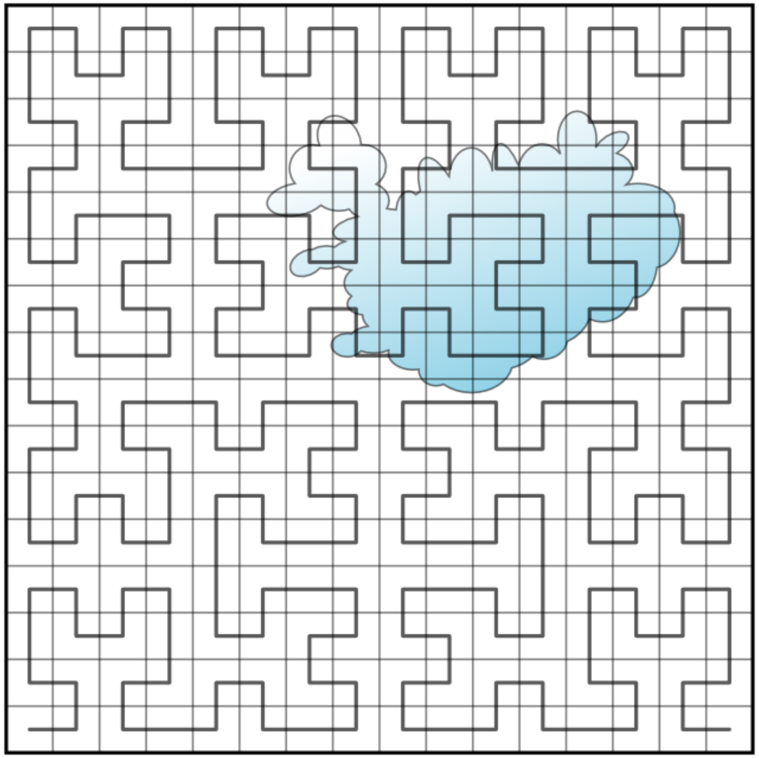
- Rectangular grid
- Block-based 3D decomposition



Framework FD4: Features

FD^4 = Four-Dimensional Distributed Dynamic Data structures

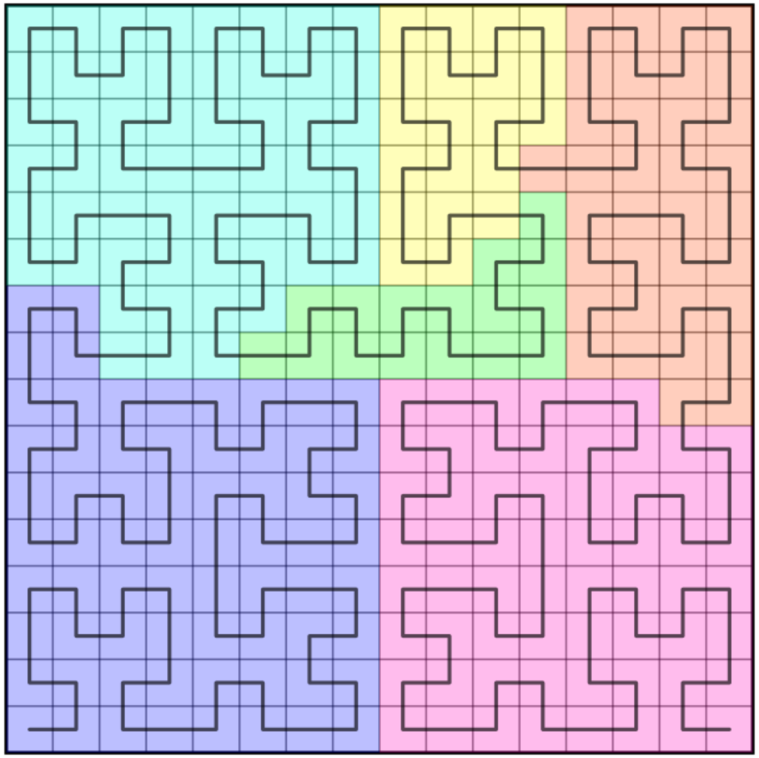
- Rectangular grid
- Block-based 3D decomposition
- Dynamic load balancing
  - Hilbert space-filling curve
  - ParMETIS



Framework FD4: Features

FD^4 = Four-Dimensional Distributed Dynamic Data structures

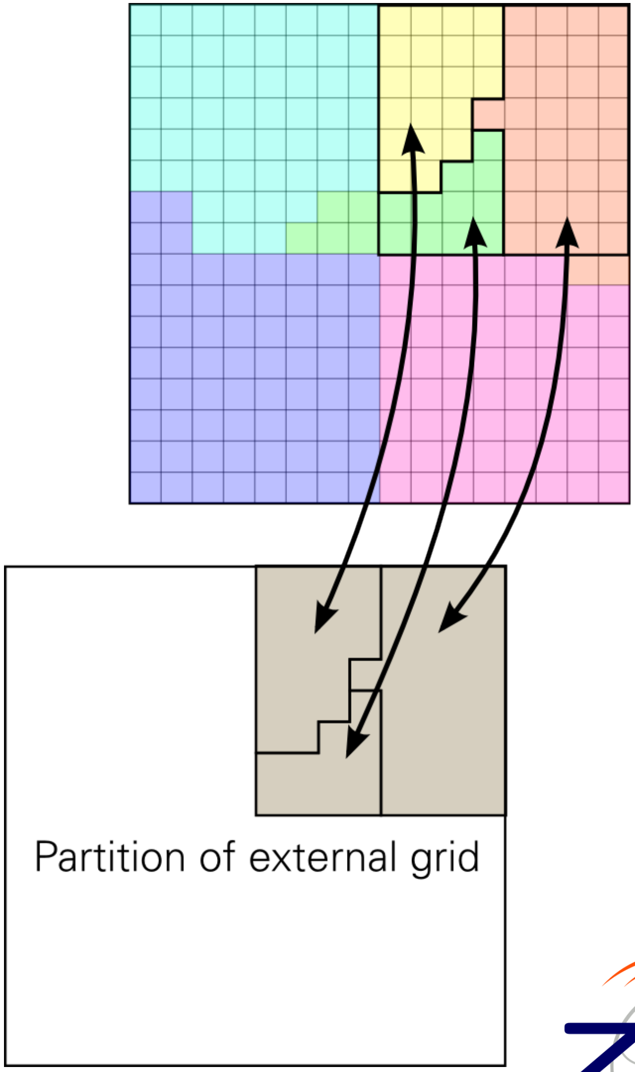
- Rectangular grid
- Block-based 3D decomposition
- Dynamic load balancing
  - Hilbert space-filling curve
  - ParMETIS



Framework FD4: Features

FD^4 = Four-Dimensional Distributed Dynamic Data structures

- Rectangular grid
- Block-based 3D decomposition
- Dynamic load balancing
  - Hilbert space-filling curve
  - ParMETIS
- Coupling to external model



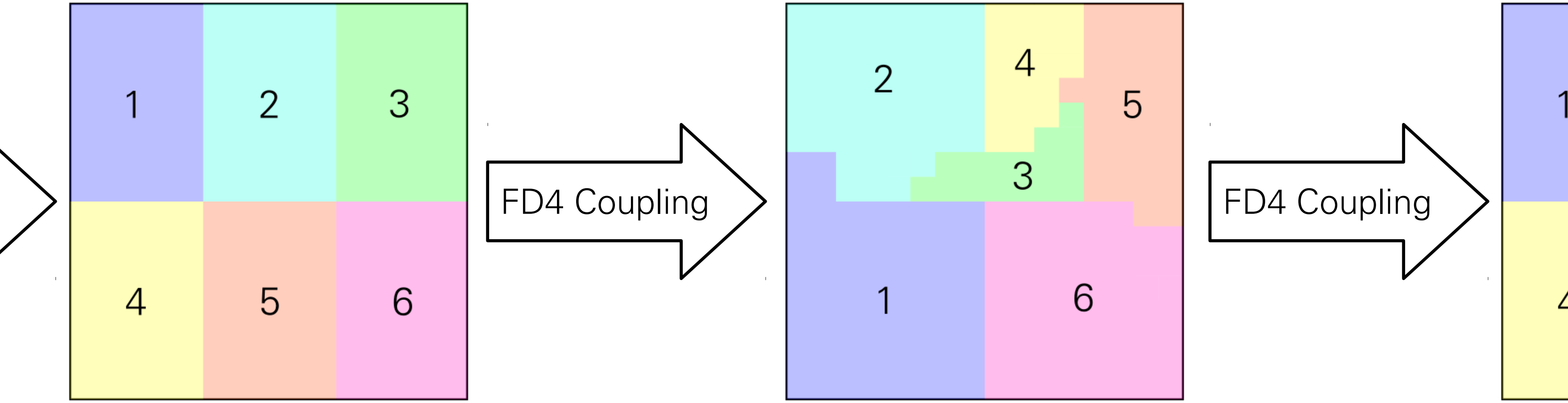
Framework FD4: Features

FD⁴ = Four-Dimensional Distributed Dynamic Data structures

- Rectangular grid
- Block-based 3D decomposition
- Dynamic load balancing
  - Hilbert space-filling curve
  - ParMETIS
- Coupling to external model
- Fortran 90, MPI

```
! MPI initialization
call MPI_Init(err)
call MPI_Comm_rank(MPI_COMM_WORLD, rank, err)
call MPI_Comm_size(MPI_COMM_WORLD, nproc, err
! create the domain and allocate memory
call fd4_domain_create(domain, nb, size, vart
call fd4_util_allocate_all_blocks(domain, err
call fd4_balance_readjust(domain, err)
! initialize ghost communication
call fd4_ghostcomm_create(ghostcomm, domain,
! loop over time steps
do timestep=1,nsteps
  ! exchange ghosts
  call fd4_ghostcomm_exch(ghostcomm, err)
  ! loop over local blocks
  call fd4_iter_init(domain, iter)
  do while(associated(iter%cur))
    ! do some computations
    call compute_block(iter)
    call fd4_iter_next(iter)
  end do
  ! dynamic load balancing
  call fd4_balance_readjust(domain, err)
end do
```

COSMO-SPECS+FD4 Coupling



COSMO

Computes dynamics

Static MxN partitioning

FD4

Send data to SPECS grid:

u, v, w, T, p, ρ, q_v

SPECS

Computes Microphysics

Data dynamically balanced by FD4

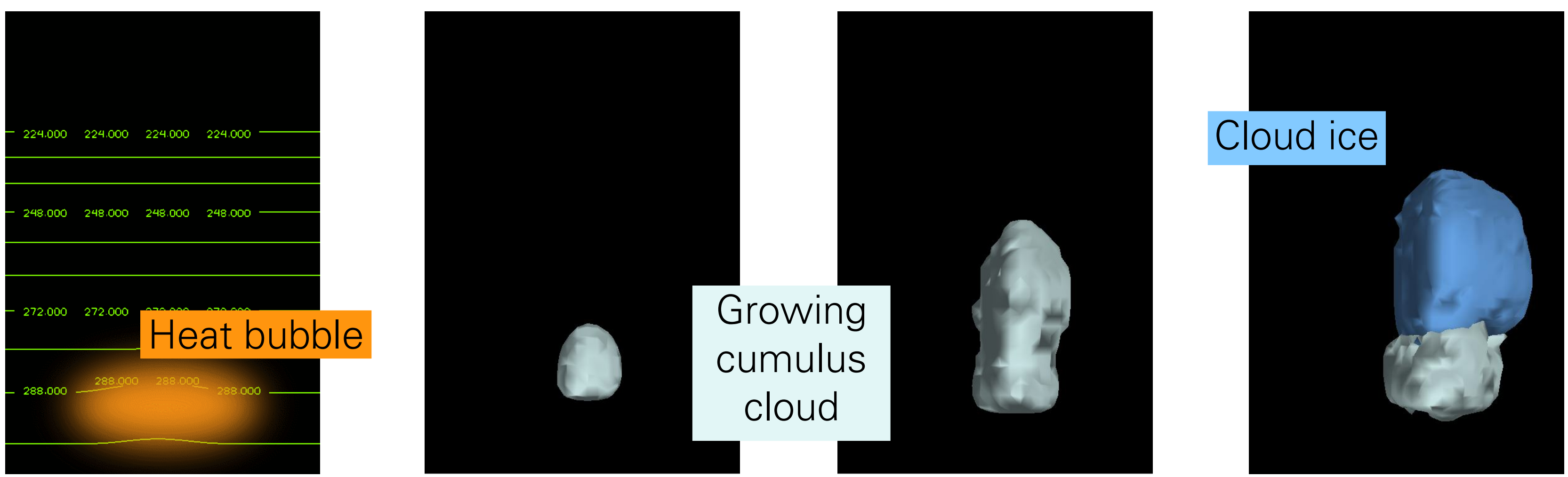
FD4

Receive data from SPECS grid:

ΔT, q_v, q_c, q_i

Strong Scaling Benchmark

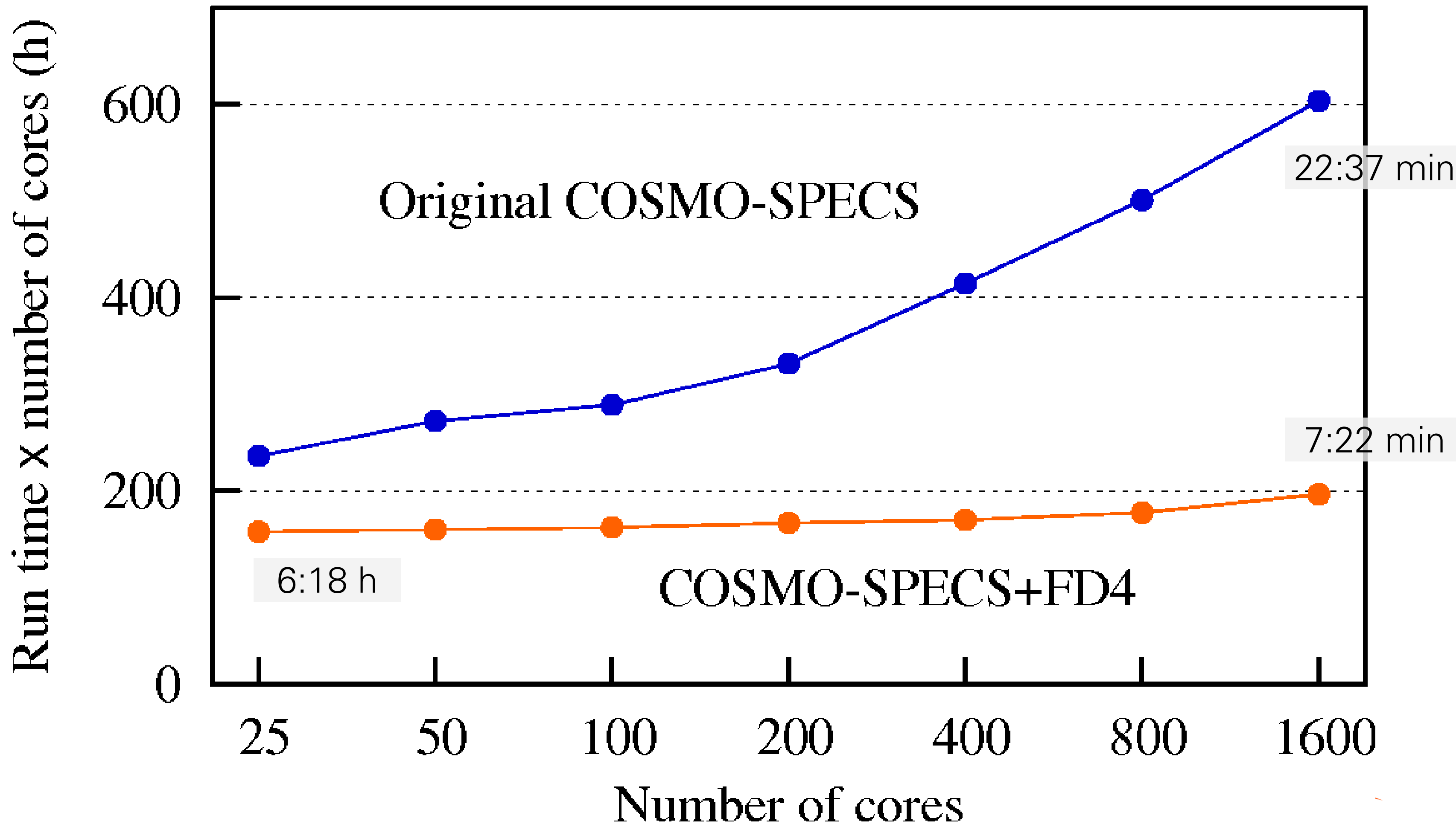
- Comparing original COSMO-SPECS with COSMO-SPECS+FD4
- Test scenario: heat bubble results in growth of cumulus cloud
- Horizontal grid: 80 x 80, 1km resolution
- Vertical grid: 48 nonuniform height levels (up to 18 km)
- 30 min forecast time
- 19 200 FD4 blocks (2 x 2 x 4 grid cells)



t = 0 min | t = 10 min | t = 20 min | t = 30 min

Strong Scaling Benchmark Results - SGI Altix 4700

Run time x #cores – flat horizontal lines indicate perfect scaling



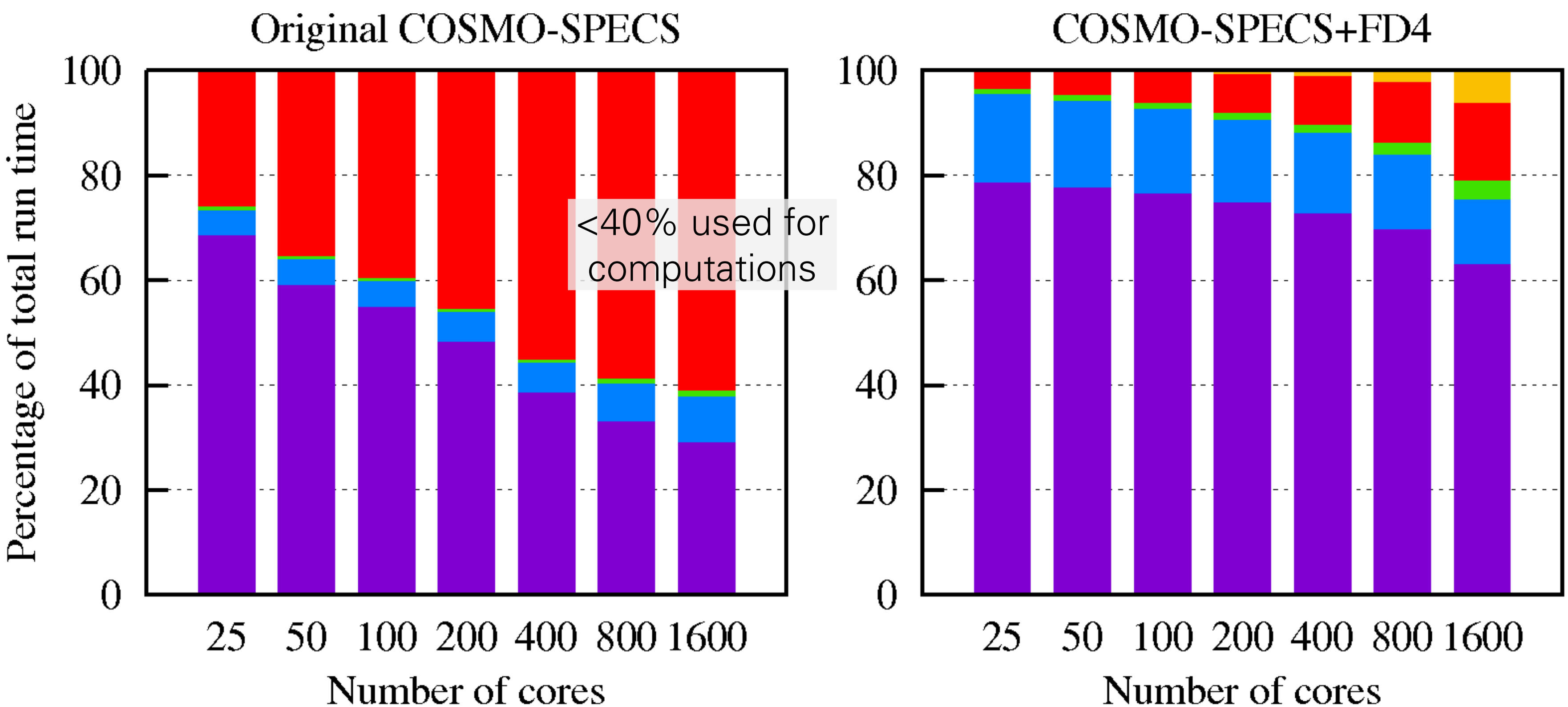
TECHNISCHE UNIVERSITÄT DRESDEN

Thanks to LRZ Munich for access to HLRB-II

ZIH Center for Information Services & High Performance Computing

Strong Scaling Benchmark Results - SGI Altix 4700

Component breakdown

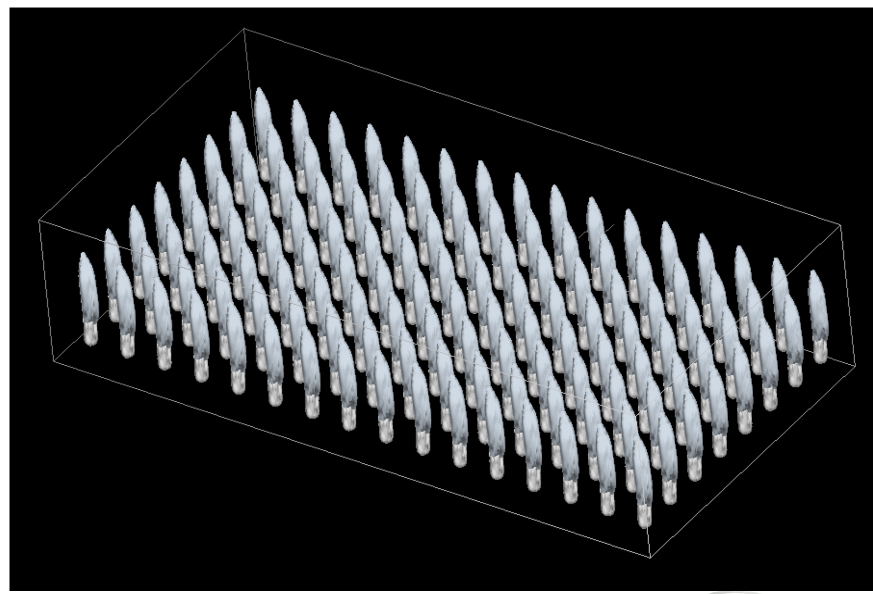


- FD4 load balancing and coupling
- Ghost exchange of hydrometeors including waiting times
- COSMO computations
- Advection of hydrometeors
- SPECS computations

Weak Scaling Benchmark

- *Replication scaling* benchmark method [Straalen09]
  - Create identical copies of same physical problem (i.e. cloud) when scaling up the grid size
- Each 32 x 32 horizontal grid tile initialized with a heat bubble
- Challenge: balance 393 216 blocks over 32k cores

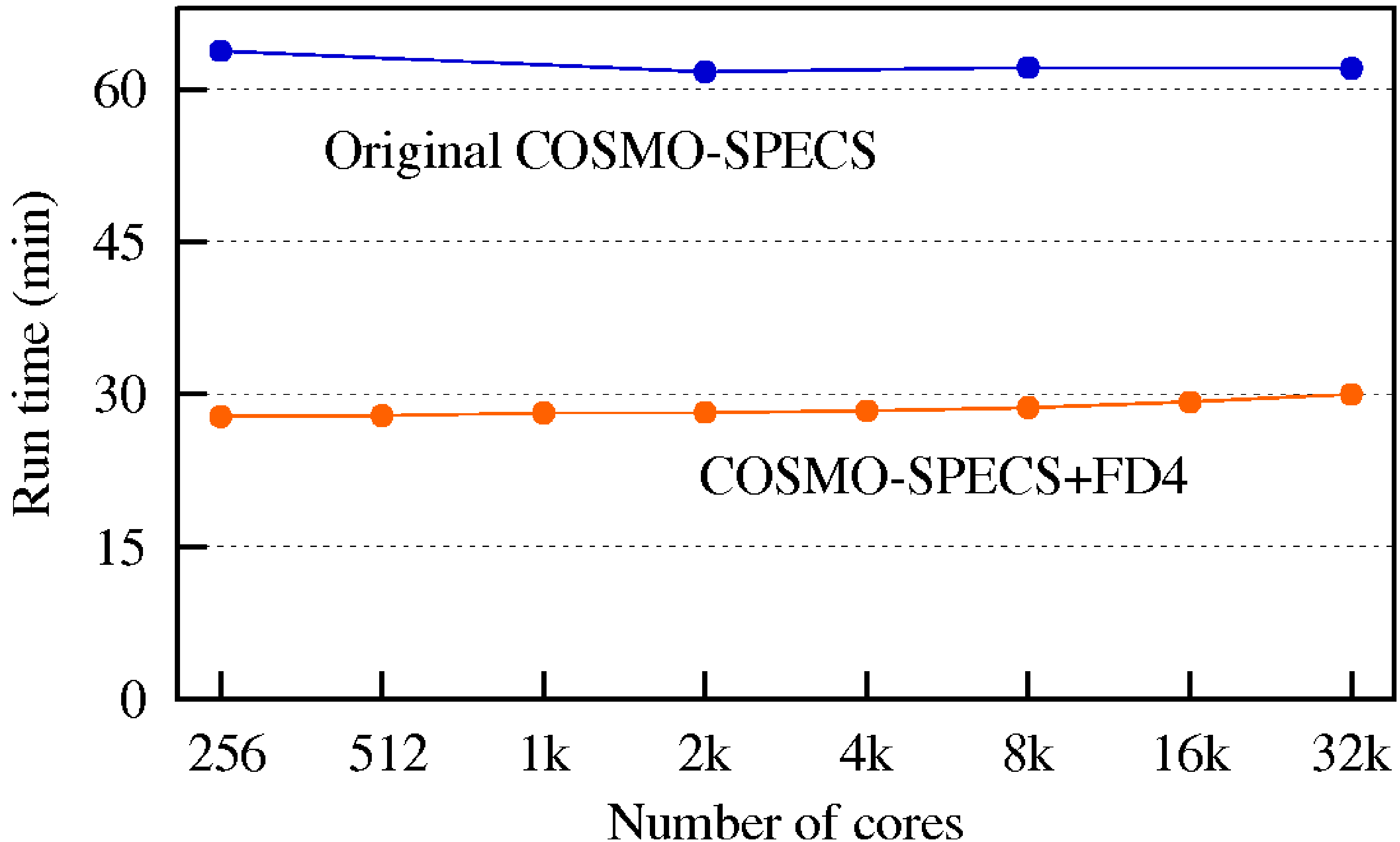
# Cores	Grid size	# Replicated clouds	# FD4 blocks
256	32x32	1x1	3072
512	64x32	2x1	6144
1024	64x64	2x2	12 288
...			
16 384	256x256	8x8	196 608
32 768	512x256	16x8	393 216



16x8 clouds after 30 min simulation

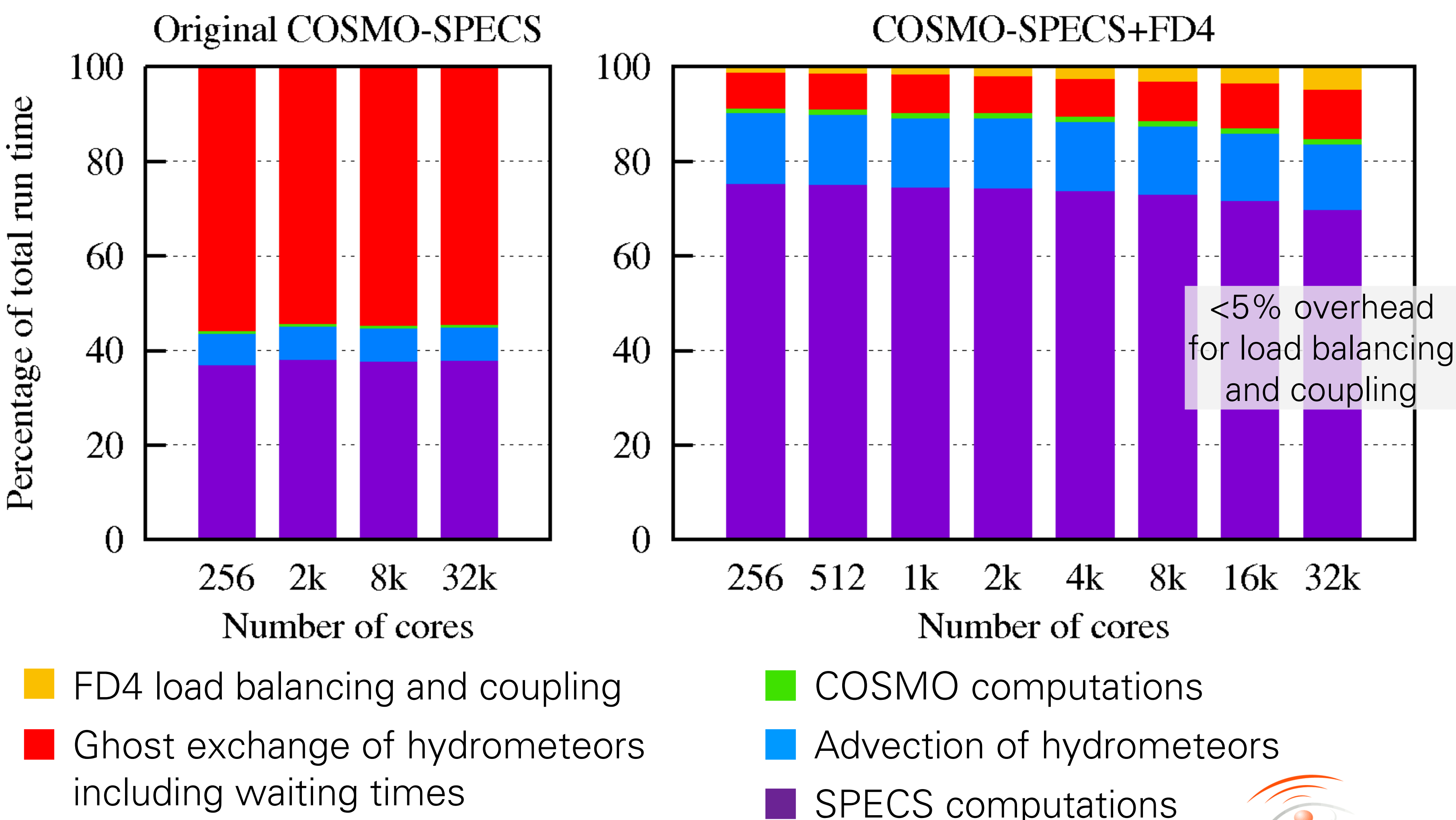
Weak Scaling Benchmark Results - BlueGene/P

Run time – flat horizontal lines indicate perfect scaling



Weak Scaling Benchmark Results - BlueGene/P

Component breakdown



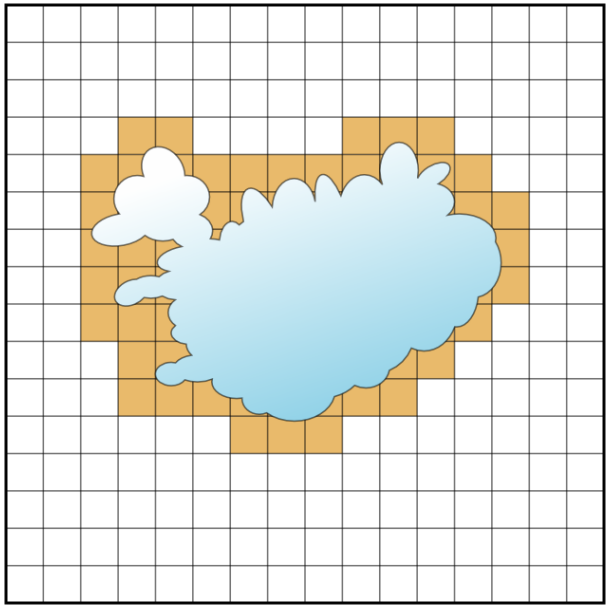
Thanks to JSC for access to JUGENE

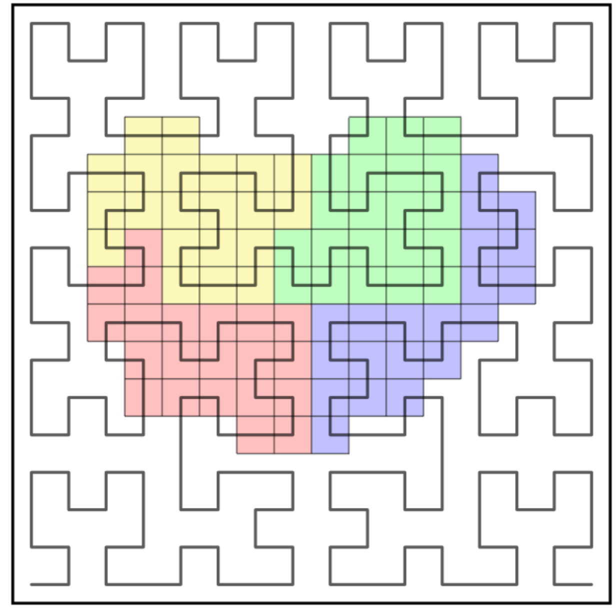
Conclusion & Outlook

- New way of coupling detailed cloud microphysics to atmospheric models
- FD4 provides highly scalable dynamic load balancing and coupling
- COSMO-SPECS performance increased significantly by FD4
- FD4 not limited to meteorology

- Next steps:
  - Adaptive time stepping
  - Parallel I/O (based on NetCDF4)
  - Avoid global metadata operations: local / hierarchical methods for load balancing and coupling

Outlook: Adaptive Block Mode in FD4





- Only those blocks are allocated which are required to capture the clouds to save memory [Lieber10]
- FD4 ensures appropriate data for stencil operations
- COSMO-SPECS+FD4: run expensive spectral bin microphysics in cloudy grid cells only

Thank you for your attention!

[CCSP09] *Atmospheric Aerosol Properties and Climate Impacts*, U.S. Climate Change Science Program and the Subcommittee on Global Change Research, 2009.

[Grützun08] V. Grützun, O. Knoth, and M. Simmel. *Simulation of the influence of aerosol particle characteristics on clouds and precipitation with LM–SPECS: Model description and first results.* Atmos. Res., 90:233–242, 2008.

[IPCC07] S. Solomon et al. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Forth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC)*, 2007.

[Lieber10] M. Lieber, R. Wolke, V. Grützun, M.S. Müller, and W.E. Nagel, *A framework for detailed multiphase cloud modeling on HPC systems*, in Parallel Computing Vol. 19, pages. 281-288, IOS Press, 2010.

[Pinar04] A. Pinar and C. Aykanat. *Fast optimal load balancing algorithms for 1D partitioning.* J. Parallel Distrib. Comput., 64(8):974-996, 2004.

[Simmel06] M. Simmel and S. Wurzler, *Condensation and activation in sectional cloud microphysical models*, Atmos. Res., 80:218-236, 2006.

[Straalen09] B. Van Straalen, J. Shalf, T. Ligocki, N. Keen, and W.-S. Yang. *Scalability challenges for massively parallel AMR applications,* in IPDPS 09, pages 1-12, IEEE, 2009.

[Teresco06] J.D. Teresco, K.D. Devine, and J.E. Flaherty. *Partitioning and Dynamic Load Balancing for the Numerical Solution of Partial Differential Equations,* in Numerical Solution of Partial Differential Equations on Parallel Computers, pages 55-88, Springer, 2006.

This work has been funded by the German Research Foundation