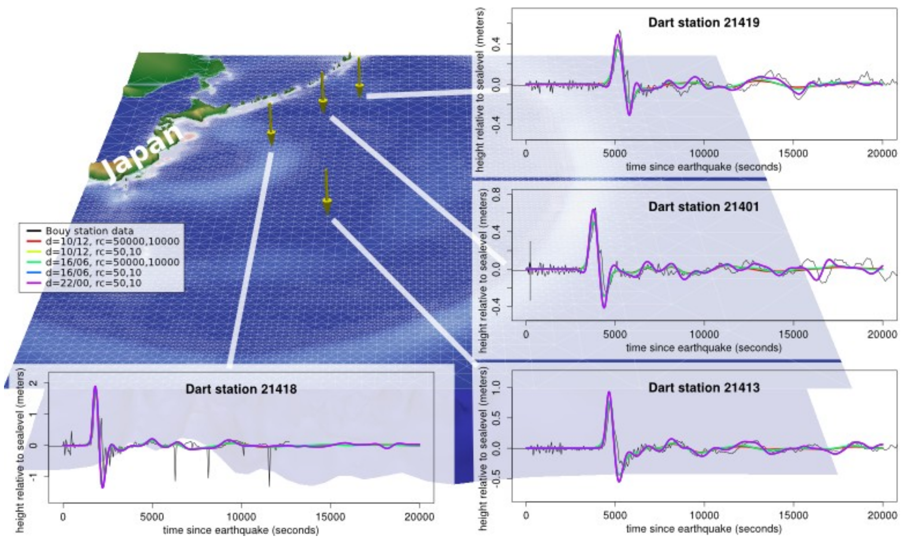
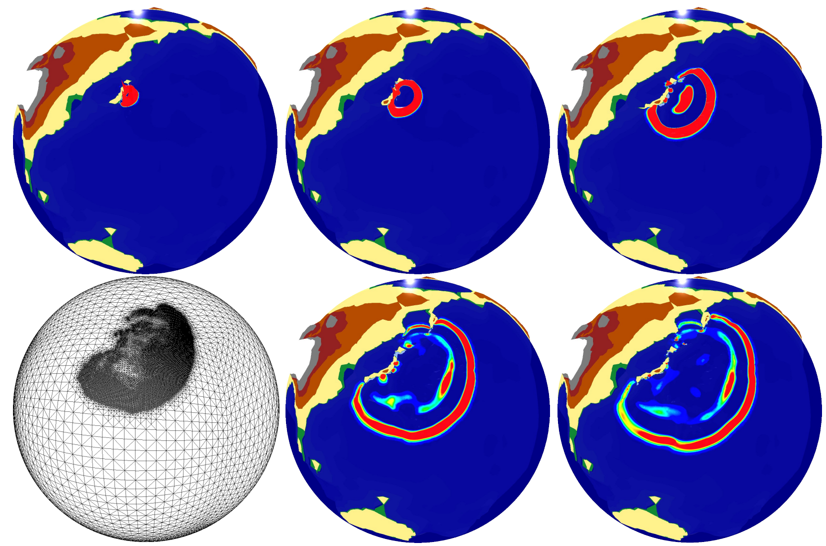
Cluster-based parallelization of simulations on dynamically adaptive grids on the sphere

Martin Schreiber, Hans-Joachim Bungartz





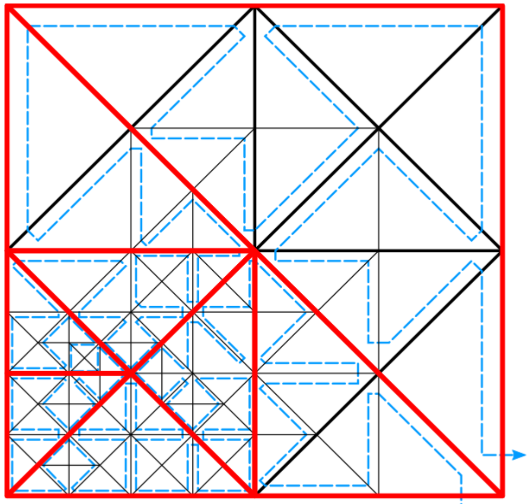
Left: Simulation of the Tohoku Tsunami, right: earth-scale simulation
(developed in collaboration with Alexander Breuer, based on Augmented Riemann solvers [1])

[1] D. L. George: Augmented Riemann solvers for the shallow water equations over variable topography with steady states and inundation, 2008

Cluster-based parallelization

Clustering:

- **Multiple partitions** in a program's context
- **Replicated** shared hyperfaces
- Our clusters are generated by **subtrees** of the bisective Sierpinski **SFC-induced space tree** [2]



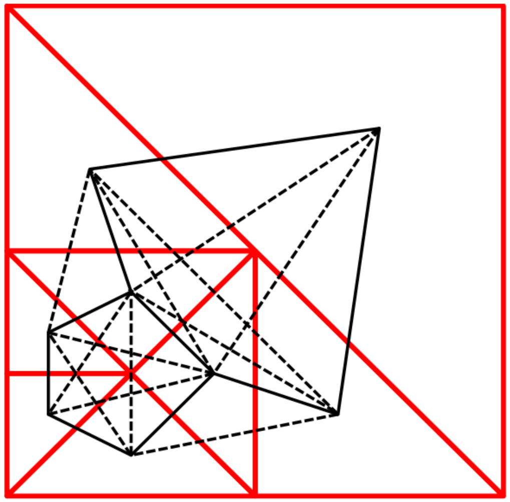
Partitioning with subtrees

Intra-cluster communication:

- Based on **location-independent** stack communication system [3] (see e.g. Peano framework)

Inter-cluster communication:

- Use SFC supporting **RLE connectivity information**
- **Implicitly update** RLE connectivity information



Node communication graph

[2] M. Bader and C. Zenger: Efficient Storage and Processing of Adaptive Triangular Grids Using Sierpinski Curves, 2006
[3] M. Mehl, T. Weinzierl and C. Zenger: A cache-oblivious self-adaptive full multigrid method, 2006

Clustering

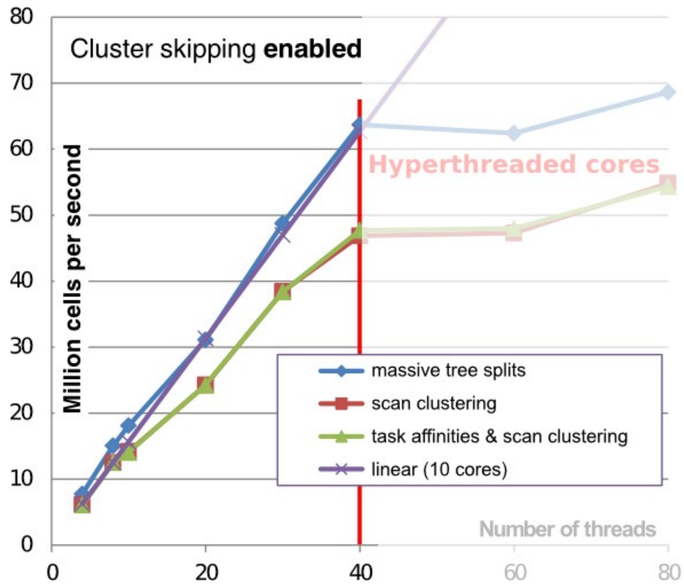
(efficient) handling of multiple partitions in a single program context

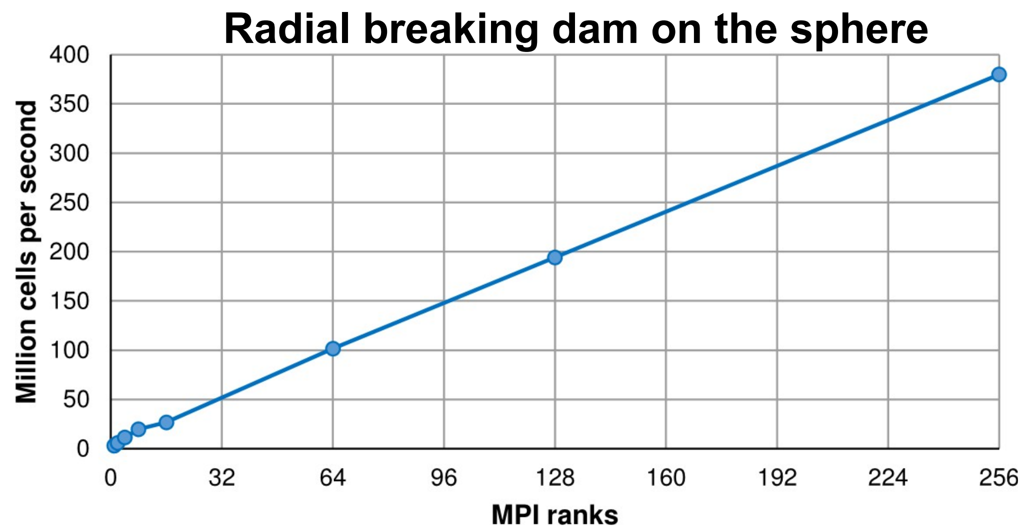
[6] **Run-length encoded (RLE) connectivity info.** for edges and nodes

OpenMP/TBB + MPI parallelization on cluster level [7]

Cluster-based **data migration** for load balancing [9]

Cluster-based **optimizations** [8]





[6] M. Schreiber, H.-J. Bungartz, M. Bader: Shared Memory Parallelization of Fully-Adapt. Sim. Using a Dynamic Tree-Split and -Join Approach, HiPC 2012
[7] M. Schreiber, T. Weinzierl and H.-J. Bungartz: SFC-based Communication Metadata Encoding for Adaptive Mesh, ParCo 2013
[8] M. Schreiber, T. Weinzierl and H.-J. Bungartz: Cluster Optimization and Parallelization of Simulations with Dynamically Adaptive Grids, Euro-Par 2013
[9] M. Schreiber, H.-J. Bungartz: Cluster-based communication and load balancing for simulations on dynamically adaptive grids, ICCS 2014, accepted

Link to video:

https://www.youtube.com/watch?v=WllGOALQerI

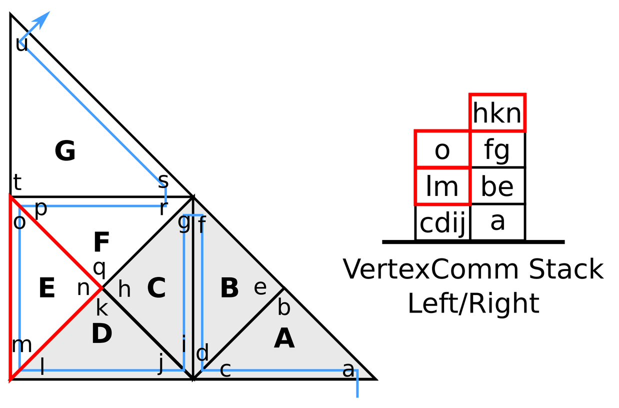
Cluster-based parallelization of simulations on dynamically adaptive grids on the sphere

Martin Schreiber
schreibm@in.tum.de

Hans-Joachim Bungartz
bungartz@in.tum.de

Sierpiński Framework - http://www5.in.tum.de/sierpinski

SFC-based domain decomposition and stack-based communication [1,2,3]



Clustering
(efficient) handling of multiple partitions in a single program context

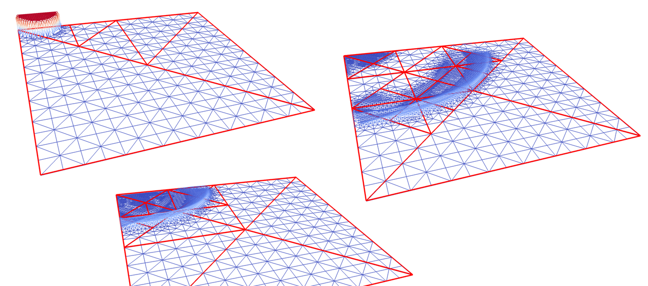
Run-length encoded (RLE) connectivity info. for edges and nodes [6]

OpenMP/TBB + MPI parallelization on cluster level [7]

Cluster-based **data migration** for load balancing [9]

Cluster-based **optimizations** [8]

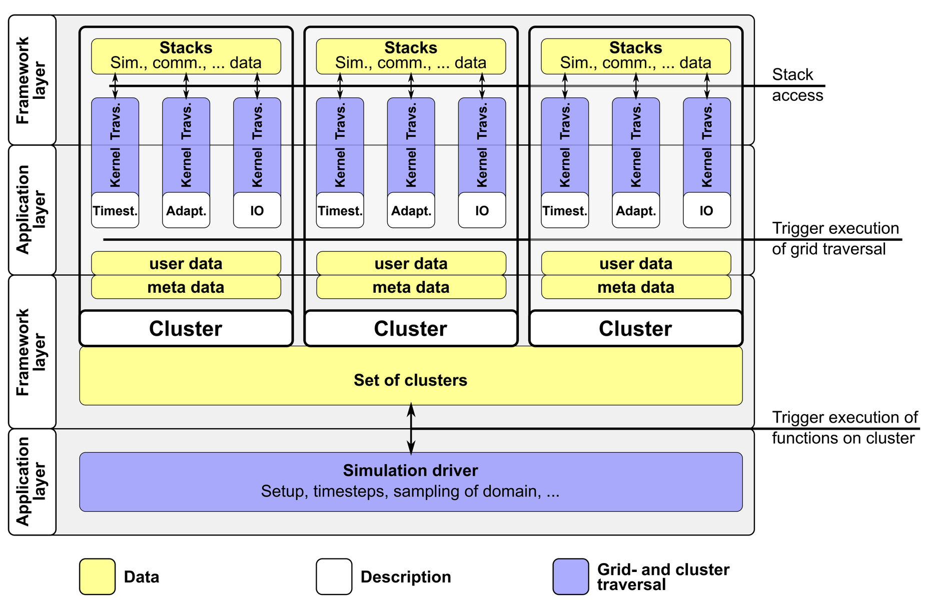
2D discontinuous Galerkin with triangles



Shallow water, Tsunami, Euler, multi layer...

Framework design

Modularized traversals/kernels:
Setup, interactive changes, simulation timestep, adaptivity, visualization, sampling (e.g. for buoy station data)

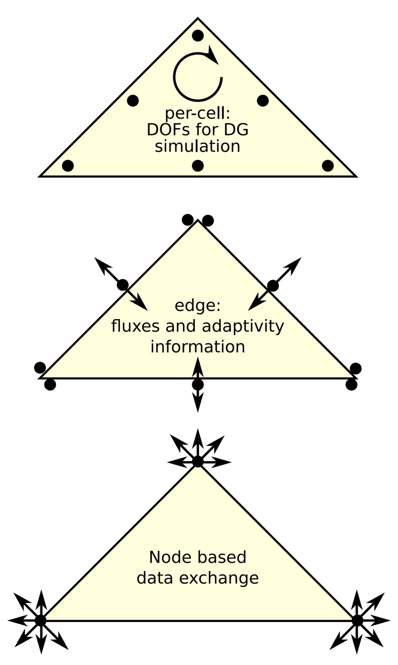


Data access patterns:

Cell data
(DG/FV simulation data, ...)

Edge data
(Flux computations, limiter, adaptivity markers, ...)

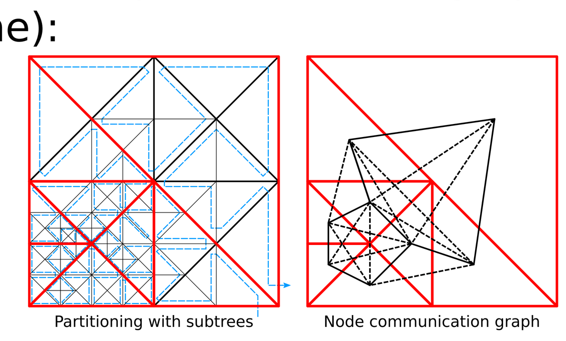
Node data:
(Visualization, limiter, ...)



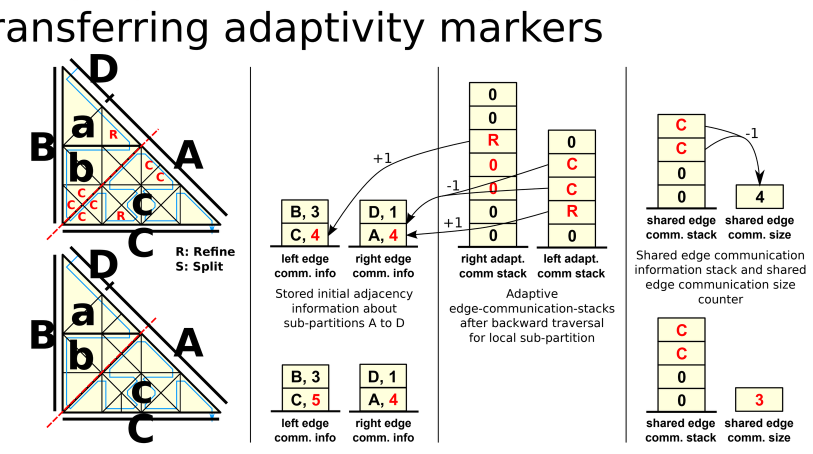
Modularized simulation framework:
Possibility of *multiple simulations* in a single program context, e.g. for coupling of oceanic and atmospheric simulations

Parallelization

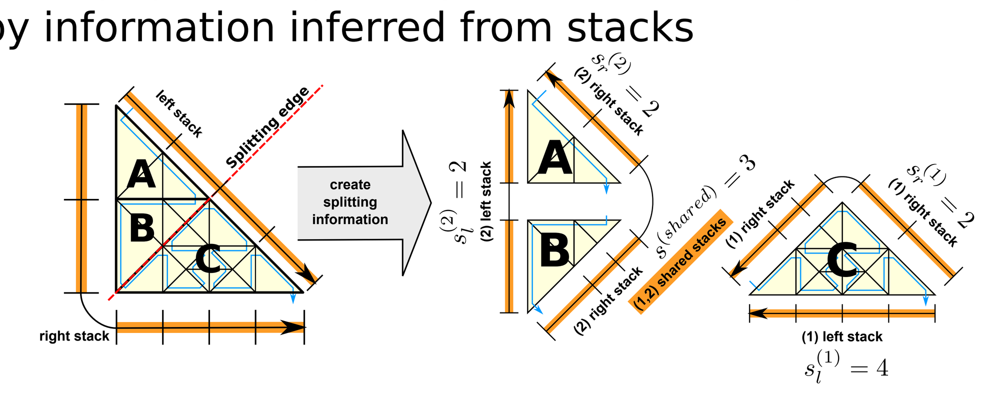
Connectivity information (Sparse connect. graph)
- edges/nodes (solid line):
  RLE entry (>0) [6]
- nodes (dashed lines):
  RLE entry (=0) [7]



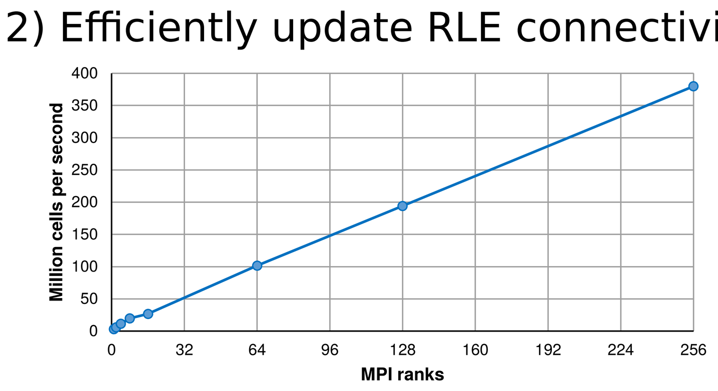
Dynamically changing grids [6]:
Implicitly update connectivity information by transferring adaptivity markers



Cluster generation [6]:
During split and join, update connectivity data by information inferred from stacks



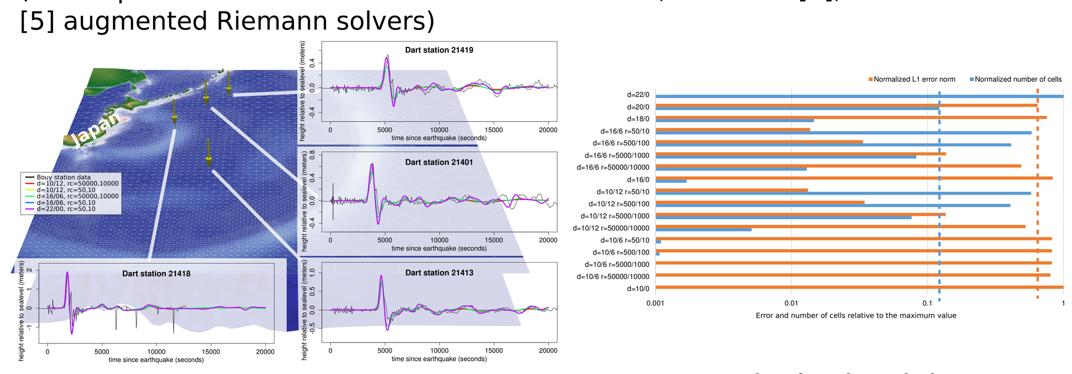
Cluster-based data migration [9]:
1) En bloc transfer of "raw" cluster data
2) Efficiently update RLE connectivity information



Example:
Simulation of radial dam break on the sphere

Applications

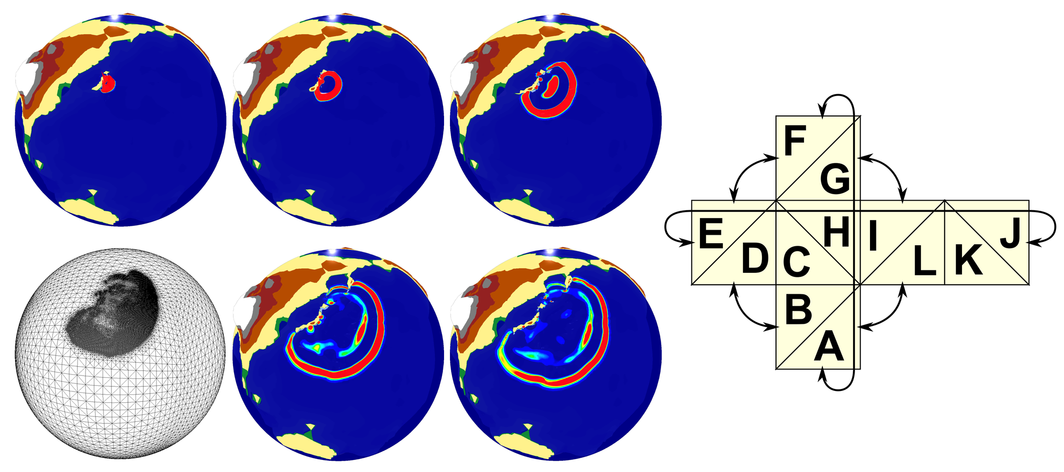
Tsunami simulation: Tohoku
(Developed in collaboration with *Alexander Breuer*, based on [4], [5] augmented Riemann solvers)



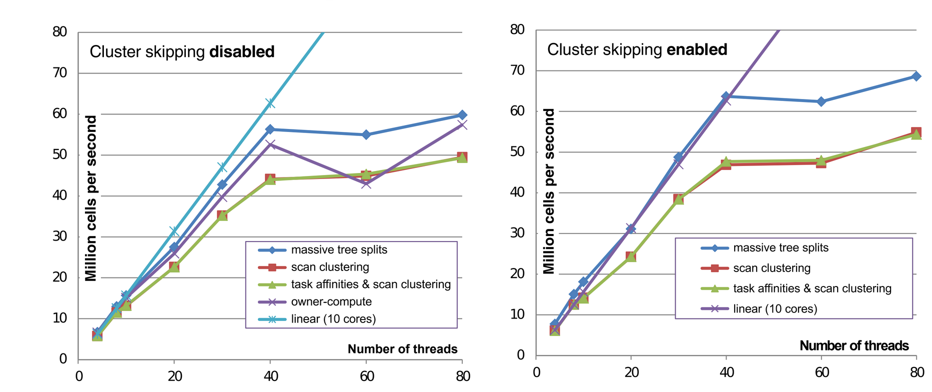
Performance improvement of 6.6 (69070 cells in average per time step with reduced error)

Dynamical adaptivity saves more than 95% of the cells

Tohoku Tsunami on hollow sphere:



Cluster-based optimizations:



Clustering allows **skipping and reordering** of the execution of operations:
- **skipping** of conformity adaptivity traversals (see above) [8]
- **local residual** corrections (skip smoother)
- compensate load imbalances with **unpredictable flux computations**, e.g. [4]
- ...

[1] M. Mehl, T. Weinzierl and C. Zenger: A cache-oblivious self-adaptive full multigrid method, 2006
[2] M. Bader and C. Zenger: Efficient Storage and Processing of Adaptive Triangular Grids Using Sierpinski Curves, 2006
[3] M. Bader, K. Rahnema and C. A. Vigh: Memory-Efficient Sierpinski-Order Traversals on Dynamically Adaptive, Recursively Structured Triangular Grids
[4] D. L. George: Augmented Riemann solvers for the shallow water equations over variable topography with steady states and inundation, 2008
[5] IHO IOC. BODC: Centenary Edition of the GEBCO Digital Atlas. British oceanographic data centre, Liverpool

[6] M. Schreiber, H.-J. Bungartz, M. Bader: Shared Memory Parallelization of Fully-Adapt. Sim. Using a Dynamic Tree-Split and -Join Approach, HiPC 2012
[7] M. Schreiber, T. Weinzierl and H.-J. Bungartz: SFC-based Communication Metadata Encoding for Adaptive Mesh, ParCo 2013
[8] M. Schreiber, T. Weinzierl and H.-J. Bungartz: Cluster Optimization and Parallelization of Simulations with Dynamically Adaptive Grids, Euro-Par 2013
[9] M. Schreiber, H.-J. Bungartz: Cluster-based communication and load balancing for simulations on dynamically adaptive grids, ICCS 2014, accepted

Outlook

Multi-layer simulations on the sphere with **dynamically changing computing resources**

This is a DFG-funded project **Invasive Computing** http://invasic.de/