FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

Scaling Finite Element Multigrid Solvers to Ten Trillion (10^{13}) Unknowns

Dominik Bartuschat

Chair for System Simulation, University of Erlangen-Nürnberg

March 3, 2017

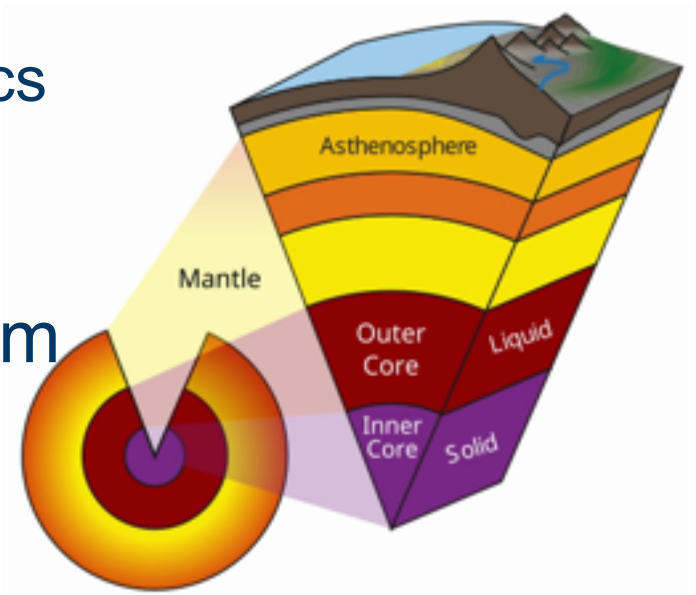
SIAM CSE 2017
Atlanta, USA

joint work with B. Gmeiner, D. Thoennes, U. Rüde (FAU)
D. Drzisga, M. Huber, L. John, C. Waluga, B. Wohlmuth (TUM)
S. Bauer, M. Mohr, P. Bunge (LMU)

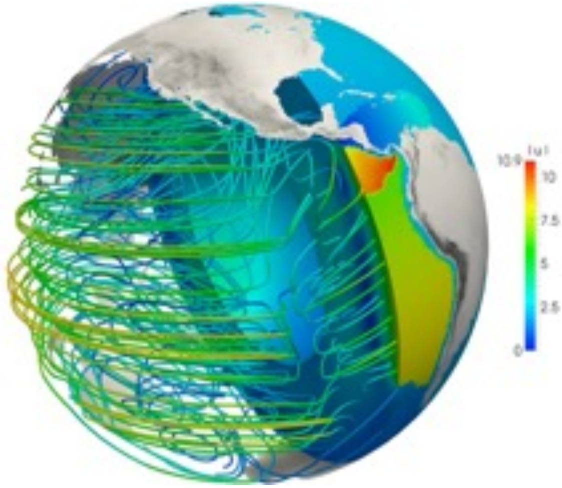
Solver for Earth Mantle Convection

- Why simulations?
  - Driving force for plate tectonics
  - Cause of earthquakes and mountain formation
- Coupled multiphysics problem
- Problem dimensions
  - 10^9 years time scale
  - 6900 km Earth radius
  - 3000 km Earth mantle thickness

➡ 10^{12} degrees of freedom (1 km resolution)



Simplified model



$$-\nabla \cdot (2\,\mu\,(\vartheta, \mathbf{x})\,\dot{\varepsilon}\,(\mathbf{u})) + \nabla p = \rho\,(\vartheta, \mathbf{x})\;\mathbf{g}$$

$$\nabla \cdot \mathbf{u} = 0$$

$$\partial_t \vartheta + \mathbf{u} \cdot \nabla \vartheta - \nabla \cdot (\kappa \nabla \vartheta) = \gamma$$

$\mathbf{u}$	velocity
μ	viscosity
$\dot{\varepsilon} = \frac{1}{2}\left(\nabla \mathbf{u} + \nabla \mathbf{u}^\top\right)$	strain rate tensor
p	dynamic pressure
ϑ	temperature
ρ	density
$\mathbf{g}$	gravity
κ	thermal conductivity
γ	heat sources

Stokes equation:

$$-\mathrm{div}\,(\nabla \mathbf{u} - p\mathbf{I}) = \mathbf{f}, \quad \mathrm{div}\,\mathbf{u} = 0$$

FEM Discretization:[1]

$$\mathbf{a}(\mathbf{u}_l, \mathbf{v}_l) + \mathbf{b}(\mathbf{v}_l, p_l) = \mathbf{L}(\mathbf{v}_l) \qquad \forall\, \mathbf{v}_l \in \mathbf{V}_l,$$

$$\mathbf{b}(\mathbf{u}_l, q_l) - \mathbf{c}(p_l, q_l) = 0 \qquad \forall\, q_l \in Q_l,$$

with: $\mathbf{a}(\mathbf{u}, \mathbf{v}) := \int_\Omega \nabla \mathbf{u} : \nabla \mathbf{v}\, dx, \quad \mathbf{b}(\mathbf{u}, q) := -\int_\Omega \mathrm{div}\,\mathbf{u} \cdot q\, dx$

Schur-complement formulation:

$$\begin{bmatrix} \mathbf{A}_l & \mathbf{B}_l^\top \\ \mathbf{0} & \mathbf{C}_l + \mathbf{B}_l \mathbf{A}_l^{-1} \mathbf{B}_l^\top \end{bmatrix} \begin{bmatrix} \underline{\mathbf{u}}_l \\ \underline{p}_l \end{bmatrix} = \begin{bmatrix} \underline{\mathbf{f}}_l \\ \mathbf{B}_l \mathbf{A}_l^{-1} \underline{\mathbf{f}}_l \end{bmatrix}$$

TERRA NEO

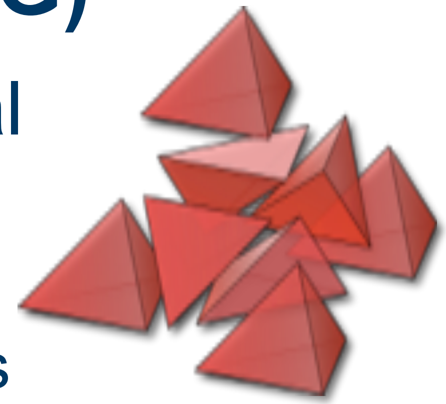
[1] B. Gmeiner et al. „Performance and Scalability of Hierarchical Hybrid Multigrid Solvers for Stokes Systems“, SIAM J. Sci. Comput. (2015)

LSS

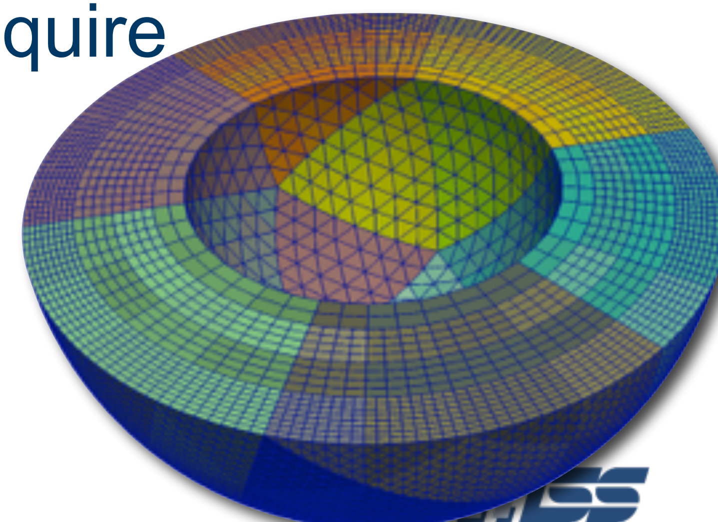
Hierarchical Hybrid Grids (HHG)

- Parallelize multigrid for tetrahedral finite elements
  - partition domain
  - parallelize all operations on all grids
  - use clever data structures
  - matrix free implementation
- Elliptic problems always require global communication

➡ Coarser grids for the global data transport



Bey‘s Tetrahedral Refinement

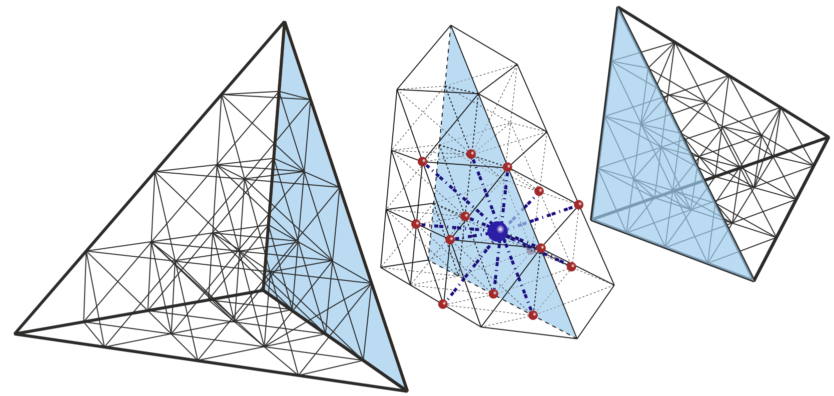


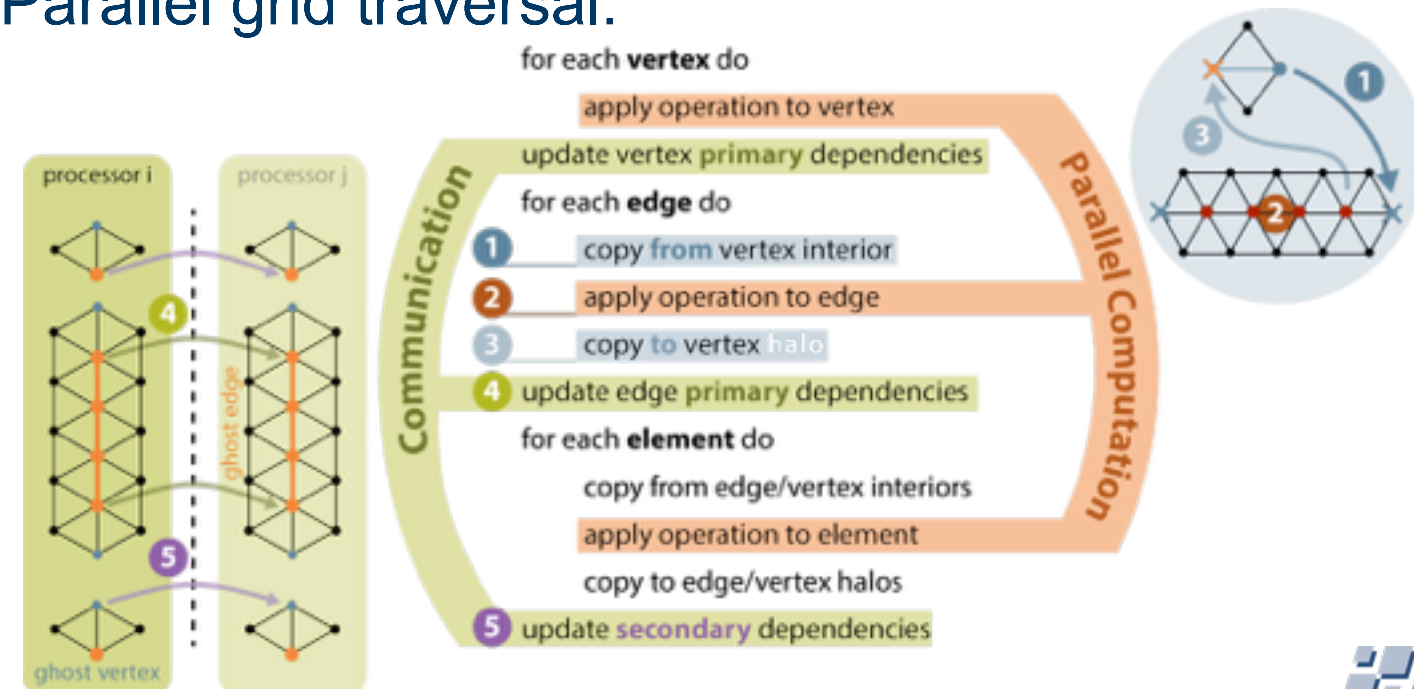
TERRA NEO

Tetrahedral discretization

HHG Concepts

- Structured refinement of unstructured base mesh
- Geometrical hierarchy: Volume, face, edge, vertex
- Parallel grid traversal:

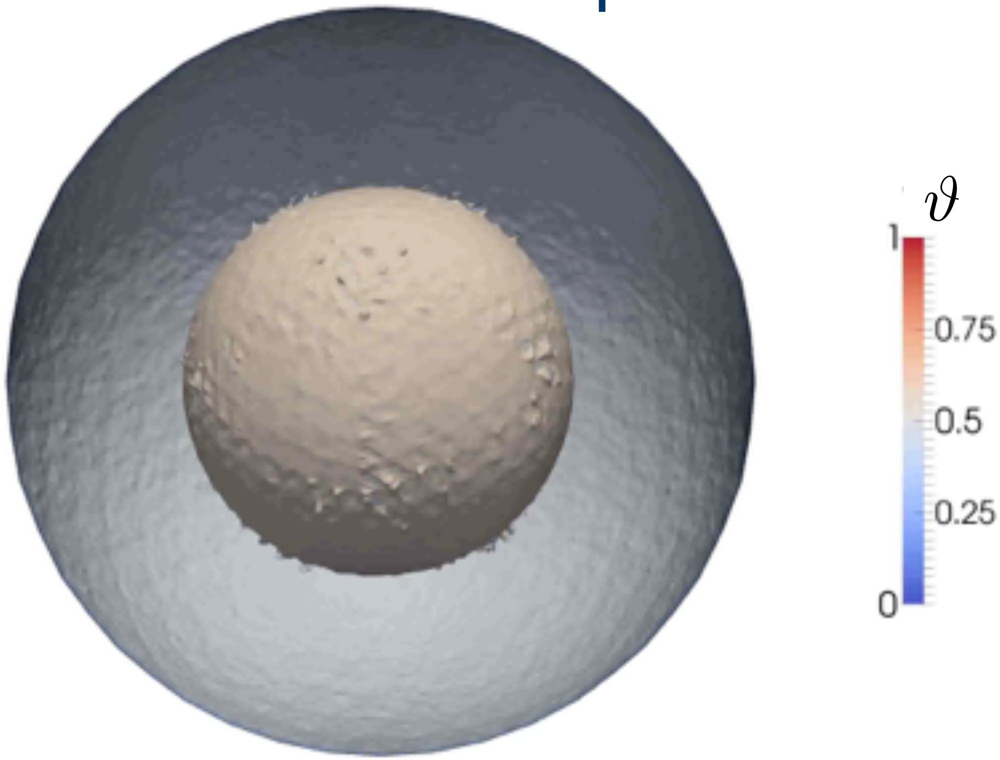




TERRA NEO

Coupled flow and transport (const. viscosity)

- Iso-surfaces of temperature distribution [1]



$$-\Delta \mathbf{u} + \nabla p = -\mathrm{Ra}\,\vartheta\,\hat{\mathbf{r}}$$

$$\nabla \cdot \mathbf{u} = 0$$

$$\partial_t \vartheta + \mathbf{u} \cdot \nabla \vartheta = \mathrm{Pe}^{-1} \Delta \vartheta$$

➡ Finite Elements with $6.5 \cdot 10^9$ DoF, 10 000 time steps

➡ Run-time 7 days on mid-size cluster (LSS): 288 cores (9 nodes)

[1] B. Gmeiner et al. "Performance and scalability of hierarchical hybrid multigrid solvers for Stokes systems", SIAM J. Sci. Comput. (2015)

Rheology in asthenospheric channel

- Viscosity model [1] (asthenosphere thickness 410 km)

$$\mu(\vartheta, r) = \left(e^{4.61 \frac{1-r}{1-r_{inner}} - 2.99\,\vartheta}\right) \begin{cases} \frac{1}{10}\, d_a^3 & \text{for} \quad r > 1 - d_a \\ 1 & \text{else} \end{cases}$$

d_a	relative asthenosphere thickness
r_{inner}	relative core radius

- Temperature distribution [2,3] recovered from seismic data

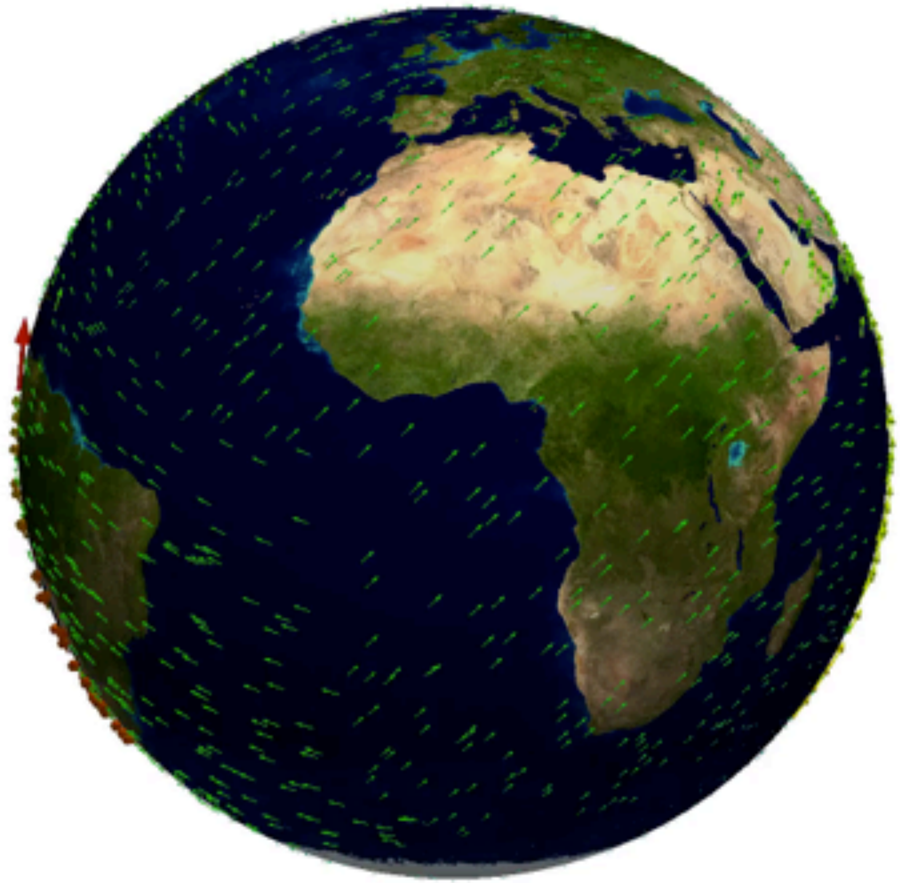
[1] D. R. Davies et al. "Reconciling dynamic and seismic models of Earth's lower mantle: The dominant role of thermal heterogeneity" Earth Planet. Sci. Lett. (2012)

[2] S. P. Grand et al. "High resolution global tomography: a snapshot of convection in the Earth" Geol. Soc. Am. Today (1997)

[3] L. Stixrude and C. Lithgow-Bertelloni "Thermodynamics of mantle minerals – I. Physical properties" Geophys. J. Int. (2005)

Temperature distribution

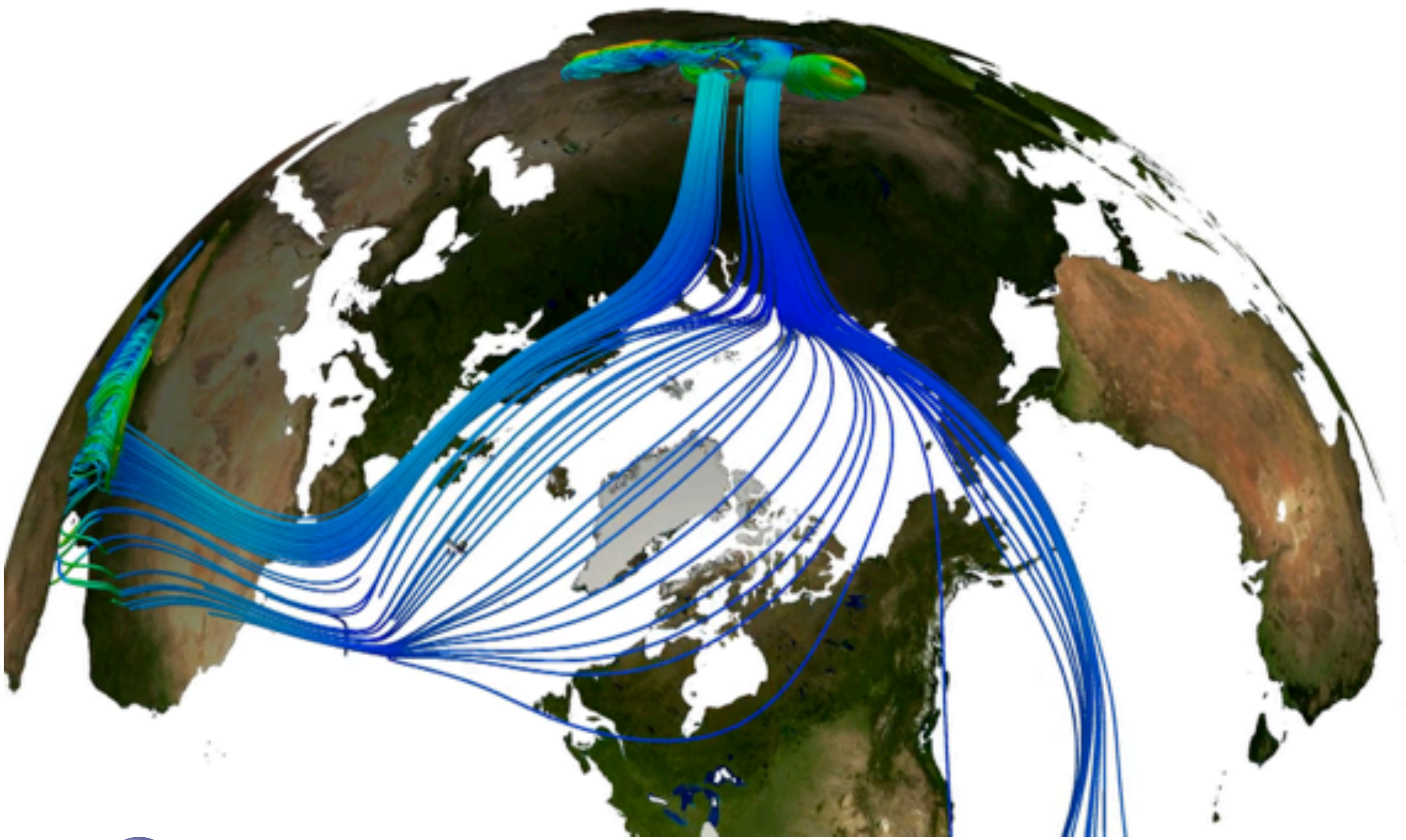
Velocity boundary conditions



[1] J. A. Boyden et al. "Next-generation plate-tectonic reconstructions using GPlates" Geoinformatics: cyberinfrastructure for the solid earth sciences, Cambridge Univ. Press (2011)

[2] M. Gurnis et al. "Plate tectonic reconstructions with continuously closing plates" Compt. Geosci. (2012)

Stationary velocity



Towards exa-scale computing

- Strong scaling is examined for two problems
  - Laplace problem with full multigrid
  - Stokes problem with pressure correction scheme [1]

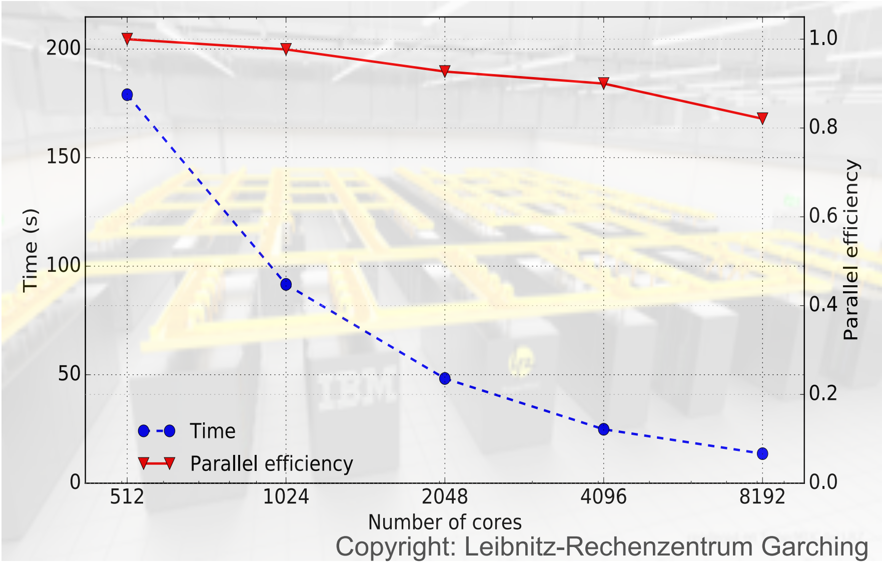
- Weak scaling of Stokes problem is investigated with an all-at-once (Uzawa) multigrid [2]

[1] B. Gmeiner et al. "Performance and Scalability of Hierarchical Hybrid Multigrid Solvers for Stokes Systems“, SIAM J. Sci. Comput. (2015)

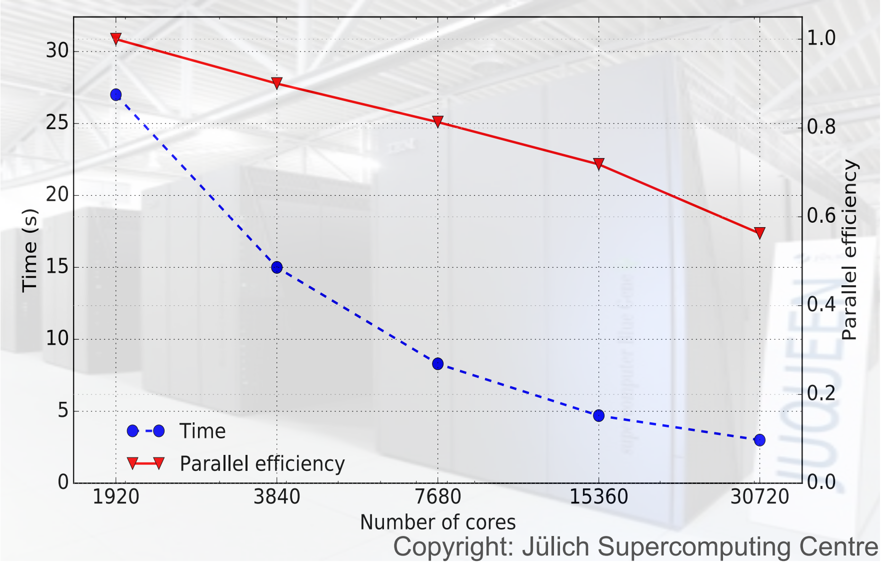
[2] B. Gmeiner et al. "A quantitative performance analysis for Stokes solvers at the extreme scale“, J. Comput. Sci. (2016)

Strong scaling results

Poisson problem



Stokes problem



➡ Reduced scalability of Stokes problem results from saddle point problem and growing coarsest grid cost

Weak scaling results

- Optimized for minimal memory consumption
  - 10^{13} unknowns correspond to 80 TByte for solution vector
  - JUQUEEN has 450 TByte memory
  - Matrix-free implementation essential

nodes	threads	DoFs	multigrid iterations	time to solution	parallel efficiency
5	80	$2.7 \cdot 10^{9}$	10	685 s	100 %
40	640	$2.1 \cdot 10^{10}$	10	704 s	97 %
320	5 120	$1.2 \cdot 10^{11}$	10	742 s	92 %
2 560	40 960	$1.7 \cdot 10^{12}$	9	720 s	95 %
20 480	327 680	$1.1 \cdot 10^{13}$	9	776 s	88 %

➡ Excellent weak scalability up to 10^{13} degrees of freedom (DoF) for Stokes problem

TERRA NEO

Conclusion

- 10^{13} degrees of freedom for FE are possible
- Multigrid scales to Peta (and beyond)
- Excellent time to solution
- HHG - lean matrix-free implementation

- Next: Re-design for higher-order finite elements