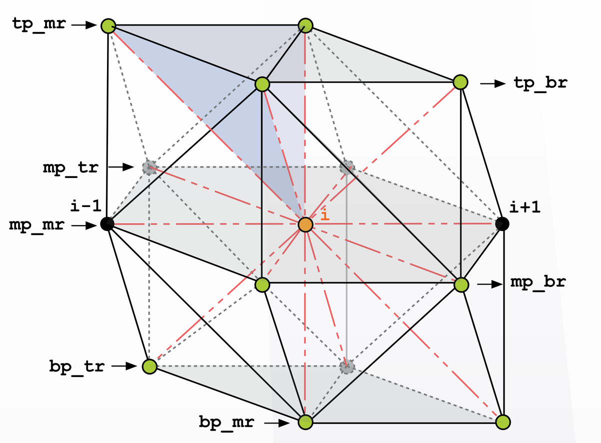
Integrated Co-Design of Future Exascale Software

M. Bauer, B. Gmeiner, M. Huber, L. John, U. Rüde, H. Stengel, C. Waluga, B. Wohlmuth

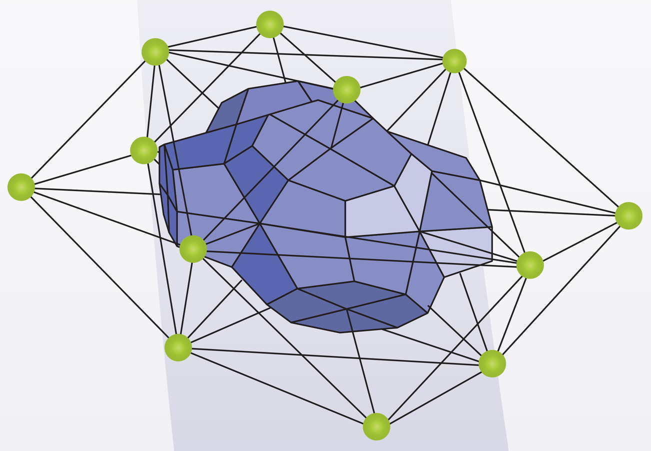
Co-designed algorithms for the numerical approximation of elliptic partial differential equations on modern supercomputers play a more and more important role in the future design of exa-scale enabled software. Here, we focus on several key ingredients such as node performance, ultra scalable multigrid methods, scheduling techniques for uncertain data and fault tolerant iterative solvers. In the case of a hard fault, we combine domain partitioning with highly scalable geometric multigrid schemes to obtain fast fault-robust solvers. The recovery strategy is based on a hierarchical hybrid concept where the values on lower dimensional primitives such as faces are stored redundantly and thus can be recovered easily. The lost volume unknowns are re-computed approximately by solving a local Dirichlet problem on the faulty subdomain. Different strategies are compared and evaluated with respect to performance, computational cost, and speed up. Locally accelerated strategies resulting in asynchronous multigrid iterations can fully compensate faults.

optimized stencil kernels




```
for (int i=1; i < (tsize-j-k-1); i=i+2) {
   u[mp_mr+i] = c[0] * (
    -c[1] *u[mp_mr+i+1] - c[2] *u[mp_tr+i-1] -
     c[3] *u[mp_tr+i]   - c[4] *u[tp_br+i]   -
     c[5] *u[tp_br+i+1] - c[6] *u[tp_mr+i-1] -
     c[7] *u[tp_mr+i]   - c[8] *u[bp_mr+i]   -
     c[9] *u[bp_mr+i+1] - c[10]*u[bp_tr+i-1] -
     c[11]*u[bp_tr+i]   - c[12]*u[mp_br+i]   -
     c[13]*u[mp_br+i+1] - c[14]*u[mp_mr+i-1] +
     f[mp_mr+i] );
  }
```

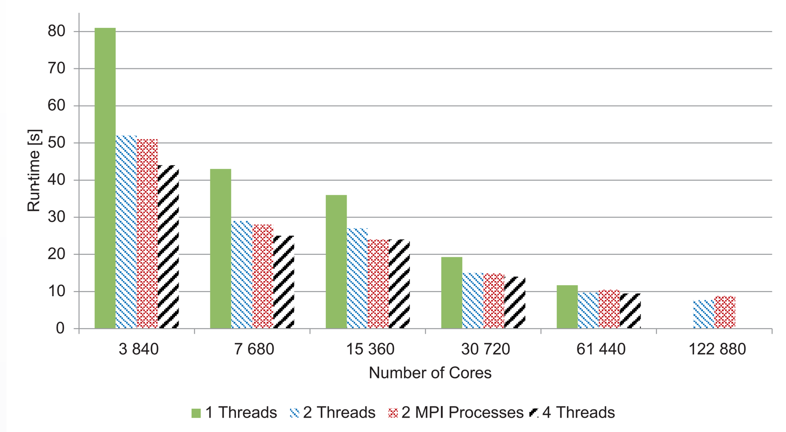
mass correction



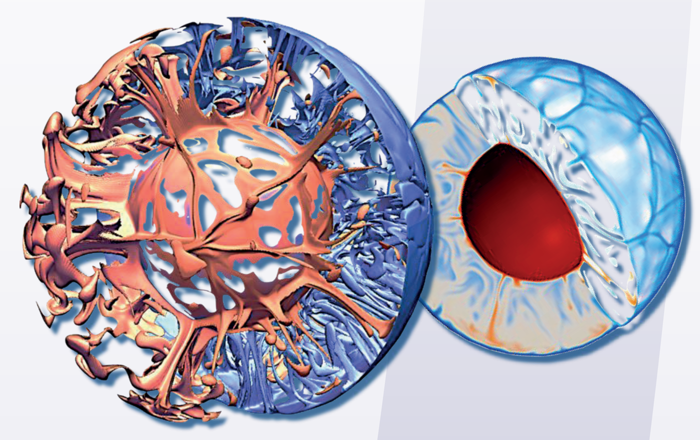
weak scalability

	JUQUEEN			SuperMUC		
	Rank 8, TOP 500, Nov. 2014			Rank 14, TOP 500, Nov. 2014		
DOF	Nodes	Threads	Time	Nodes	Threads	Time
$8.4 \cdot 10^7$	1	30	30 s	1	4	16 s
$6.4 \cdot 10^8$	4	240	38 s	2	30	20 s
$5.2 \cdot 10^9$	30	1920	40 s	15	240	24 s
$4.4 \cdot 10^{10}$	240	15260	44 s	120	1920	27 s
$3.4 \cdot 10^{11}$	1920	122880	48 s	960	15360	34 s
$2.8 \cdot 10^{12}$	15360	983040	54 s	7680	122880	41 s

strong scalability



coupled systems



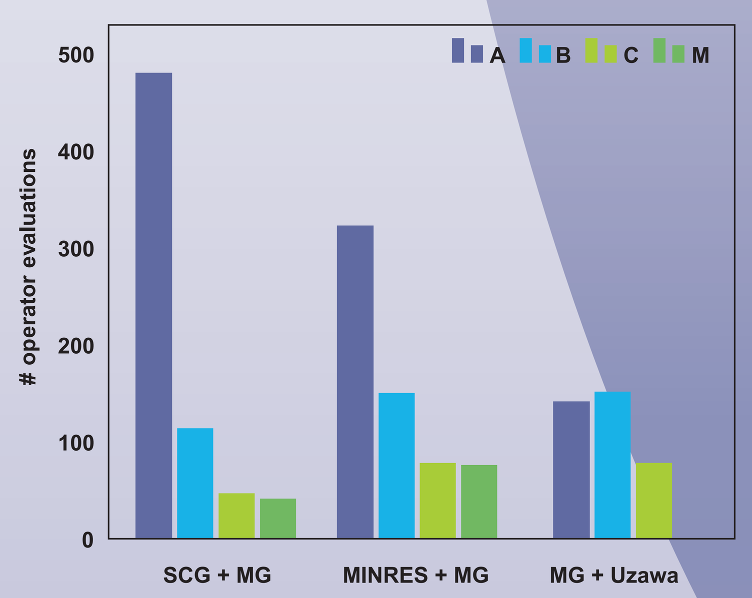
algorithmic resilience

Dirichlet-Neumann-strategy: Two consequtive failures at $k_F = 5$ and $k_F = 9$ with $n_s = 4$

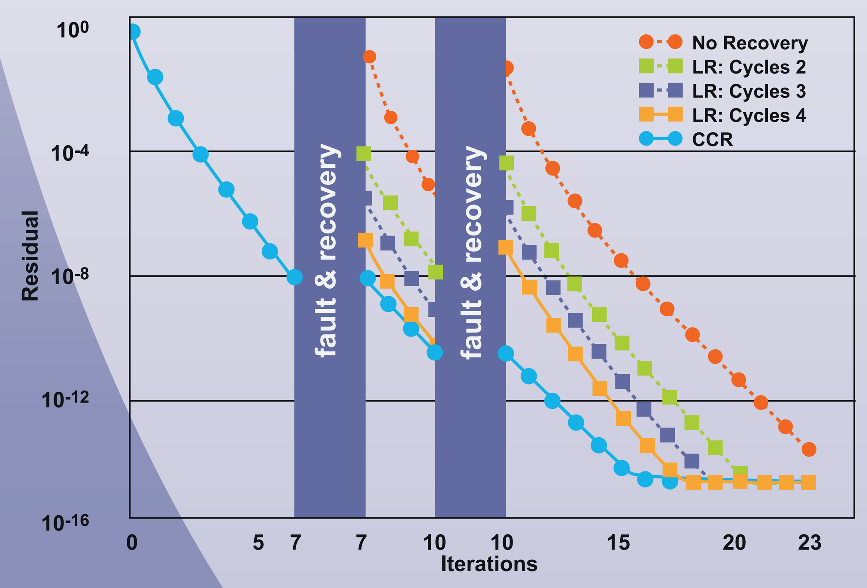
Size	No Rec	$n_I = 1$	2	3	4
769^3	19,21	8,01	0,05	– 0,38	4,22
1281^3	21,27	10,58	– 0,15	– 0,68	3,95
2305^3	18,50	7,91	– 0,33	– 0,87	3,76
4353^3	19,74	5,81	2,58	4,61	9,24

Global recovery can fully compensate faults w.r.t. time-to-solution.

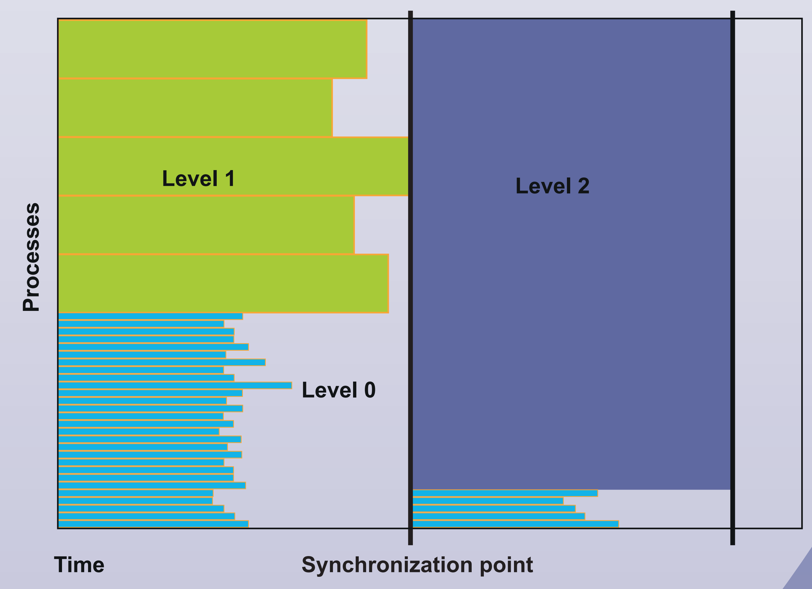
operator counts



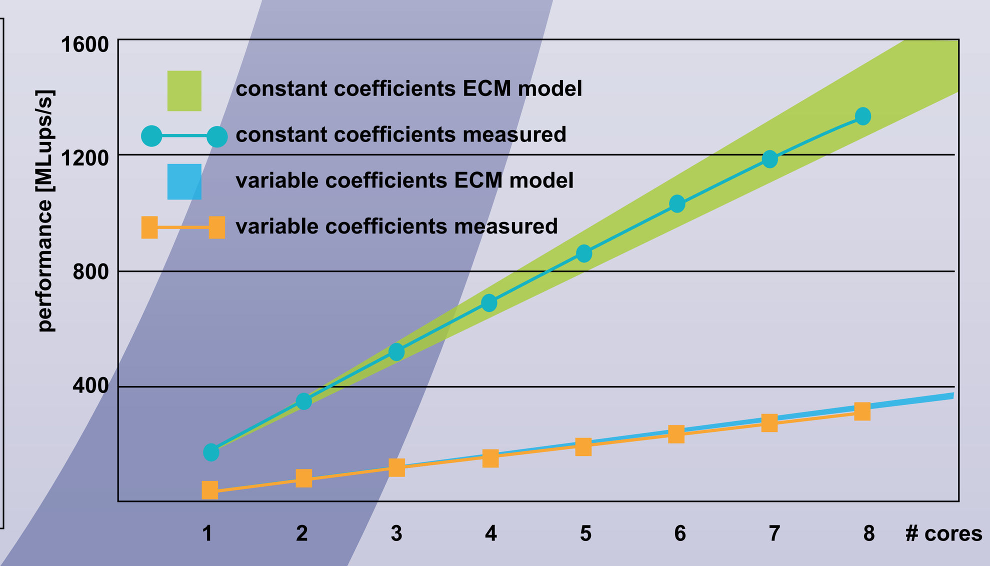
fault tolerance



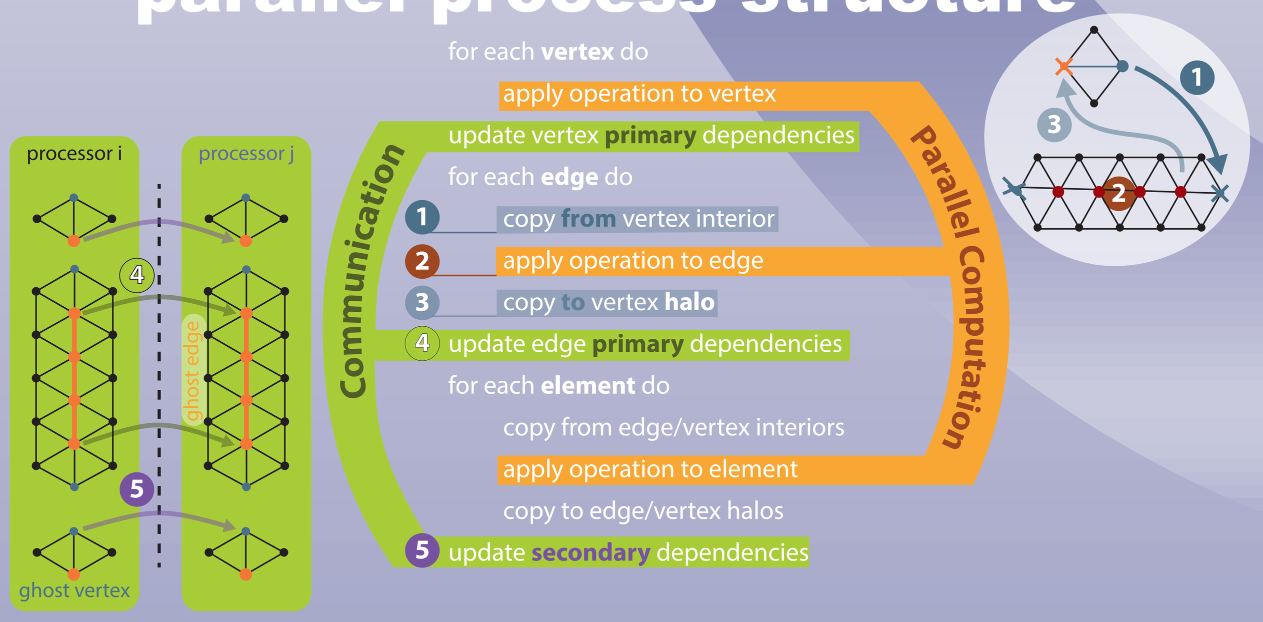
uncertain data



node performance

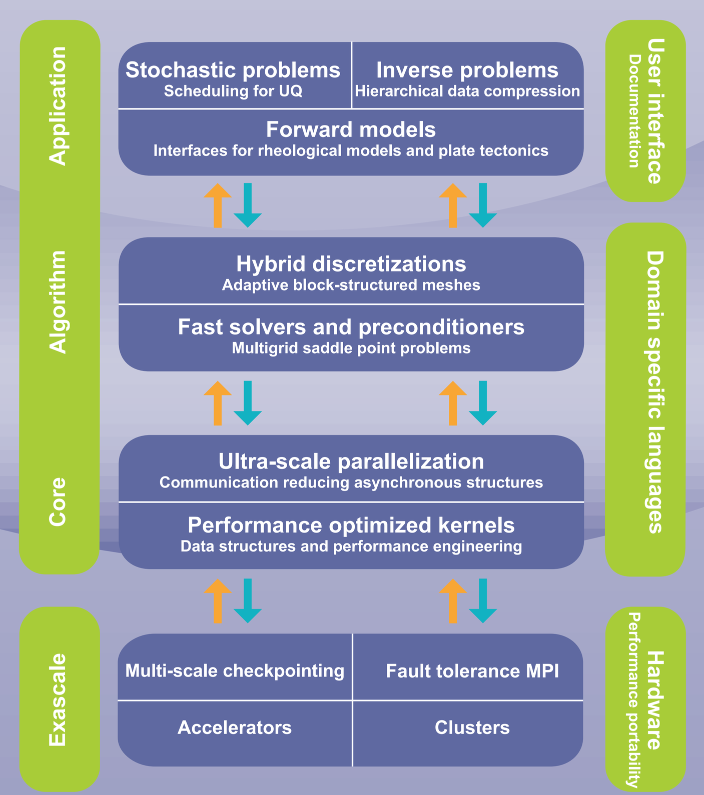


parallel process structure

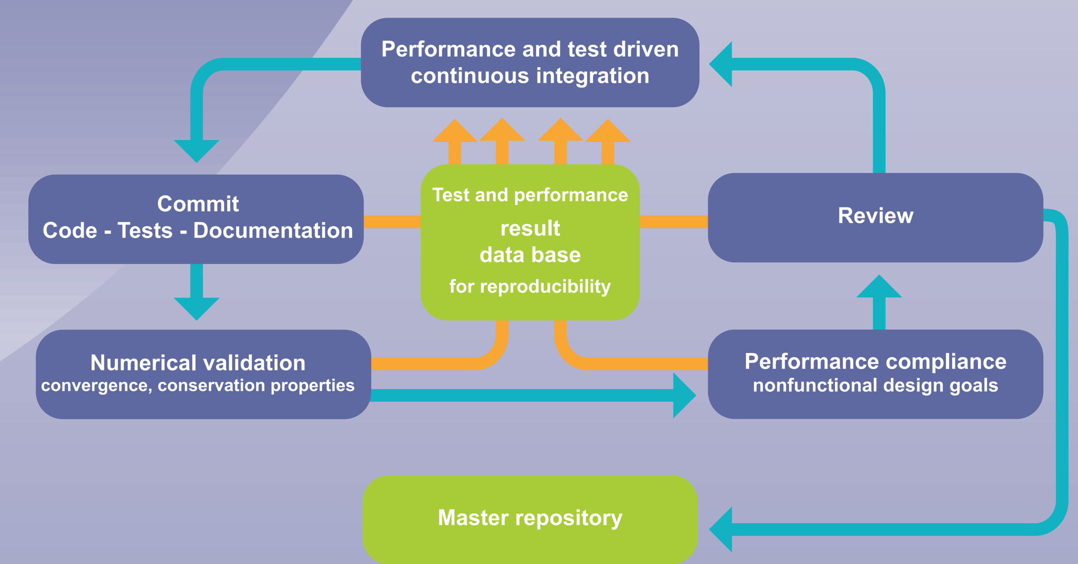


- block hybrid smoothers enhance parallelization
- lower primitive halos reduce communication
- stencil design supports optimized kernel performance
- matrix free structures push DoF to the limit

software structure



development process



- integrated co-design
- long term sustainability
- outreach and training
- developer ecosystem

challenges in co-design

computer sciences: software design for future exascale

mathematics: HPC performance oriented metrics

applications: model complexity and uncertainty

bridging disciplines: integrated co-design

exascale relevance

algorithmically supported resilience and reliability

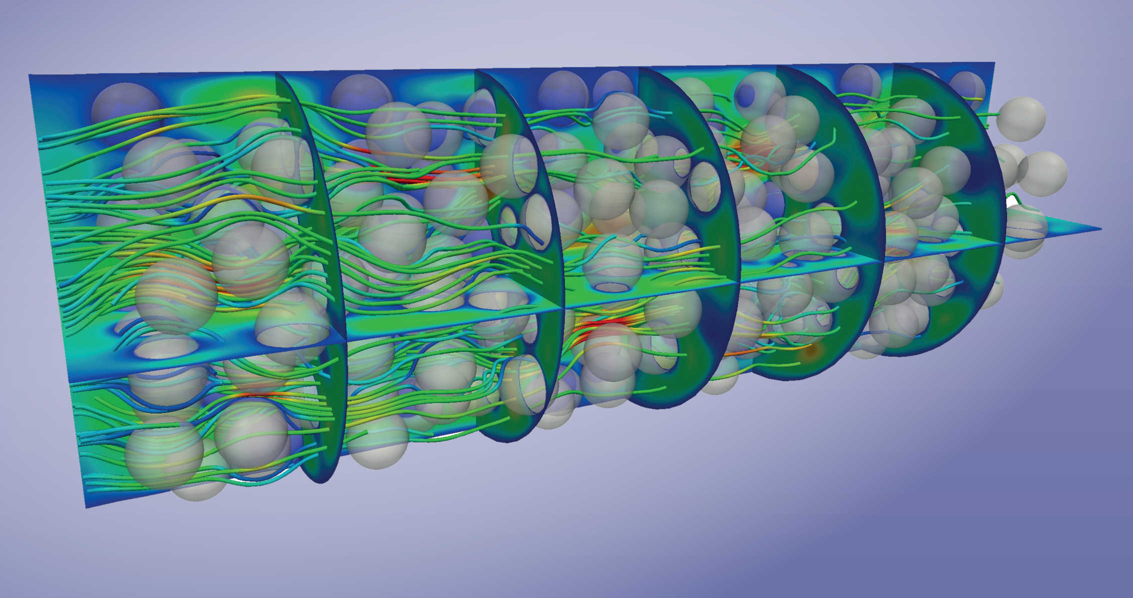
fast and ultra-scalable solvers and discretizations

hybrid scheduling for multi-scale complexity

asynchronous execution and hierarchical compression

architecture-aware optimized node performance

flow around a sphere package



references

B. Gmeiner, U. Rüde, H. Stengel, C. Waluga, and B. Wohlmuth. Performance and Scalability of Hierarchical Hybrid Multigrid Solvers for Stokes Systems. SIAM J. Sci. Comput., 37(2):C143–C168, 2015.

B. Gmeiner, U. Rüde, H. Stengel, C. Waluga, and B. Wohlmuth. Towards textbook efficiency for parallel multigrid. Numer. Math. Theory Methods Appl., 8, 2015.

M. Huber, L. John, B. Gmeiner, U. Rüde, and B. Wohlmuth. Massively parallel hybrid multigrid methods for the Stokes system, 2015. In preparation.

M. Huber, B. Gmeiner, U. Rüde, and B. Wohlmuth. Resilience for mutlgrid software at the extreme scale, 2015. ArXiv e-prints, 1506.06185.