Extreme scale multi-physics simulations of the tsunamigenic 2004 Sumatra megathrust earthquake

Carsten Uphoff[1], Sebastian Rettenberger[1]*, Leonhard Rannabauer[1], Michael Bader[1], Stephanie Wollherr[2], Thomas Ulrich[2], Elizabeth Madden[2], Alice-Agnes Gabriel[2]
(1) Department of Informatics, Technical University of Munich
(2) Department of Earth and Environmental Sciences, LMU Munich
* former member

Methods and algorithms

Our code SeisSol uses an ADER-DG method, which is suitable for linear hyperbolic PDEs; e.g. the elastic wave equations:

$$\frac{\partial}{\partial t}\sigma_{ij} - \lambda\delta_{ij}\frac{\partial}{\partial x_k}u_k - \mu\left(\frac{\partial}{\partial x_j}u_i + \frac{\partial}{\partial x_i}u_j\right) = 0$$
$$\rho\frac{\partial}{\partial t}u_i - \frac{\partial}{\partial x_j}\sigma_{ij} = 0$$

The numerical scheme consists of element-local and small matrix chain multiplications. Data exchange only happens between immediate face-neighbours.

ADER time integration

$$\frac{\partial^{k+1}Q_n}{\partial t^{k+1}} = -\tilde{K}_\xi\frac{\partial^k Q_n}{\partial t^k}A^* - \tilde{K}_\eta\frac{\partial^k Q_n}{\partial t^k}B^* - \tilde{K}_\zeta\frac{\partial^k Q_n}{\partial t^k}C^*$$
$$\mathcal{I}_n = \sum_{k=0}^{N}\frac{\Delta t^{k+1}}{(k+1)!}\frac{\partial^k Q_n}{\partial t^k}$$

Element-local, matrix-based update scheme

$$Q_{n+1} = Q_n + \tilde{K}_\xi\mathcal{I}_n A^* + \tilde{K}_\eta\mathcal{I}_n B^* + \tilde{K}_\zeta\mathcal{I}_n C^* - \sum_{i=1}^{4}F_i^-\mathcal{I}_n A_i^+ - \sum_{i=1}^{4}F_{i,j_i,h_i}^+\mathcal{I}_n^{(m_i)}A_i^-$$

Local time stepping (LTS)

Static adaptivity, and especially low-quality elements may lead to severe time step restrictions, due to the CFL condition

$$\Delta t \leq \frac{C}{2N+1}\cdot\frac{2r_{in}}{c_p},$$

where r_{in} is the element's insphere radius. As a remedy we use clustered LTS [1, 2]:

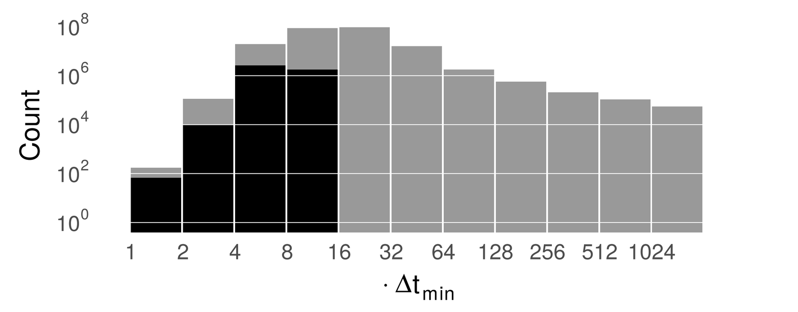


Figure 1 : Elements are partitioned into time clusters for a more regular update scheme suitable for modern hardware.

Code generation

We use an embedded domain-specific language (eDSL), specifically written for matrix chain products [3], to generate our computation-intensive kernels:

- Automatic detection of irrelevant matrix entries in matrix chain products

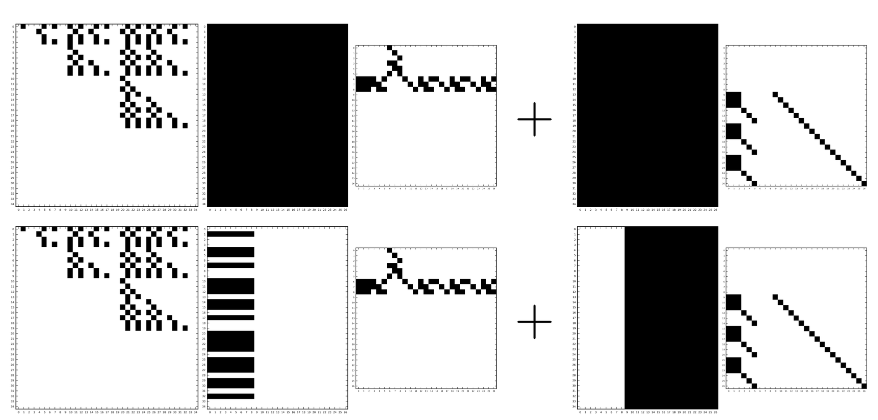


Figure 2 : Partial kernel before (top) and after (bottom) removing irrelevant entries in matrix chain products.

- Automatic determination of zero-paddings for aligned loads and stores
- Enabling block decompositions for structured sparse matrices
- Automatic solution of matrix chain ordering problem

To make our eDSL more broadly applicable, we are currently extending it to also support tensor contractions; e.g.

```
D['abckl'] <= A['ijmc'] * B['mkab'] * C['ijl']
```

generates a kernel for

$$D_{abckl} = \sum_{i,j,m} A_{ijmc}B_{mkab}C_{ijl}$$

HW-SW-codesign: L2 cache efficiency

We use $F^+_{i,j,h} = R_i f_h(R_j)^T$.

	Memory		Multiplications	
Deg. N	$48\times F^+$	R, f, R^T	$F^+\mathcal{I}A^-$	$R((f(R^T\mathcal{I}))A^-)$
4	459 KB	38.1 KB	13860	12960
5	1176 KB	83.8 KB	32760	26838
6	2646 KB	165.4 KB	70308	51660

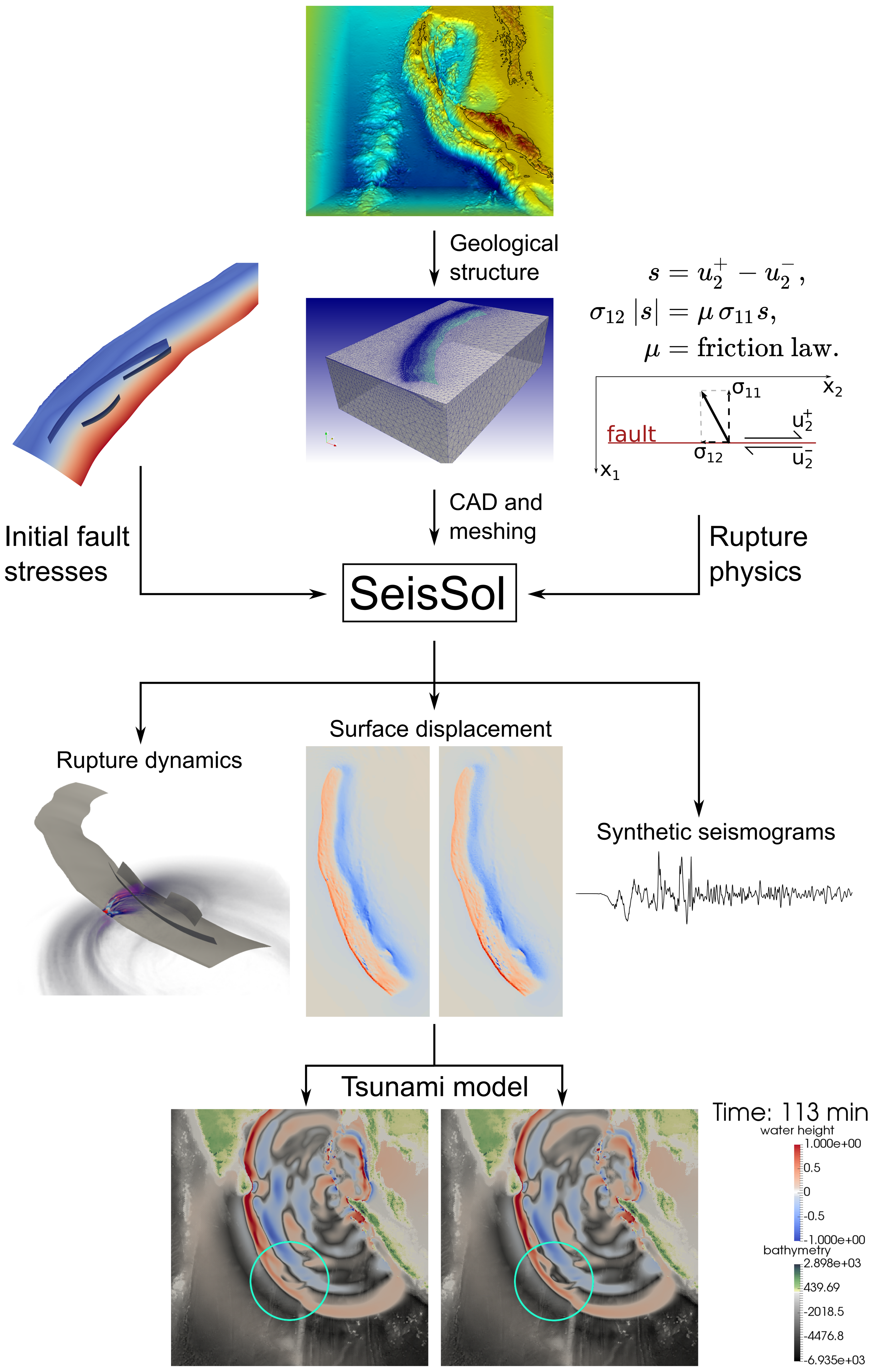
The 2004 Sumatra-Andaman earthquake

An extreme event in every respect:

- Failure of 1300 km–1500 km of the Sumatra subduction zone
- 8 minutes of shaking
- 230,000–280,000 fatal victims
- Caused a tsunami up to 30 m high
- M_w 9.1–9.3

Largest model has 221 million elements (111 billion degrees of freedom) with 400 m fault resolution and 1.3–2.2 km material layer resolution [2].

ASCETE modelling workflow



Results

- Scenario matches key observations: magnitude, slip distribution, and displacement.
- 2.2 Hz frequency content.
- Splay faults generate additional displacement.
- New high-resolution sea-floor displacement output is used as initial condition for tsunami models:

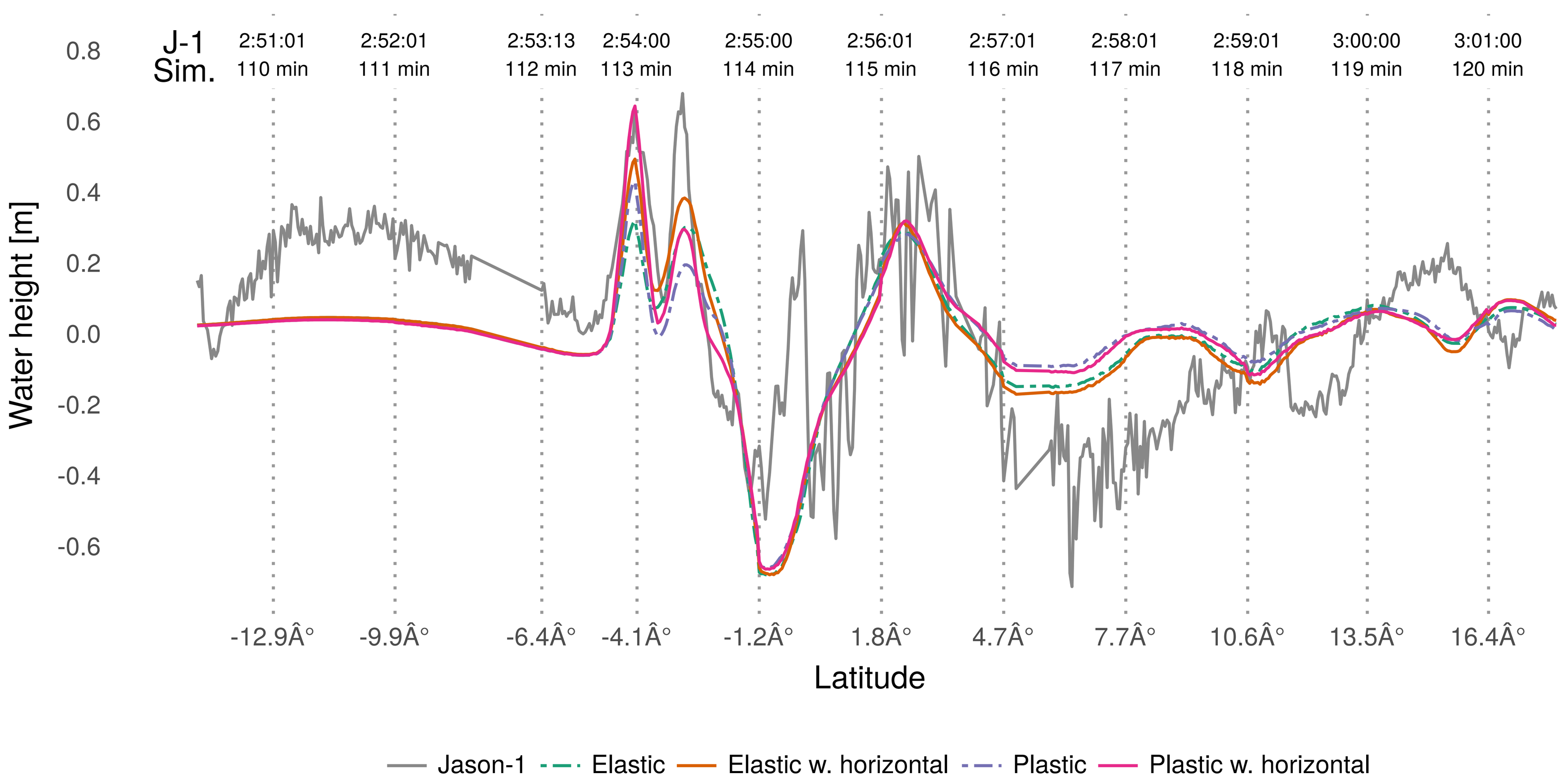


Figure 3 : Comparison of simulated water height with satellite data [4]. Surface displacement output of SeisSol is used as initial condition.

Performance

- **M** SuperMUC Phase 2, 3072 nodes, 2x Intel Xeon E5-2697 v3 / node, Infiniband FDR14 interconnect with 4:1 pruned fat tree topology.
- **S** Shaheen II, 6174 nodes, 2x Intel Xeon E5-2698 v3 / node, Cray Aries interconnect with Dragonfly topology.
- **C** Cori, 9688 nodes, 1x Intel Xeon Phi 7250 / node, Cray Aries interconnect with Dragonfly topology.

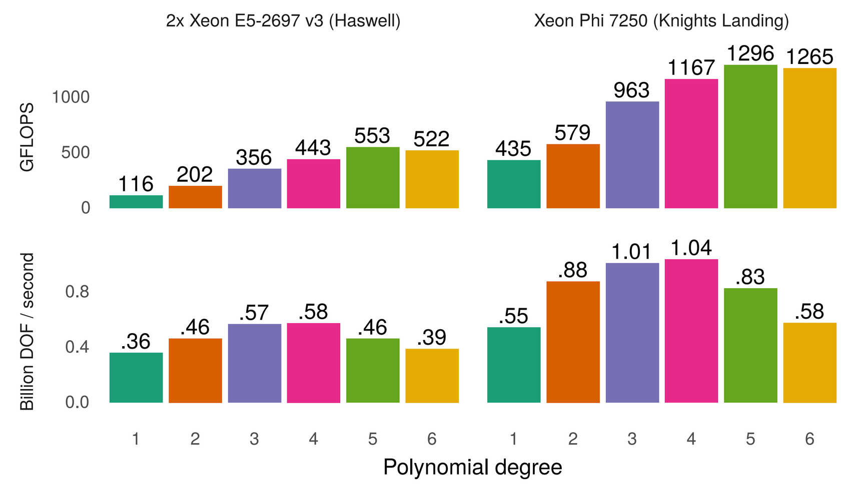


Figure 4 : Shared memory performance for 100000 elements.

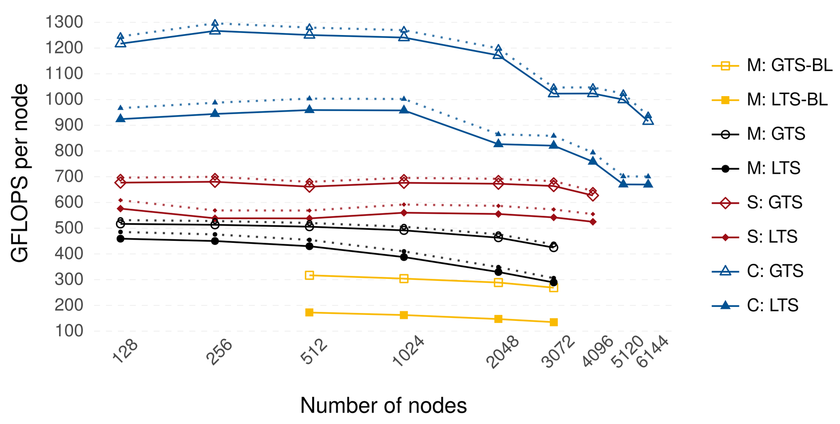


Figure 5 : Strong scaling with 221 million elements and 5th degree polynomials.

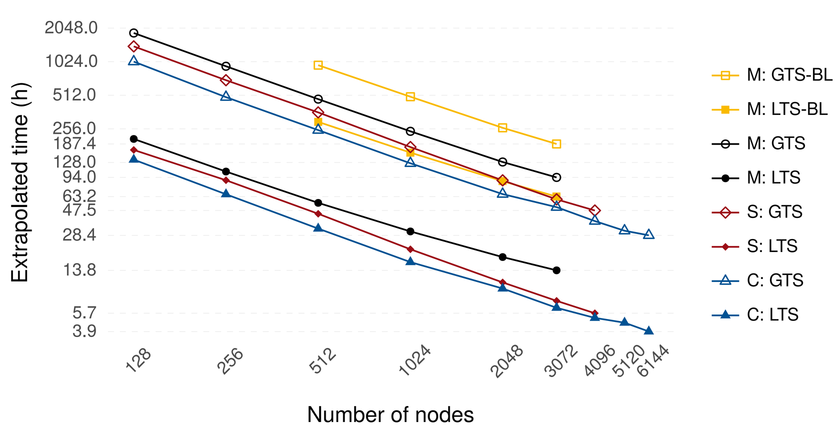


Figure 6 : Local time-stepping reduces time to solution.

Reproducible science

SeisSol is selected for the SC18 Student Cluster Competition Reproducibility Challenge. The foundation is the artifact appendix in [2].

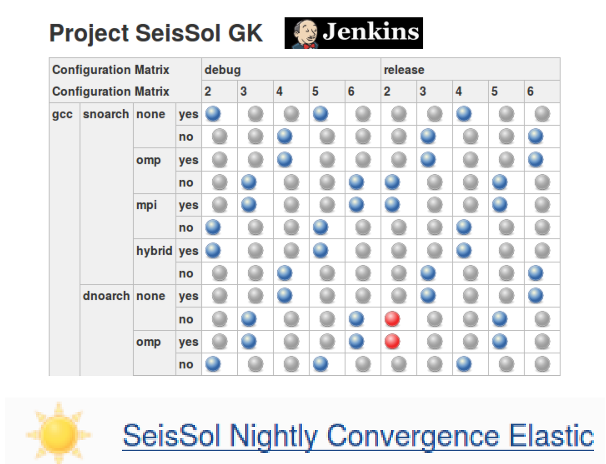
Continuous integration

Continuous integration (CI) strongly helps in the development process of SeisSol, simply because our software is tested automatically and thus more often.

Low effort CI for code consistency across developers and simple build or unit tests:

git Travis CI

High effort CI for complex tests such as (sparse) build matrix of all configurations and nightly convergence tests:



References

[1] A. Breuer, A. Heinecke, and M. Bader, Petascale local time stepping for the ADER-DG finite element method, in *2016 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, pages 854–863, 2016.

[2] C. Uphoff et al., Extreme scale multi-physics simulations of the tsunamigenic 2004 Sumatra megathrust earthquake, in *SC '17: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, Denver, CO, USA, 2017, SC17 Best Paper.

[3] C. Uphoff and M. Bader, Generating high performance matrix kernels for earthquake simulations with viscoelastic attenuation, in *Proceedings of HPCS 2016*, 2016.

[4] Sea surface height anomalies derived from the Jason-1 data record, `ftp://podaac-ftp.jpl.nasa.gov/allData/jason1/L2/j1_ssha/c109/`.

uphoff@in.tum.de

github.com/SeisSol/SeisSol