SDSC SAN DIEGO SUPERCOMPUTER CENTER

UC San Diego

Fused Seismic Simulations with the Discontinuous Galerkin Method at Extreme-Scale

Alexander Breuer, Alexander Heinecke (Intel), Yifeng Cui

Extreme-scale Discontinuous Galerkin Environment (EDGE)

- ★ Focus: Problem settings with high geometric complexity, e.g., mountain topography
- ★ Unique support for fused simulations exploiting inter-simulation parallelism
- ★ Rapid prototyping through support for different element types: Line elements, rect. quads, 3-node triangles, rect. hexes, 4-node tets
- ★ Equations: Advection (FV+ADER-DG: 1D, 2D, 3D), Shallow Water Equations (FV: 1D), Elastic Wave Equations (FV+ADER-DG: 2D, 3D)
- ★ Parallelization: Assembly kernels for WSM, SNB, HSW, KNC (non-fused), KNL (fused & non-fused), OpenMP (custom), MPI (overlapping)
- ★ Continuity: Continuous Integration (sanity checks), Continuous Delivery (automated convergence + benchmarks runs), code coverage, license checks, container bootstrap
- ★ License: BSD 3-Clause (software), CC0 for supporting files, e.g., user guide

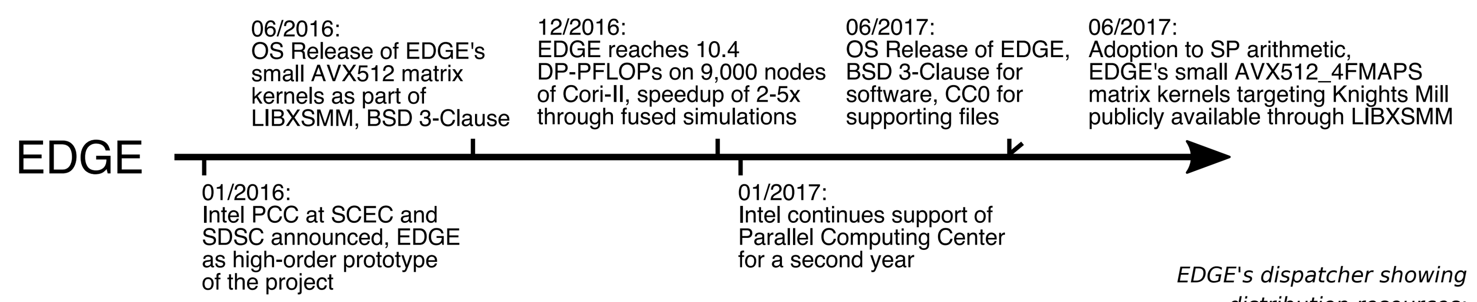
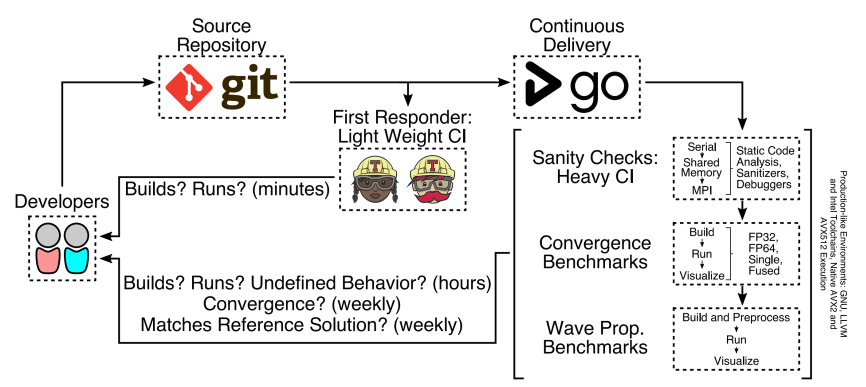
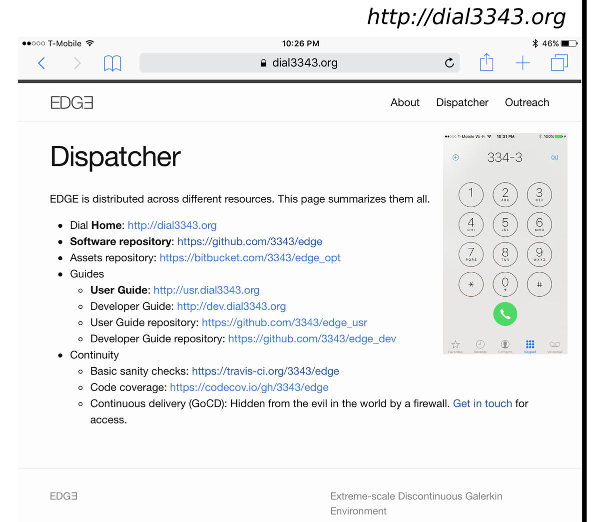


Illustration of EDGE's continuous integration and continuous delivery pipelines. Git Logo by Jason Long is licensed under the Creative Commons Attribution 3.0 Unported License.



EDGE's dispatcher showing distribution resources: http://dial3343.org



Fused Simulations Exploit Inter-Simulation Parallelism

"Why is this a good idea?"

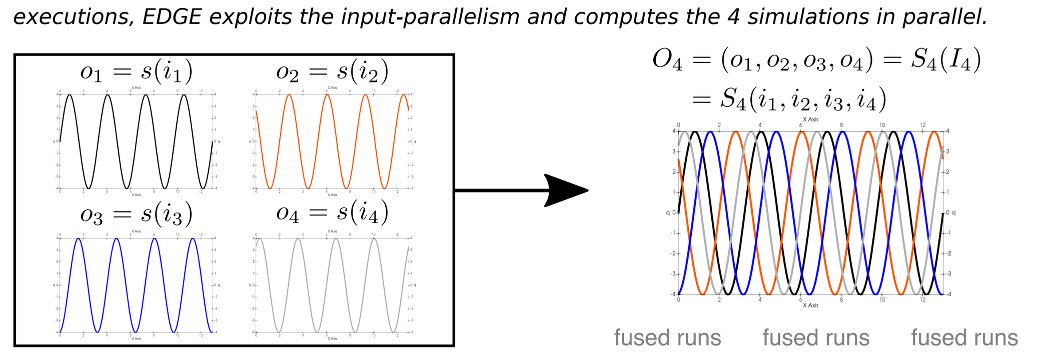
- ★ Idea: Exploit input parallelism by fusing multiple, similar simulations in a single execution of the solver
- ★ EDGE supports this idea at all levels of parallelism, starting at a single vector op
- ★ Fusing multiples of the vector-width (KNL: 8 simulations in double precision) allows for perfect vectorization without zero ops
- ★ Fusion of multiples of 64 bytes (8 simulations) leads to alignment to cache-lines without artificial zero-padding
- ★ Read-only data structures are shared among all fused simulations

"Similar simulations?"

- ★ EDGE imposes restrictions on fused seismic simulations:
  - ★ Identical mesh for all fused simulations
  - ★ Identical simulation parameters: start and end time, convergence rate, "frequency" of wave field output, "frequency" and location of seismic receivers
  - ★ Identical material parameters (velocity model)
  - ★ "Sources" mostly arbitrary: Arbitrary initial DOFs, kinematic sources: arbitrary location and moment rates, spontaneous rupture: identical friction law, other initial parameters arbitrary

$$q(x,t)_t + v \cdot q(x,t)_x = 0,\ v \in \mathbb{R}$$

Illustration of EDGE's fused approach (4 simulations), applied to the advection equation with sinusoidal initial values. While a traditional solver handles different initial values in multiple executions, EDGE exploits the input-parallelism and computes the 4 simulations in parallel.



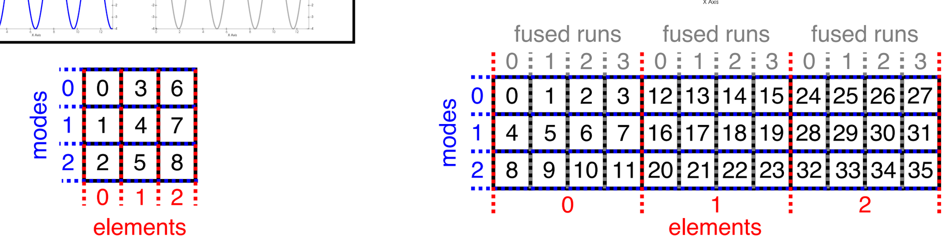
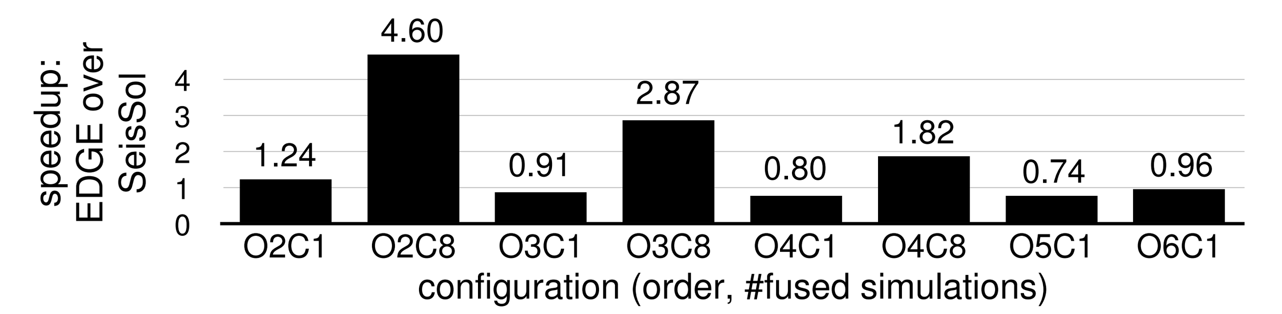


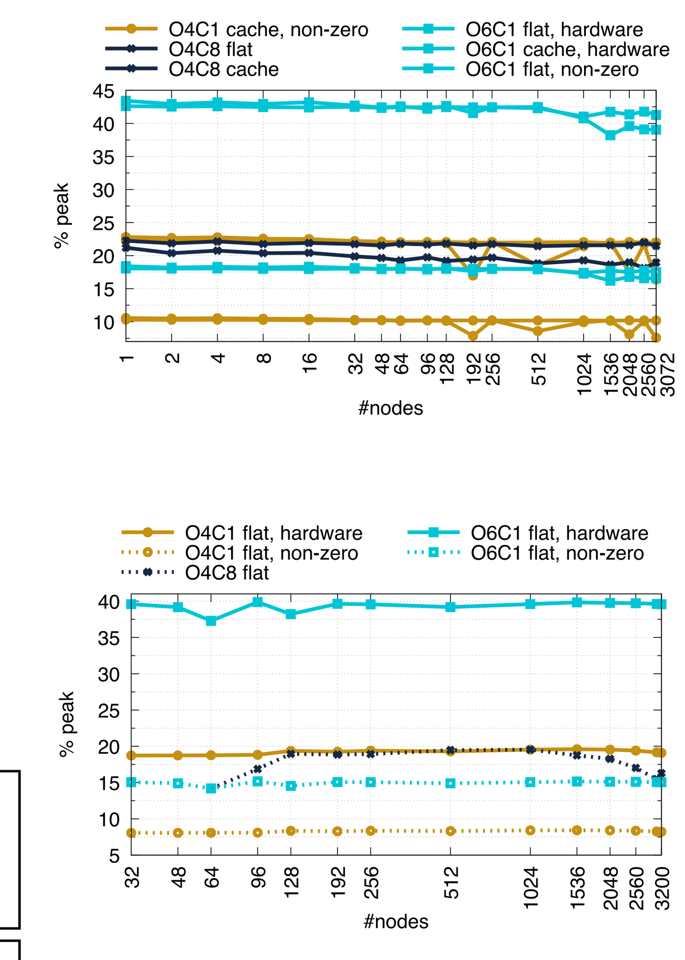
Illustration of EDGE's memory layout for a third order (P2 elements) ADER-DG solver for the advection equation. In the case of the traditional, non-fused approach, the DG-modes are the fastest dimension in memory, followed by the elements. In case of fused simulations, the fused runs are the fastest dimension, followed by the DG-modes and the elements.

Seismic Simulations at Extreme-Scale

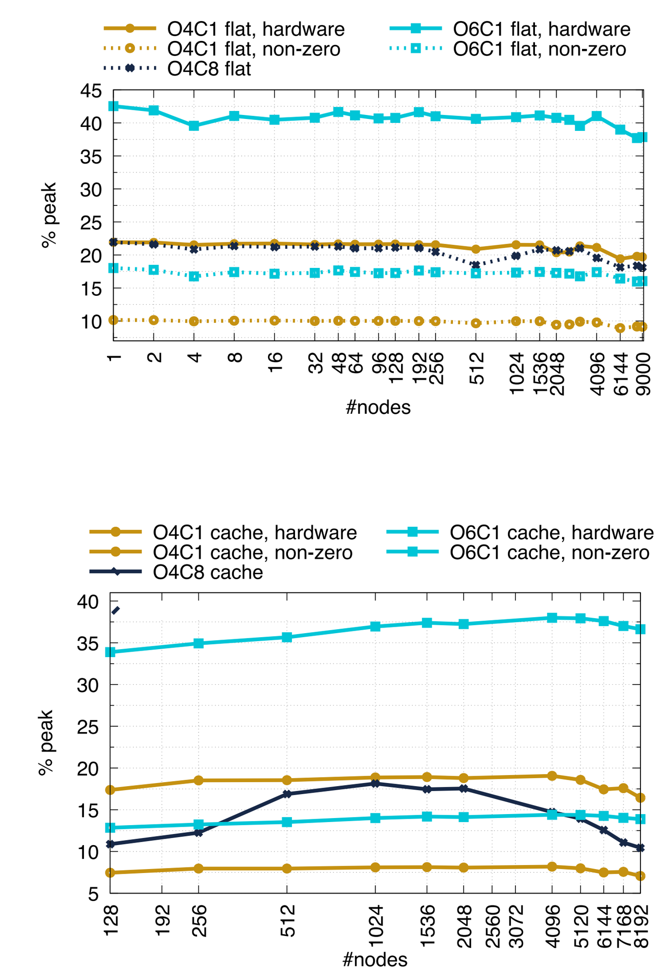
- ★ Fused simulations greatly outperform non-fused simulations, low orders of convergence gain from increased arithmetic intensity, higher orders from more science per FLOP (sparse operators)
- ★ EDGE features a highly optimized data layout, which splits elements participating in MPI-communication from those which are independent of MPI within a time step
- ★ Background progression of MPI-messages is ensured through a dedicated communication core, a minimal number of implicit barriers maximizes core utilization
- ★ EDGE prioritizes critical paths to maximize overlap of communication and computation
- ★ EDGE sustained 10.4 PFLOPS (double precision) on 9,000 nodes (612,000 cores) of Cori Phase 2 and surpassed the previous 8.6 PFLOPS record performance of SeisSol on 24,576 cards of Tianhe-2 [1]
- ★ Strong scaling the LOH.1 benchmark using an unstructured tetrahedral mesh of 172E6 elements let to near-perfect parallel efficiencies for an 50x (O4C8) and 100x (O4C1, O6C1) increase in computer power on Theta
- ★ Fusion of eight fourth order simulations (O4C8) outperforms the non-fused counterpart (O4C1) by 2.0x in flat mode at scale [1]



Speedup of EDGE over SeisSol (GTS, git-tag 201511, [2]). For convergence rates O2-O6 results for single non-fused forward simulations (O2C1-O6C1) are presented. Additionally, respective per-simulation speedups for orders O2-O4 are presented when using EDGE's full capabilities by fusing eight simulations (O2C8-O4C8). [1]

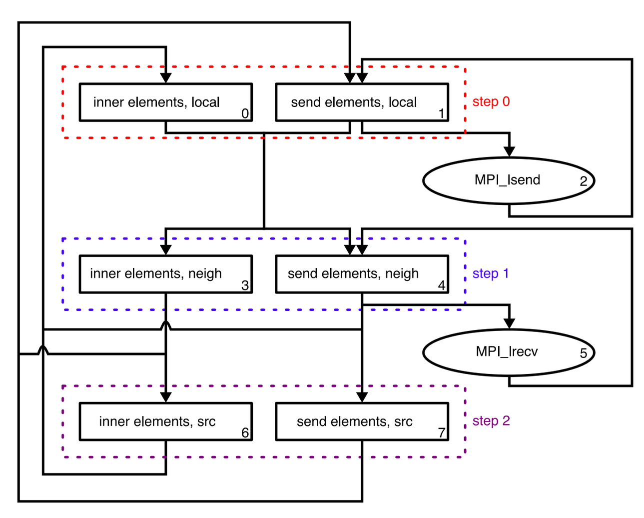


Weak scaling study (top) and strong scaling study (bottom) on Theta. The weak scaling used 276,480 tets per node and the strong scaling a total of 172,386,915 tets. Shown are hardware and non-zero peak efficiencies in cache and flat mode. O denotes the order and C the number of fused simulations.



Weak scaling study (top) and strong scaling study (bottom) on Cori Phase II. The weak scaling used 276,480 tets per node and the strong scaling a total of 340,727,199 tets. Shown are hardware and non-zero peak efficiencies in cache and flat mode. O denotes the order and C the number of fused simulations.

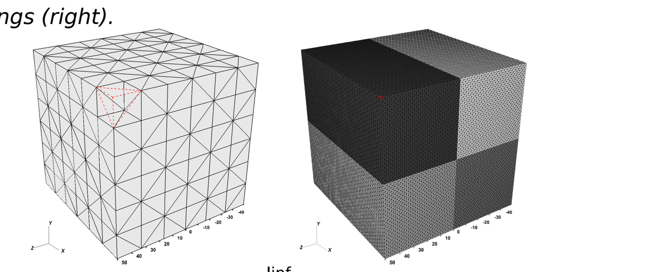
Illustration of EDGE's control flow and dependencies when solving the elastic wave equations with source terms. Shown is a single partition, where send elements are adjacent to elements owned by other MPI-ranks (faces as bridge). Inner elements are independent of neighboring ranks within a time step. Each of the three ADER-DG operations, 1) update with local contributions, 2) update with neighboring contr., and 3) kinematic sources, is split into inner- and send-regions. Send elements have priority.

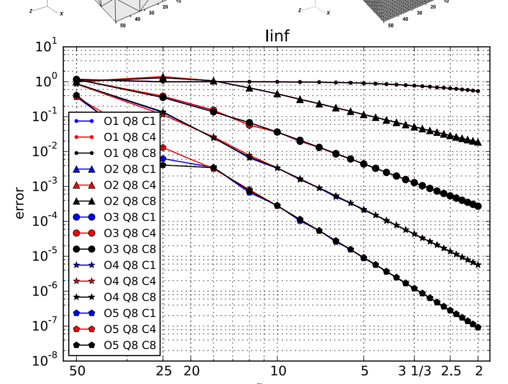


EDGE in Action

- ★ Example setups available as CC0 (public domain) for data and BSD 3-Clause for scripts
- ★ Support for convergence setups through periodic boundary conditions for all element types (incl. unstructured tet meshes)
- ★ Supported verification benchmarks: HHS1, HSP1a, LOH.1, LOH.2, Can4

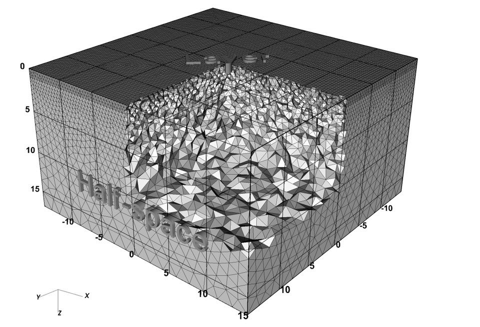
Two regular tetrahedral meshes. By imposing periodic boundary conditions, the plane wave initial solution is reproduced after diagonal propagation through the domain, even in MPI-parallel settings (right).

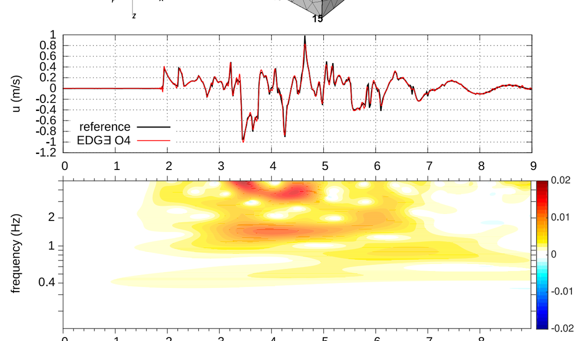




Convergence of EDGE for regular, tetrahedral meshes in the Linf-norm for the elastic wave equations in velocity-stress form. Shown are orders O1-O5 for quantity v (Q8) when utilizing EDGE's fusion capabilities with shifted initial conditions. [1]

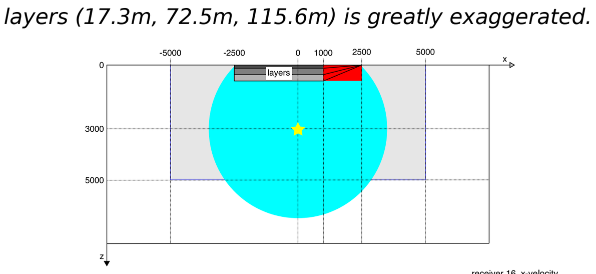
LOH.1 Benchmark: Example mesh and material regions.

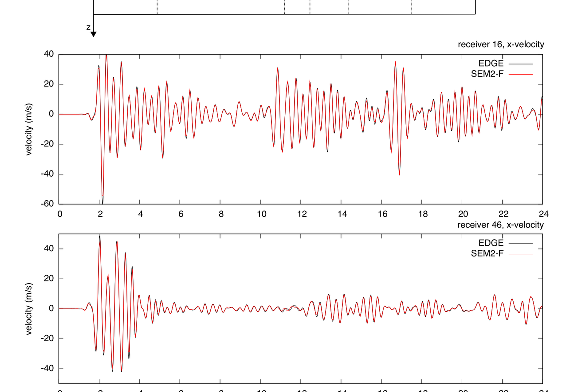




Synthetic seismogram [1] and time-frequency misfit (0.13Hz-0.5Hz) for quantity u at the ninth seismic receiver located at (8647 m, 5764 m, 0). Detailed setup: [1], Misfit visualization: TF-MISFIT_GOF_CRITERIA, http://nuquake.eu

Illustration showing the problem-adapted mesh resolution (colors) of the Can4 setup. The height of the layers (17.3m, 72.5m, 115.6m) is greatly exaggerated.





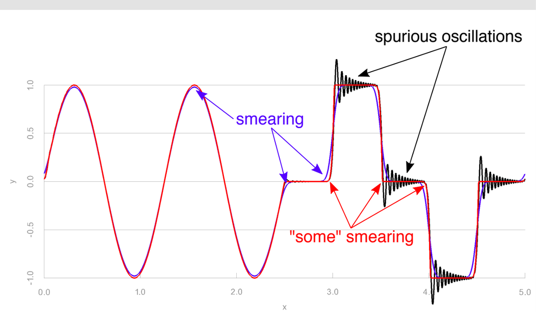
Comparison of EDGE's solution (black) to the reference solution SEM2-F of [3] (red) for the first 24s of the Can4 benchmark. Upper plot: Receiver above the wedge, 1km away from the northern edge of the valley. Lower plot: Receiver 1km away from the southern edge.

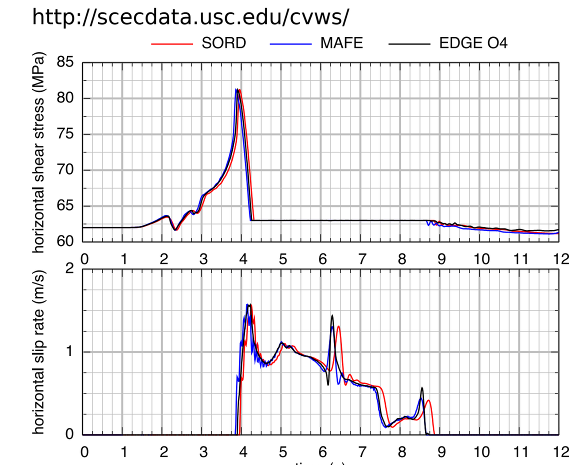
Outlook: A Glance into the Future

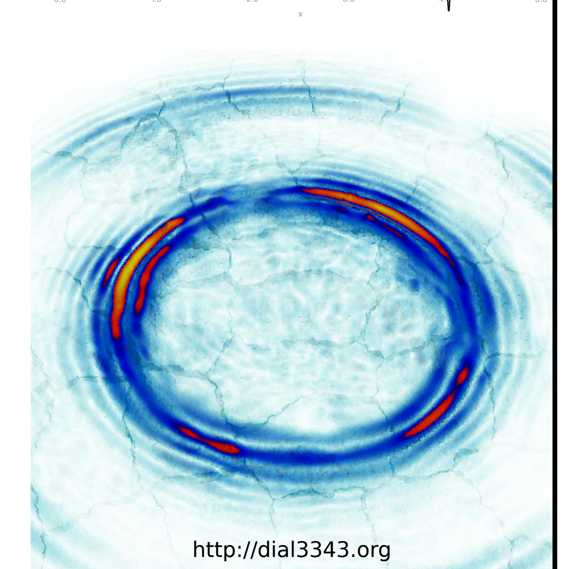
- ★ Support for multiphysics solvers at internal boundaries, targeting fused spontaneous rupture simulations
- ★ Grouped Local Time Stepping (LTS) for increased resolution at (internal) boundaries and to allow for "mistakes" of the volume mesher
- ★ DG-limiter to cope with large gradients in the solution, enabling EDGE for nonlinear hyperbolic PDEs
- ★ Last but not least: If you are interested in working with us, get in touch!

Right: Illustration of different solvers for the advection equation with periodic boundary conditions. Shown is a Comparison of 1) an unlimited O2 ADER-DG solver (black), 2) a FV solver (blue), and 3) a limited ADER-DG solver (red). 2500 elements were used for the DG solvers, 7500 for the FV solver.

Bottom: Fourth order simulation results (200m char. length) for the TPV5 benchmarks at fault receiver faultst075dp075 (strike 7.5 km, dip 7.5 km). EDGE's results are compared with results of the two codes SORD and MAFE, obtained from http://scecdata.usc.edu/cvws/







References & Support

[1] EDGE: Extreme Scale Fused Seismic Simulations with the Discontinuous Galerkin Method - A. Breuer, A. Heinecke, Y. Cui. In High Performance Computing: 32nd International Conference, ISC High Performance 2017, Frankfurt, Germany, June 18-22, 2017, Proceedings

[2] High Order Seismic Simulations on the Intel Xeon Phi Processor (Knights Landing) - A. Heinecke, A. Breuer, M. Bader, and P. Dubey. In High Performance Computing: 31st International Conference, ISC High Performance 2016, Frankfurt, Germany, June 19-23, 2016, Proceedings

[3] 3-D numerical simulations of earthquake ground motion in sedimentary basins: testing accuracy through stringent models - E. Chaljub, E. Maufroy, P. Moczo, J. Kristek, F. Hollender, P.-Y. Bard, E. Priolo, P. Klin, F. de Martin, Z. Zhang, W. Zhang, X. Chen In Geophys. J. Int. (2015) 201, Issue 1, 90–111

This work was supported by SCEC through award #16247.

This work is supported by Intel through the Intel Parallel Computing Center "Accurate and Efficient Earthquake Simulations on Intel Xeon Phi".

This research used resources of the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

This research used resources of the Argonne Leadership Computing Facility (ALCF), which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357.

This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575.

This research is part of the Blue Waters sustained-petascale computing project, which is supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the state of Illinois.