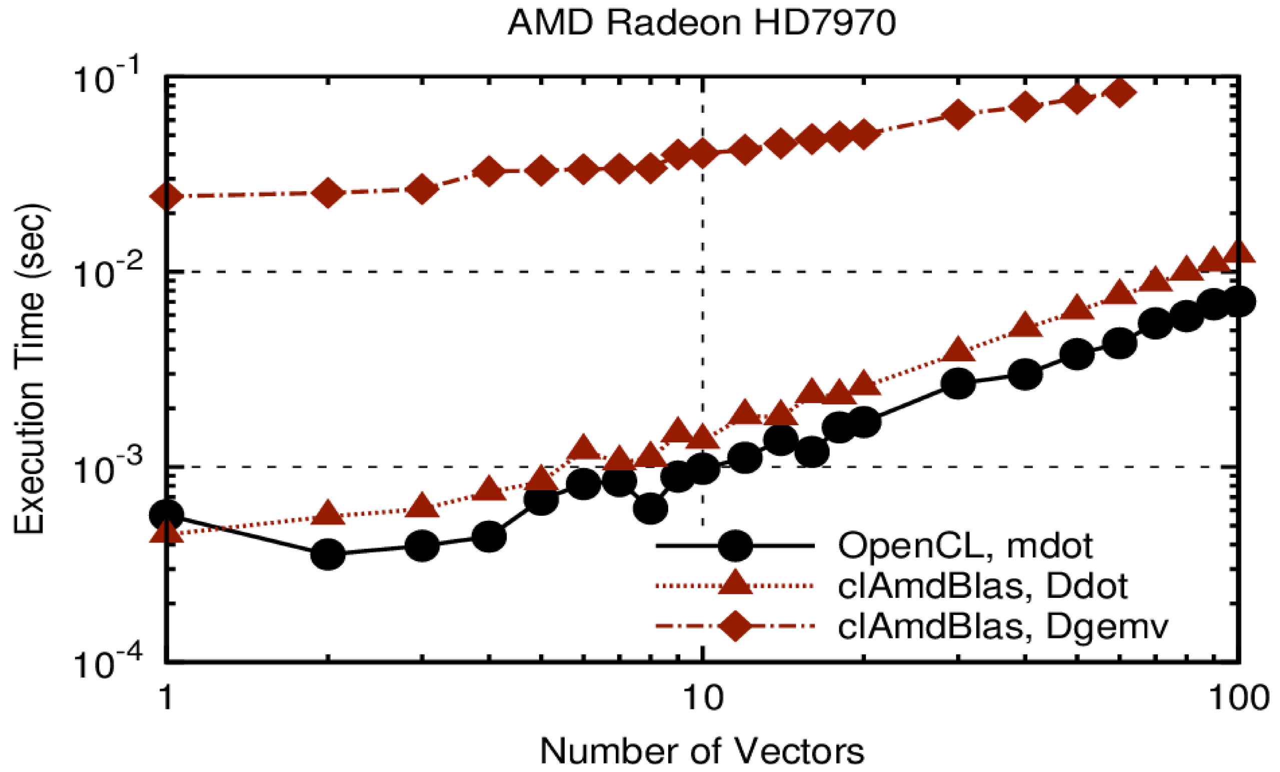
FASTMath Iterative Solver Technologies

Work in FASTMath in the area of iterative solutions focuses on commonly used methods, such as lower level kernels, Krylov subspace methods, and algebraic multigrid (AMG), that can be used in many different applications. Efforts have improved both the algorithms themselves and their implementations.

Broad Accelerator Support via OpenCL

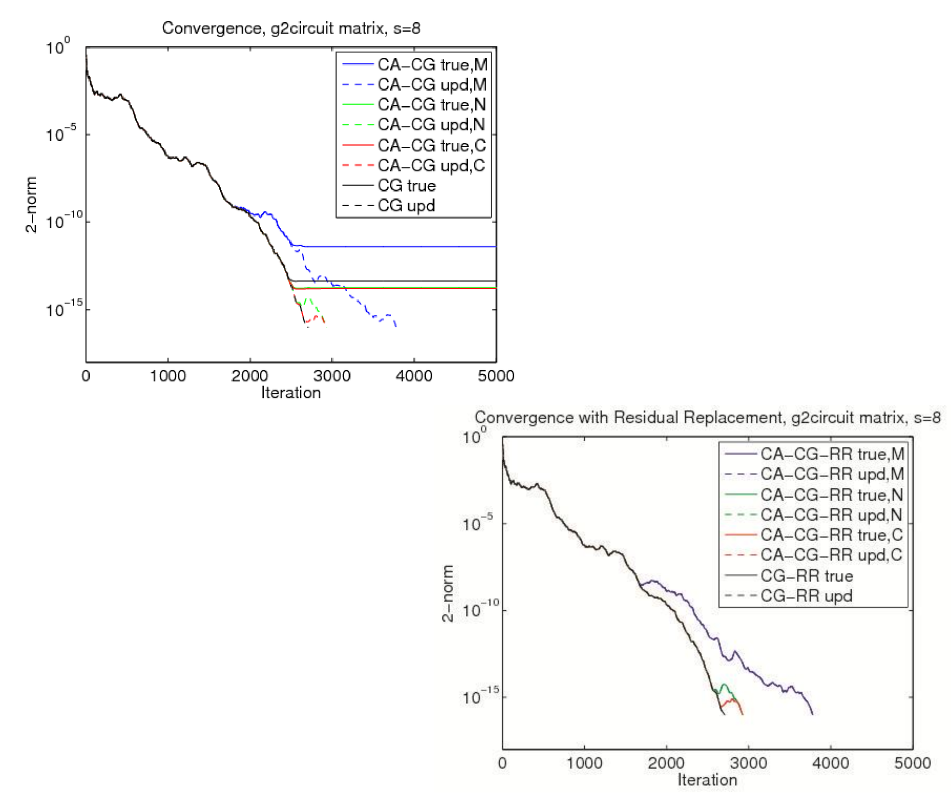
- Porting our existing CUDA-based operations to OpenCL allows for leveraging the performance of both current and future hardware from different vendors. By supplementing vendor-provided BLAS implementation by domain-specific kernels, high performance is obtained.



Source: ANL PETSc Team

Communication-Avoiding Krylov methods

- Fixed a source of numerical instability in Communication-Avoiding (CA) Krylov Methods via Residual Replacement.



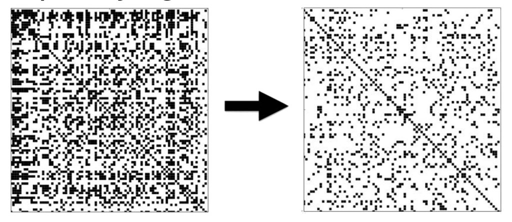
- Extended CA Krylov Methods to widely used nonsymmetric solvers like BiCGStab.

Source: UC Berkeley

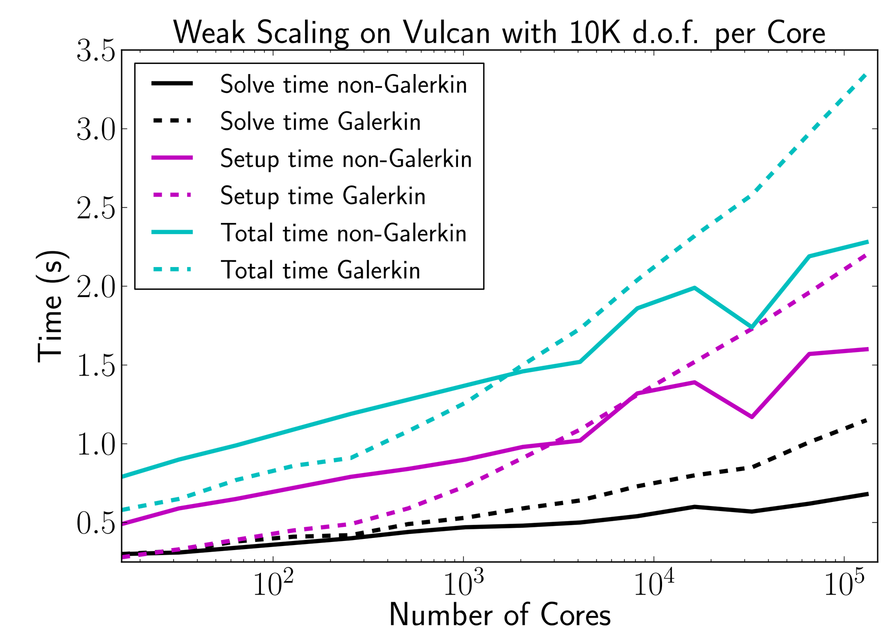
Non-Galerkin AMG

- Standard AMG suffers from fill-in on coarse-levels, which in turn induces increasing communication complexities at large scale.
- This project builds a mathematical and algorithmic framework to reduce this fill-in and hence also reduce parallel communication and runtime

Sparsifying a coarse-level matrix



- The approach safely eliminates matrix entries in the standard Galerkin coarse-level matrix to yield a non-Galerkin AMG method, while preserving important near null-space components to maintain good AMG convergence.
- Parallel test uses a set of best practices AMG parameters for the Galerkin AMG data, and then turns on non-Galerkin for a comparison.
  - Total speedup is significant (~50%) and grows with core count.
  - Solve and setup phases are accelerated



Source: LLNL hypre Team

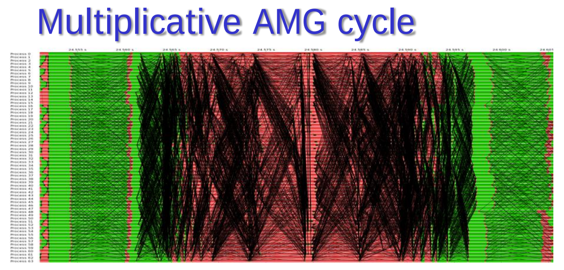
Reducing Communication with Additive AMG Variants

- Increasing communication complexities on coarser levels in AMG can significantly decrease performance.
- Classical additive AMG methods have improved communication complexities per cycle, but converge significantly slower than multiplicative AMG.
- Mult-additive AMG, which uses a smoothed truncated prolongator, converges significantly faster than additive AMG. Omitting most of the smoothing portion in the mult-additive V-cycle, and with it communication leads to simplified mult-additive AMG.
- Solve times on a Linux cluster with Infiniband fat tree network for an unstructured 3D problem with jumps on a sphere (left) and a diffusion problem with a 27 pt stencil on BG/Q (right)

- multiplicative AMG
  (--- GS, — Jac. smoothing)
- mult-additive AMG.
- simplified mult-additive AMG

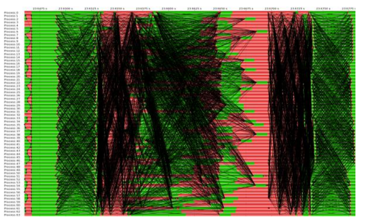
- Improved computation /communication overlap leads to faster cycle times.

Multiplicative AMG cycle

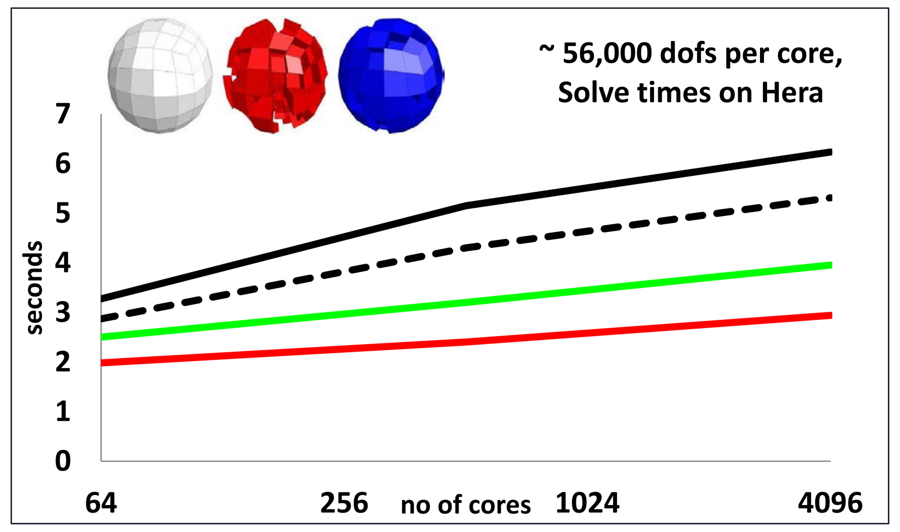


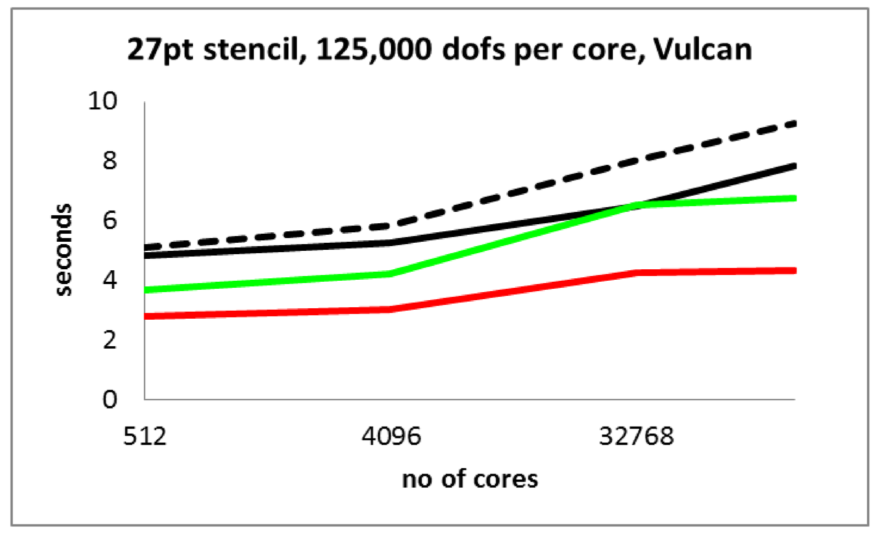
- **Performance profiles for 64 MPI tasks on Hera**

computation
idle time
MPI calls



Mult-Additive AMG with truncated smoothed interpolation



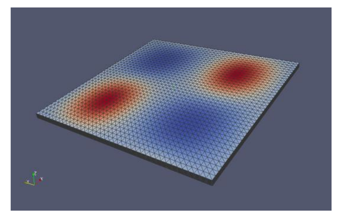


Source: LLNL hypre Team

Parallel AMG Based on Energy Minimization

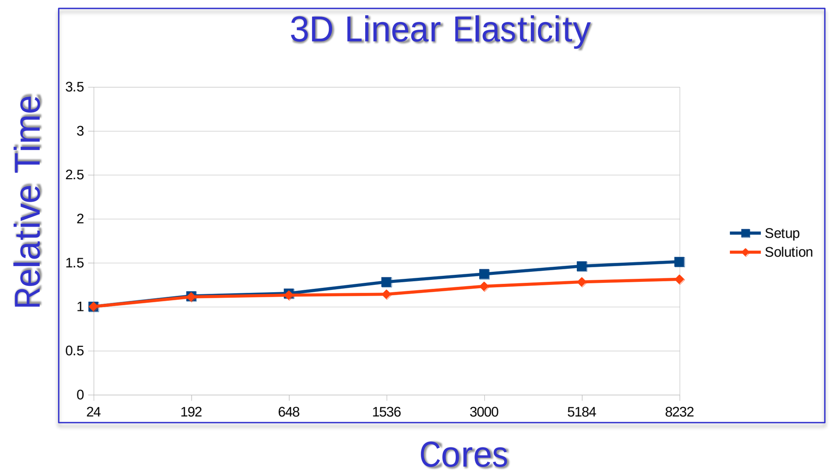
- Generate interpolation operators P
- Input: sparsity pattern, near nullspace
  - Can reuse prior interpolation operators

$$\text{Find } P = \operatorname{argmin} \sum \|P_k\|_X$$
$$\text{s.t. sparsity}(P) \in S$$
$$N_f = P N_c$$



- Benefits multiple solve situations
- Setup expense offset by reusing prior P as initial guess to current emin setup

- Promising initial weak scaling



Krylov Iterations Versus #Emin Setup Steps

Step	Emin(6)	Emin(1)	Emin(6,1)
2	17	30	17
8	16	32	17
12	17	33	18
18	17	36	18
23	17	36	18
28	17	34	18

Ice Sheet Model (75x75x25)

Source: Sandia Trilinos Team

More Information: http://www.fastmath-scidac.org or contact Lori Diachin, LLNL, diachin2@llnl.gov, 925-422-7130