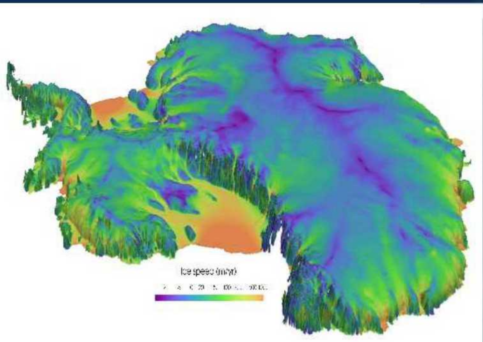
SAND2019-0564PE


Performance Subtask: Update from Sandia National Laboratories

Irina Tezaur and **Jerry Watkins**

<u>*SNL Staff:*</u> L. Bertagna, M. Perego, I. Tezaur (subtask lead), R. Tuminaro, J. Watkins

Sandia National Laboratories, Livermore, CA 94587

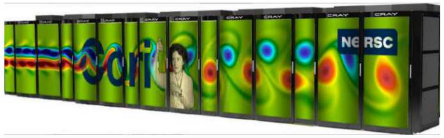
Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Summary of performance subtask

Sub-project focus:

- **Performance** of Albany Land-Ice (ALI) stress-velocity: solver must be scalable, fast, robust.
  - ALI diagnostic solve is comprised of **finite element assembly (FEA)** and **linear solve**. Each takes ~50% of CPU time.
- **Performance portability** of ALI solver stress-velocity: solver must run correctly and efficiently on up-and-coming and future architectures (e.g., multi-core, many-core, GPU, Intel Xeon Phi, etc.)
  - SNL's strategy for performance portability focuses on ***Kokkos***: C++ open-source library that provides performance-portability across diverse devises with different memory models using MPI+X
- Task includes **improved time-stepping methods** for e.g., temp.-velocity coupling in MALI.

Importance/broader ties:



- For initialization, UQ, and predictive dynamic runs, solver needs to run as fast as possible.
- Running on up-and-coming architectures (e.g., Summit, Aurora) is critical.
- ALI is the basis of MPAS-Albany-Land-Ice (MALI) dynamical core.
- Pieces developed (e.g., linear solvers implemented in Trilinos) could be useful to other dycores.

Year 1 and 2 SNL tasks/deliverables

Sandia National Laboratories

Year	Task	Institutions	Joint with subgroup	Status
1	Finish Kokkos ALI FE refactor for velocity; port/optimize for current GPU/Phi	SNL		mostly complete
2	Kokkos ALI FE refactor for temperature & hydrology; port/optimize for current GPU/Phi	SNL		deferred to year 3
1	Deliver robust solvers/preconditioners for temperature model	SNL		in progress
2	Improve time-stepping method for thickness evolution (implicit/obstacle formulation)	SNL, LBNL, LANL	Other development*	deferred to year 3

More Detailed Status

Performance portability via Kokkos [J. Watkins, I. Tezaur]:

- Kokkos refactor/optimization of FEA in stress-velocity solver: mostly complete
- Kokkos refactor of FEA for temperature/hydrology: deferred to year 3

Linear solver performance [R. Tuminaro, I. Tezaur, M. Perego]:

- Robust solvers/preconditioners for temperature model: in progress
- Performance improvements of semi-coarsening-based AMG preconditioner in MueLu for stress-velocity solver (remnant from PISCEES): in progress

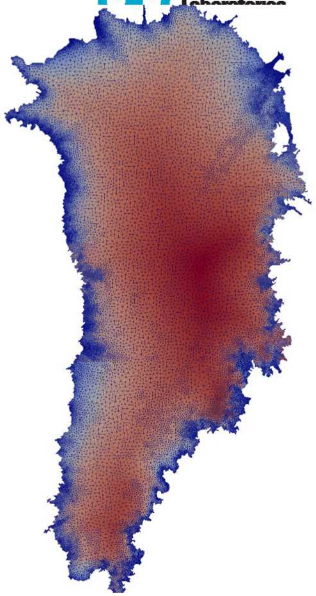
Improved time-stepping method for thickness evolution [M. Perego, L. Bertagna, I. Tezaur, R. Tuminaro]: deferred to year 3

- Infrastructure work in Albany to enable multi-physics/block preconditioning: in progress

* Albany infrastructure work to enable modular multi-physics and block preconditioning (SNL, Yrs 1-2)

Performance portability via Kokkos

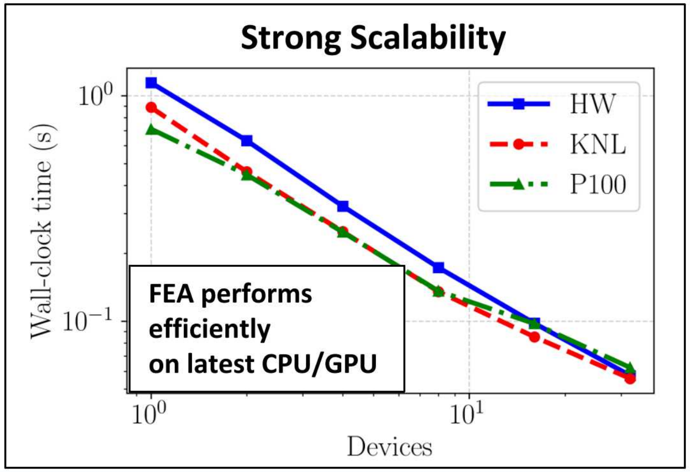
SNL staff: J. Watkins, I. Tezaur

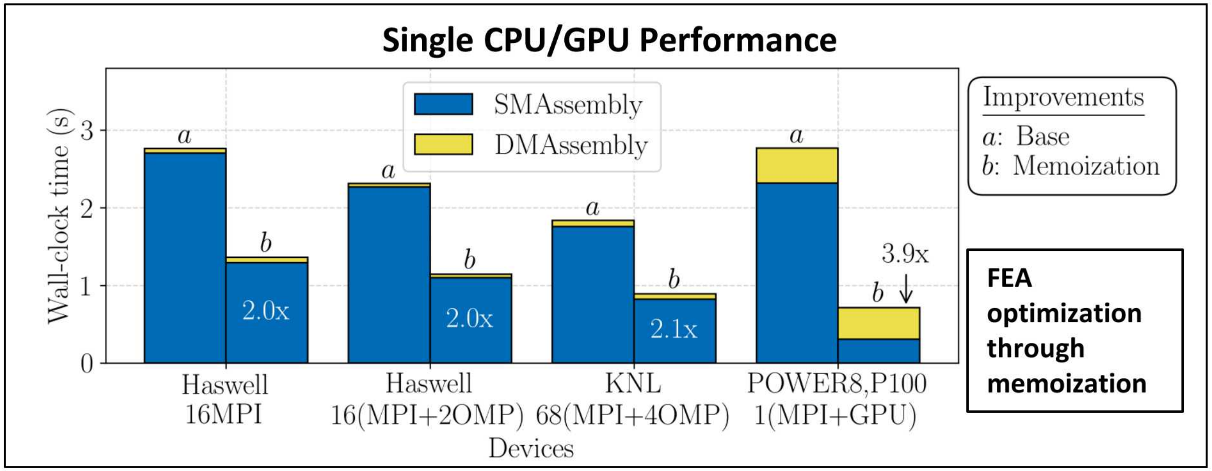


Task 1 [yr 1]: Finish Kokkos ALI FE refactor for velocity; port/optimize for current GPU/Phi
Task 2 [yr 2]: Kokkos ALI FE refactor for temperature & hydrology

Status:

- Task 1 is complete* (see figures) but still needs work:
  - ➢ Boundary conditions need to be refactored to work well on GPUs and future CPUs.
  - ➢ Optimization strategies need to be generalized for more flexibility.
- Task 2 deferred to year 3, as temperature and hydrology solvers are not mature yet.





Performance portability via Kokkos

SNL staff: J. Watkins, I. Tezaur

Task 1 [yr 1]: Finish Kokkos ALI FE refactor for velocity; port/optimize for current GPU/Phi
Task 2 [yr 2]: Kokkos ALI FE refactor for temperature & hydrology

Challenges:

- Task 1: Component based design makes it difficult to refactor/optimize
  - flexibility vs. performance
- Task 2: This task will share similar challenges as task 1

Mitigation strategy/next steps:

- Task 1: Currently testing different strategies on miniapp (faster design cycle) to achieve good performance while still maintaining some flexibility
- Task 2: Performance optimization of the temperature and hydrology solvers will benefit from the optimization approach adopted for the velocity solver

Looking forward:

- Task 3 [yr 3]: Port/tailor ALI code performance to Summit (GPU), Aurora (Xeon)
  - Tpetra – performance portable global matrix assembly (WIP – Q1, FY19)
  - Aurora shifted to 2021; hierarchical parallelism required for strong scaling
  - Trilinos – simd support for CPU optimization (WIP – Q3, FY19)

Nightly testing

SNL staff: I. Tezaur

Performance monitored via **CDash nightly testing** on a variety of architectures.

Sandia National Laboratories

Login All Dashboards

Albany

Dashboard Calendar Previous Current Next Project

Project

Project	Configure Error	Configure Warning	Configure Pass	Build Error	Build Warning	Build Pass	Test Not Run	Test Fail	Test Pass
Albany	0	18	24	1	20	4	0	6	3201

SubProjects

Project	Configure Error	Configure Warning	Configure Pass	Build Error	Build Warning	Build Pass	Test Not Run	Test Fail	Test Pass
Peridigm	0	0	1	1	1	0			
TrilinosIntel	0	1	1	0	1	0			
AlbanyIntel	0	0	1	0	0	1	0	1	347
IKTCismAlbany	0	1	1	0	1	0	0	0	5
IKTCismAlbanyEpetra	0	1	1	0	1	0	0	0	5
IKTAlbanyFunctorOpenMP	0	1	1	0	1	0	0	0	278
Trilinos	0	1	1	0	1	0			
TrilinosClang	0	1	1	0	1	0			
Albany64BitClang	0	0	1	0	1	0	0	1	166
IKTRideTrilinosCUDA	0	1	1	0	1	0			
IKTRideAlbanyCUDA	0	2	2	0	2	0	0	2	186
albany_cluster-toss3_skybridge-login5_serial-intel-release	0	1	1	0	0	1	0	1	288
Albany64Bit	0	0	1	0	1	0	0	1	317
IKTAlbany	0	1	1	0	1	0	0	0	372
IKTAlbanyNoEpetra	0	1	1	0	1	0	0	0	261
TrilinosDbg	0	1	1	0	1	0			
Albany64BitDbg	0	0	1	0	0	1	0	0	225
trilinos_cluster-toss3_skybridge-login5_serial-intel-release	0	1	1	0	1	0			
IKTMayerARMTrilinos	0	1	1	0	1	0			
IKTMayerARMAlbany	0	0	1	0	0	1	0	0	294
IKTWatermanTrilinosCUDA	0	1	1	0	1	0			
IKTWatermanAlbanyCUDA	0	1	1	0	1	0	0	0	94
IKTAlbanyFPECheckDbg	0	1	1	0	1	0	0	0	363

P100 GPU

ARM

V100 GPU

Albany

Dashboard Calendar Previous Current Next Project

Project

Project	Configure Error	Configure Warning	Configure Pass	Build Error	Build Warning	Build Pass	Test Not Run	Test Fail	Test Pass
Albany	0	3	0	0	2	1	0	0	608

SubProjects

SubProject	Configure Error	Configure Warning	Configure Pass	Build Error	Build Warning	Build Pass	Test Not Run	Test Fail	Test Pass
EdisonTrilinos	0	1	0	0	1	0			
CoriTrilinos	0	1	0	0	1	0			
CoriAlbanyFELIX	0	1	0	0	1	0			
CoriCismAlbany	0	1	0	0	1	0			
EdisonAlbanyFELIX	0	1	0	0	1	0			
EdisonCismAlbany	0	1	0	0	1	0			
IKTBlakeTrilinosSerial	0	1	0	0	0	1			
IKTBlakeAlbanySerial	0	0	1	0	0	1	0	0	316
IKTBlakeTrilinosOpenMP	0	1	0	0	0	1			
IKTBlakeAlbanyOpenMP	0	0	1	0	0	1	0	0	292

NERSC (Cori, Edison)

Skylake

Nightly testing

SNL staff: I. Tezaur

Performance monitored via **CDash nightly testing** on a variety of architectures.

Sandia National Laboratories

Login All Dashboards

Albany — Dashboard | Calendar | Previous | Current | Next | Project

Project

Project	Configure Error	Configure Warning	Configure Pass	Build Error	Build Warning	Build Pass	Test Not Run	Test Fail	Test Pass
Albany	0	18	24	1	20	4	0	6	3201

SubProjects

Project	Configure Error	Configure Warning	Configure Pass	Build Error	Build Warning	Build Pass	Test Not Run	Test Fail	Test Pass
Peridigm	0	0	1	1	1	0			
TrilinosIntel	0	1	1	0	1	0			
AlbanyIntel	0	0	1	0	0	1	0	1	347
IKTCismAlbany	0	1	1	0	1	0	0	0	5
IKTCismAlbanyEpetra	0	1	1	0	1	0	0	0	5
IKTAlbanyFunctorOpenMP	0	1	1	0	1	0	0	0	278
Trilinos	0	1	1	0	1	0			
TrilinosClang	0	1	1	0	1	0			
Albany64BitClang	0	0	1	0	1	0	0	1	166
IKTRideTrilinosCUDA	0	1	1	0	1	0			
IKTRideAlbanyCUDA	0	2	2	0	2	0	0	2	186
albany_cluster-toss3_skybridge-login5_serial-intel-release	0	1	1	0	0	1	0	1	288
Albany64Bit	0	0	1	0	1	0	0	1	317
IKTAlbany	0	1	1	0	1	0	0	0	372
IKTAlbanyNoEpetra	0	1	1	0	1	0	0	0	261
TrilinosDbg	0	1	1	0	1	0			
Albany64BitDbg	0	0	1	0	0	1	0	0	225
trilinos_cluster-toss3_skybridge-login5_serial-intel-release	0	1	1	0	1	0			
IKTMayerARMTrilinos	0	1	1	0	1	0			
IKTMayerARMAlbany	0	0	1	0	0	1	0	0	294
IKTWatermanTrilinosCUDA	0	1	1	0	1	0			
IKTWatermanAlbanyCUDA	0	1	1	0	1	0			
IKTAlbanyFPECheckDbg	0	1	1	0	1	0			

P100 GPU (IKTRideTrilinosCUDA, IKTRideAlbanyCUDA)

ARM (IKTMayerARMTrilinos, IKTMayerARMAlbany)

V100 GPU (IKTWatermanTrilinosCUDA, IKTWatermanAlbanyCUDA)

Albany — Dashboard | Calendar | Previous | Current | Next | Project

Project

Project	Configure Error	Configure Warning	Configure Pass	Build Error	Build Warning	Build Pass
Albany	0	3	0	0	2	1

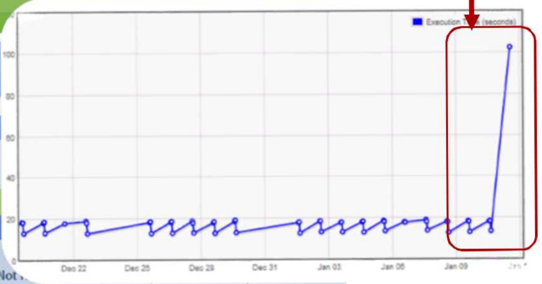
SubProjects

SubProject	Configure Error	Configure Warning	Configure Pass	Build Error	Build Warning	Build Pass	Not …		
EdisonTrilinos	0	1	0	0	1	0			
CoriTrilinos	0	1	0	0	1	0			
CoriAlbanyFELIX	0	1	0	0	1	0			
CoriCismAlbany	0	1	0	0	1	0			
EdisonAlbanyFELIX	0	1	0	0	1	0			
EdisonCismAlbany	0	1	0	0	1	0			
IKTBlakeTrilinosSerial	0	1	0	0	0	1			
IKTBlakeAlbanySerial	0	0	1	0	0	1	0	0	316
IKTBlakeTrilinosOpenMP	0	1	0	0	0	1			
IKTBlakeAlbanyOpenMP	0	0	1	0	0	1	0	0	292

NERSC (Cori, Edison) (EdisonTrilinos … EdisonCismAlbany)

Skylake (IKTBlakeTrilinosSerial … IKTBlakeAlbanyOpenMP)

Slowdown identified by nightly CDash testing!

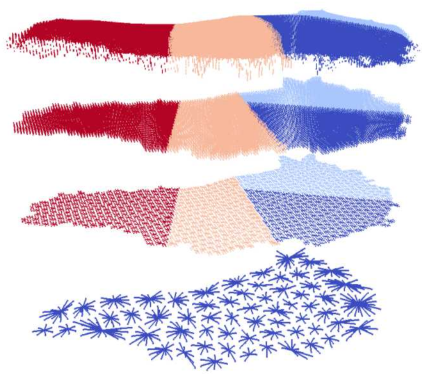


Linear solver performance

SNL staff: R. Tuminaro, I. Tezaur, M. Perego

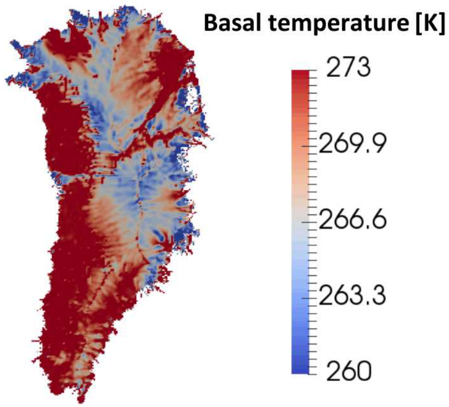
Task 1 [yr 1]: deliver robust solvers/preconditioners for temperature model in ALI.

Task 2 [from PISCEES]: for FO-Stokes solver, improve performance of MueLu implementation of aggressive-semi-coarsening-based AMG* preconditioner.



Status:

- Task 1 deferred to year 2 b/c Ray was unavailable in year 1
  - Some progress made on nonlinear solvers for temperature equation last year (presented at AGU 2017).
- Task 2 replaced Task 1 for year 1, and is in progress:
  - Inconsistencies b/w ML and MueLu have been resolved.
  - Seg faults/FPEs/overflow in MueLu have been resolved.
  - Total # of iterations b/w ML and MueLu are similar though not identical (right).
  - MueLu still runs ~2× slower than ML (right) – culprit has been identified as banded relaxation line-smoothing in Ifpack2.



2km AIS/ 20 layers	ML	MueLu
Total Time (s)	413.3	513.6
Total Linear Solve Time (s)	97.1	190.6
# linear iters	426	400

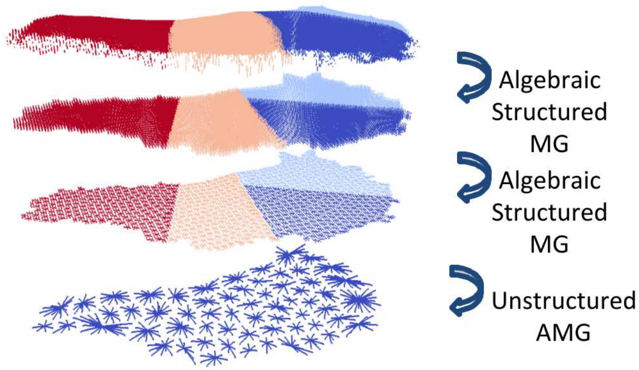
* R. Tuminaro, et al. "A matrix dependent/algebraic multigrid approach for extruded meshes with applications to ice sheet modeling", *SISC* 2016.

Linear solver performance

SNL staff: R. Tuminaro, I. Tezaur, M. Perego

Task 1 [yr 1]: deliver robust solvers/preconditioners for temperature model in ALI.

Task 2 [from PISCEES]: improve performance of MueLu implementation of aggressive-semi-coarsening-based AMG* preconditioner to make it comparable to ML.



Challenges:

- Task 1: highly nonlinear and advection-dominated problems.
- Task 2: currently lacking manpower to speed-up relevant routine in Ifpack2 (banded relaxation line-smoothing).

Mitigation strategy/next steps:

- Task 1: Ray + Mauro will start on this in FY19.
- Task 2: currently consulting with Ifpack2/MueLu developers; hire summer student to work on Ifpack2 improvements?

Looking forward:

- Task 3 [yr 4]: study of Trilinos linear solvers performance on next generation architectures for ALI
  - ➢ Ifpack2 and MueLu are designed for performance portability (integrate Kokkos), and are being matured for up-and-coming architectures by Trilinos team.

* R. Tuminaro, et al. "A matrix dependent/algebraic multigrid approach for extruded meshes with applications to ice sheet modeling", *SISC* 2016.

Improved time-stepping method for thickness evolution

SNL staff: L. Bertagna, M. Perego, I. Tezaur, R. Tuminaro

Task 1 [yr 2]: Improve time-stepping method for thickness evolution (implicit/obstacle formulation)

Status: deferred to year 2.

- We decided to focus on infrastructure (block data structure) work that is needed by the thickness evolution equation and temperature/hydrology couplings with FO solver
  - This will enable a variety of couplings and block solvers (Teko) for improved performance.

Challenges:

- Hard highly nonlinear free-surface problem on fixed mesh
- Albany is a complex code to refactor; we have limited manpower.

Mitigation strategy/next steps:

- Hiring post-doc to help with the task

Looking forward:

- Task 2 [yr 3]: ALI robust solvers/preconditioners for coupled thermo-mechanical model; scalability studies
  - Block data structures will enable block solvers for improved performance

Summary/looking forward

Year	Task	Institutions	Joint with subgroup	Status
1	Finish Kokkos ALI FE refactor for velocity; port/optimize for current GPU/Phi	SNL		mostly complete
2	Kokkos ALI FE refactor for temperature & hydrology; port/optimize for current GPU/Phi	SNL		deferred to year 3
1	Deliver robust solvers/preconditioners for temperature model	SNL		in progress
2	Improve time-stepping method for thickness evolution (implicit/obstacle formulation)	SNL, LBNL, LANL	Other development*	deferred to year 3
3	ALI: robust solvers/preconditioners for coupled thermo-mechanical model; scalability studies	SNL		
3	Port & tailor ALI code performance to Aurora (Phi) & Summit (GPU) machines	SNL		
4	Study of Trilinos linear solvers performance on next generation architectures for ALI	SNL		
4	Numerical convergence study of ice evolution on simplified experiments	SNL, LBNL		
5	Kokkos refactor of remaining ALI evaluators; & optimize for new state-of-the-art HPC	SNL		

* Albany infrastructure work to enable modular multi-physics and block preconditioning (SNL, Yrs 1-2)

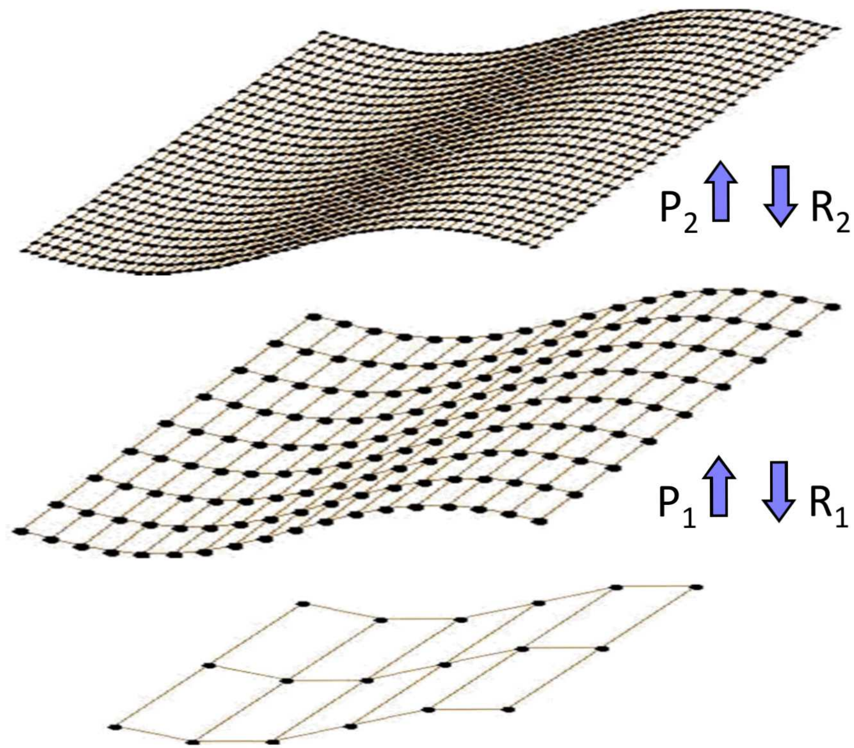
Start of Backup Slides

How Does Multi-Grid Work?

Solve $A_3 u_3 = f_3$

Basic idea:

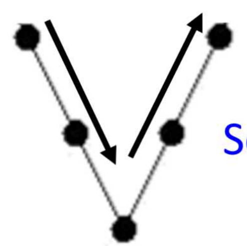
- Develop coarse approximations
- Accelerate convergence via coarse iterations to efficiently propagate information



Smooth $A_3 u_3 = f_3$. Set $f_2 = R_2 r_3$.

Smooth $A_2 u_2 = f_2$. Set $f_1 = R_1 r_2$.

Solve $A_1 u_1 = f_1$ directly.

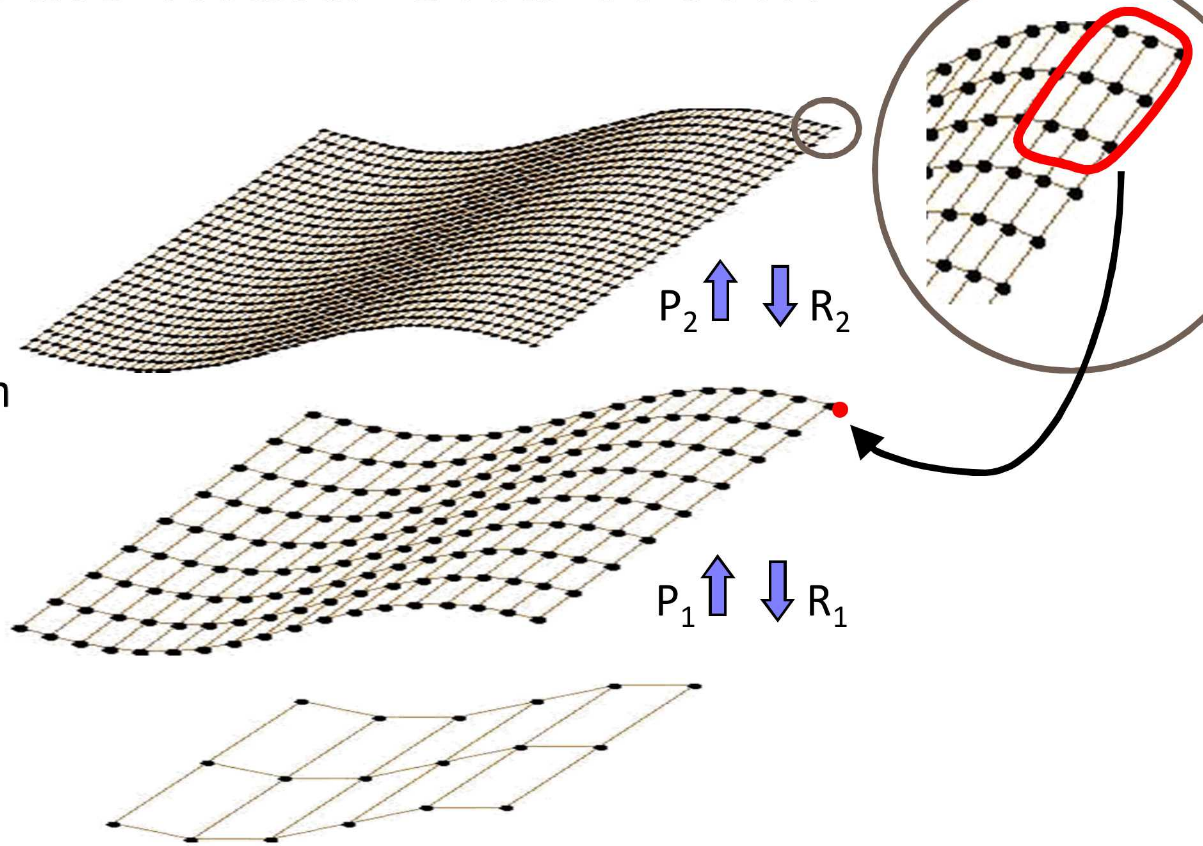


Set $u_3 = u_3 + P_2 u_2$. Smooth $A_3 u_3 = f_3$.

Set $u_2 = u_2 + P_1 u_1$. Smooth $A_2 u_2 = f_2$.

How Does *Algebraic* Multi-Grid Work?

Solve $A_3 u_3 = f_3$



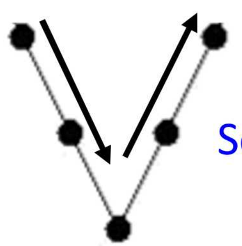
Basic idea:

- Construct graph & coarsen
- Determine P_i & R_i sparsity pattern
- Determine P_i & R_i's coeffs
- Project: $A_i = R_i A_{i+1} P_i$

Smooth $A_3 u_3 = f_3$. Set $f_2 = R_2 r_3$.

Smooth $A_2 u_2 = f_2$. Set $f_1 = R_1 r_2$.

Solve $A_1 u_1 = f_1$ directly.



Set $u_3 = u_3 + P_2 u_2$. Smooth $A_3 u_3 = f_3$.

Set $u_2 = u_2 + P_1 u_1$. Smooth $A_2 u_2 = f_2$.

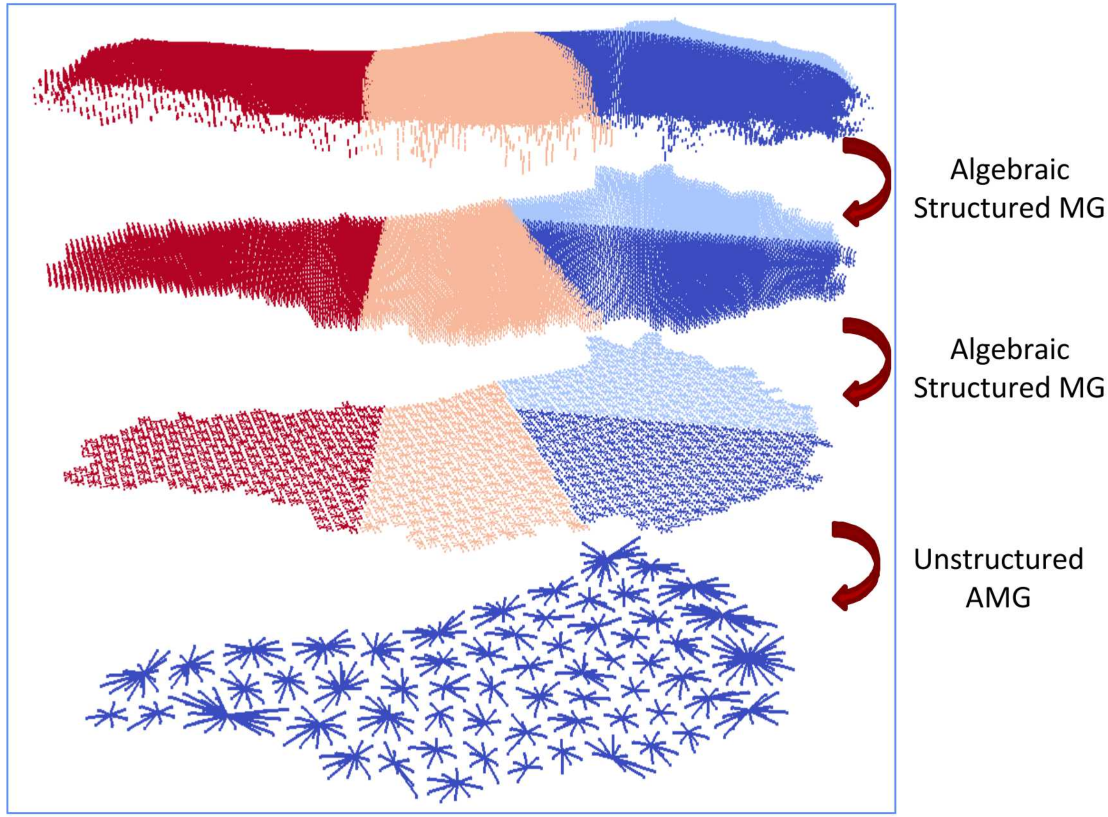
Scalability via Algebraic Multi-Grid Preconditioning with Semi-Coarsening

Bad aspect ratios ($dx \gg dz$) ruin classical AMG convergence rates!

- relatively small horizontal coupling terms, hard to smooth horizontal errors

$\Rightarrow$ Solvers (AMG and ILU) must take aspect ratios into account

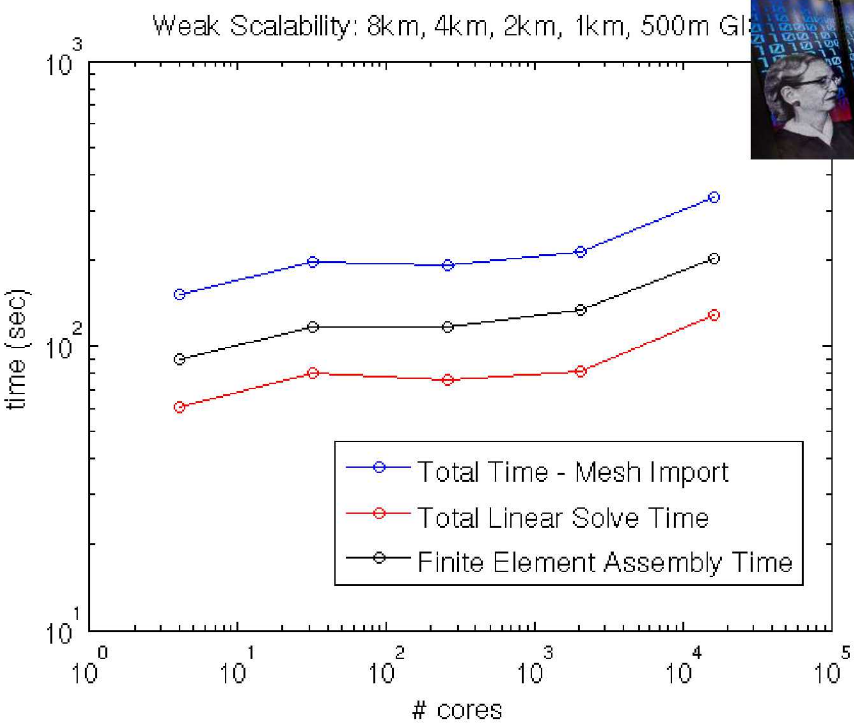
We developed a **new AMG solver** based on aggressive **semi-coarsening** (available in *ML/MueLu* packages of *Trilinos*)



See (Tezaur *et al.*, 2015),
(Tuminaro *et al.*, 2016).

Scaling studies (next slides):
New AMG preconditioner vs. ILU

Greenland Controlled Weak Scalability Study



- Weak scaling study with fixed dataset, 4 mesh bisections.
- ~70-80K dofs/core.
- ***Conjugate Gradient (CG) iterative method*** for linear solves (faster convergence than GMRES).
- ***New AMG preconditioner*** developed by R. Tuminaro based on ***semi-coarsening*** (coarsening in z-direction only).
- ***Significant improvement*** in scalability with new AMG preconditioner over ILU preconditioner!

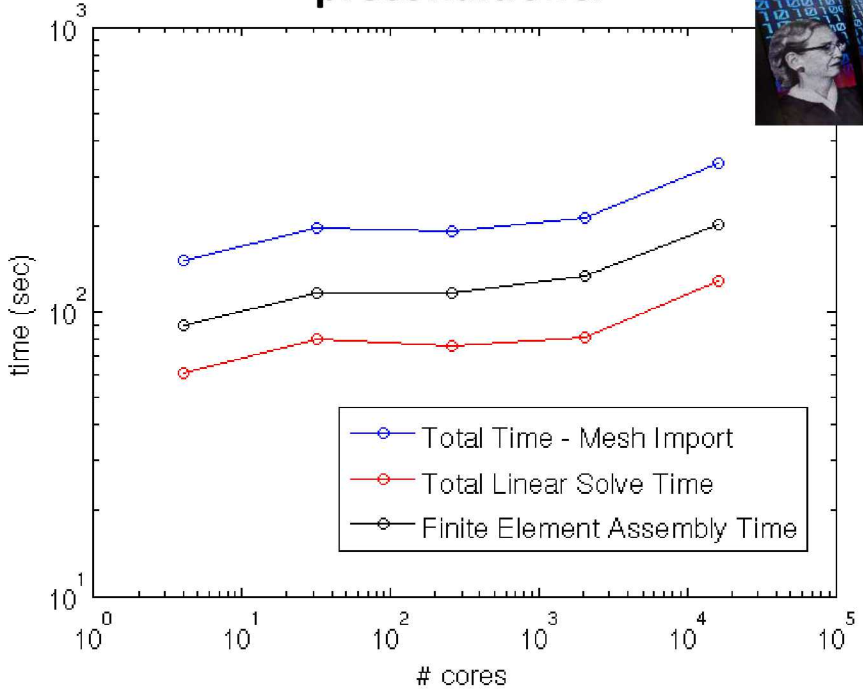
4 cores
334K dofs
8 km Greenland,
5 vertical layers

→ $\times 8^4$ scale up

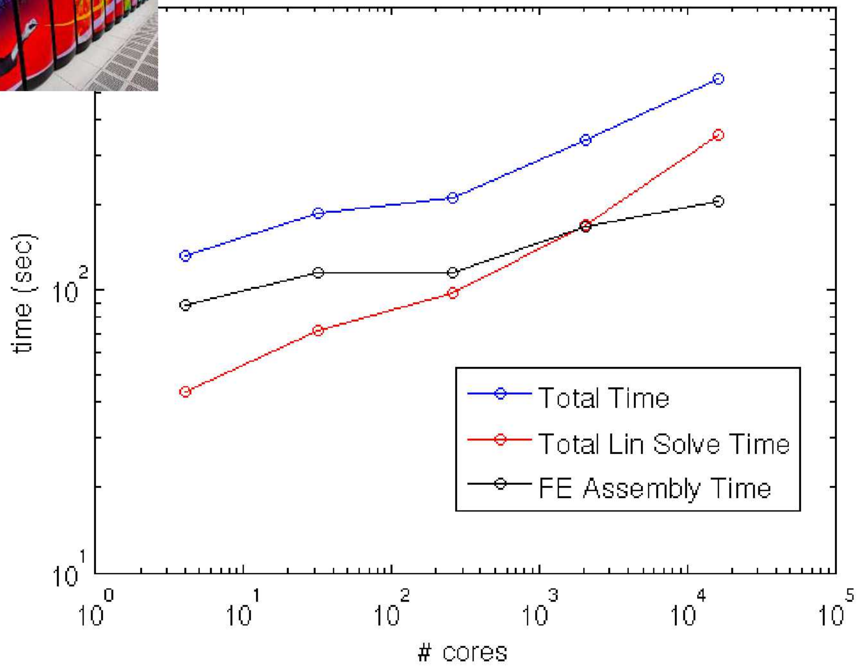
16,384 cores
1.12B dofs(!)
0.5 km Greenland,
80 vertical layers

Greenland Controlled Weak Scalability Study

New AMG preconditioner preconditioner



ILU preconditioner



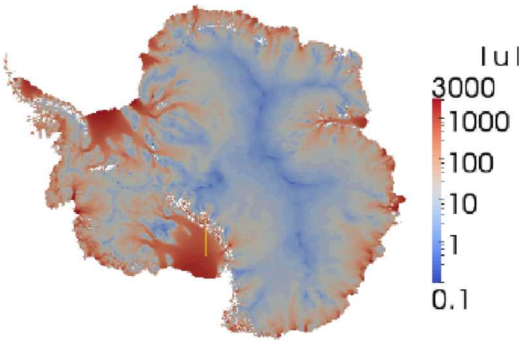
4 cores
334K dofs
8 km Greenland,
5 vertical layers

→ $\times 8^4$ scale up →

16,384 cores
1.12B dofs(!)
0.5 km Greenland,
80 vertical layers

- ***Significant improvement*** in scalability with new AMG preconditioner over ILU preconditioner!

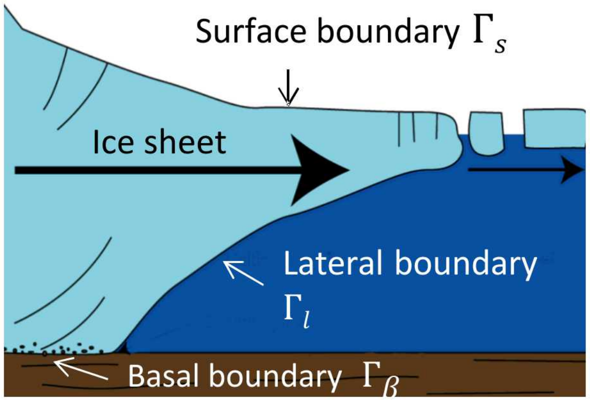
Moderate Resolution Antarctica Weak Scaling Study



Antarctica is fundamentally different than Greenland:
AIS contains large ice shelves (floating extensions of land ice).

- ***Along ice shelf front:*** open-ocean BC (Neumann).
- ***Along ice shelf base:*** zero traction BC (Neumann).

⇒ For vertical grid lines that lie within ice shelves, top and bottom BCs resemble Neumann BCs so sub-matrix associated with one of these lines is almost* singular.

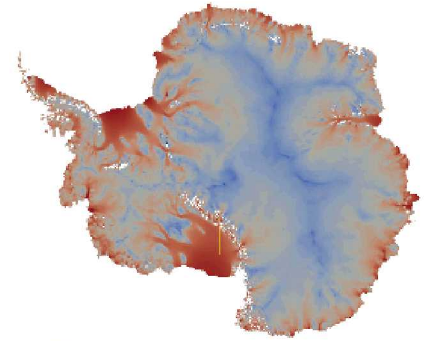


(vertical > horizontal coupling)
+
Neumann BCs
=
nearly singular submatrix associated with vertical lines

⇒ Ice shelves give rise to severe ill-conditioning of linear systems!

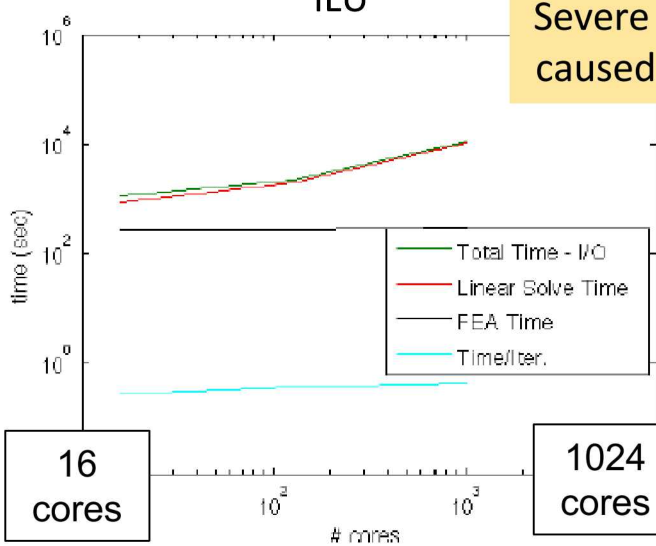
*Completely singular in the presence of islands and some ice tongues.

Moderate Resolution Antarctica Weak Scaling Study



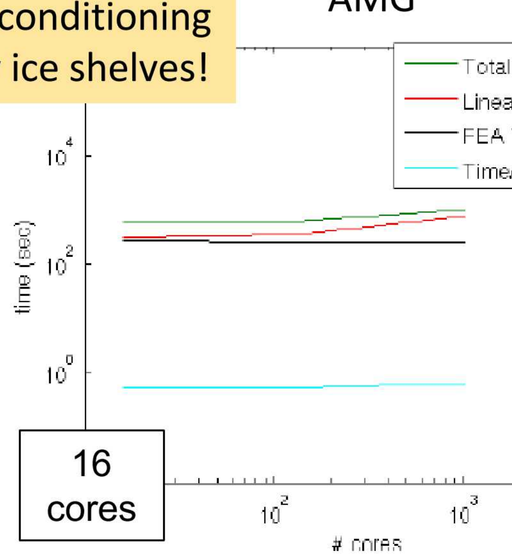
- Weak scaling study on Antarctic problem (8km w/ 5 layers → 2km w/ 20 layers).
- Initialized with realistic basal friction (from deterministic inversion) and temperature field from BEDMAP2.
- **Iterative linear solver**: GMRES.
- **Preconditioner**: ILU vs. new AMG based on aggressive semi-coarsening.

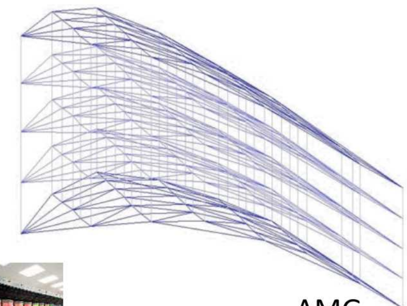
ILU



Severe ill-conditioning caused by ice shelves!

AMG





AMG preconditioner

(vertical > horizontal coupling)
+
Neumann BCs
=
nearly singular submatrix associated with vertical lines

cores   # cores

AMG preconditioner less sensitive than ILU to ill-conditioning (ice shelves → Green's function* with modest horizontal decay → ILU is less effective).

* Tuminaro *et al.*, *SISC*, 2016.