Memory Bandwidth vs FLOPs - Tradeoffs on Modern Hardware

Jeremy L Thompson

University of Colorado Boulder

jeremy@jeremylt.org

30 March 2026

Interrupt Me!

Please interrupt with questions/comments

Overview

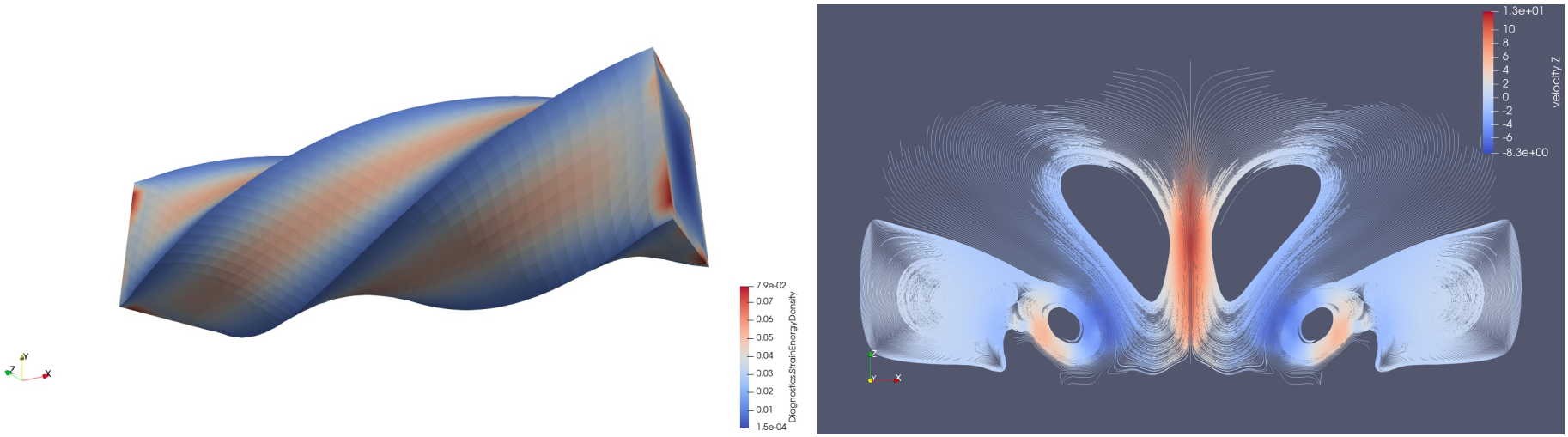
1. End Goal
2. Hardware
3. Limitation
4. Ratel
5. Questions

libCEED Team

libCEED Repo: https://github.com/CEED/libCEED

Developers: Ahmad Abdelfattah, Zach R. Atkins, Valeria Barra, Natalie Beams, Jed Brown, Jean-Sylvain Camier, Veselin Dobrev, Yohann Dudouit, Leila Ghaffari, Sebastian Grimberg, Tzanio Kolev, David Medina, Will Paznel, Thilina Ratnayaka, Rezgar Shakeri, Stan Tomov, James Wright III, Jeremy L Thompson

Background



libCEED solid mechanics (left) and fluid dynamics (right) mini-apps

- Physics based simulations important in science/engineering
- Intuition: FEM solves equations with piecewise polynomial solution
- libCEED supports FEM-like simulations on modern hardware

libCEED Projects

Several projects built using libCEED

- Ratel - solid mechanics FEM (H1) and iMPM (PSAAP)
- HONEE - fluid dynamics FEM (H1) & differential filtering (PHASTA)
- MFEM - various applications, libCEED integrators (LLNL)
- Palace - Electromagnetics FEM with MFEM + libCEED (Amazon)
  H(div) and H(curl) elements
- RDycore - FV river dynamical core with PETSc + libCEED (SciDAC)

Large Simulations

Large physics based simulations require lots of hardware

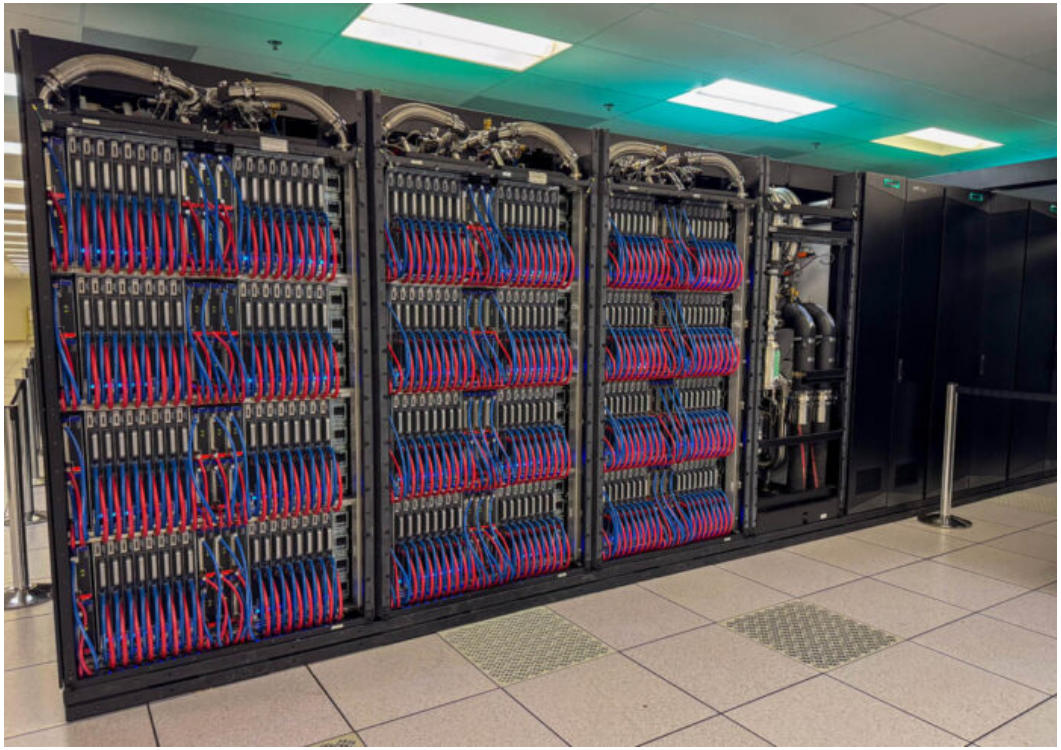
NCAR's weather models as a great example

High continuous demand for computational resources

Supercomputer?

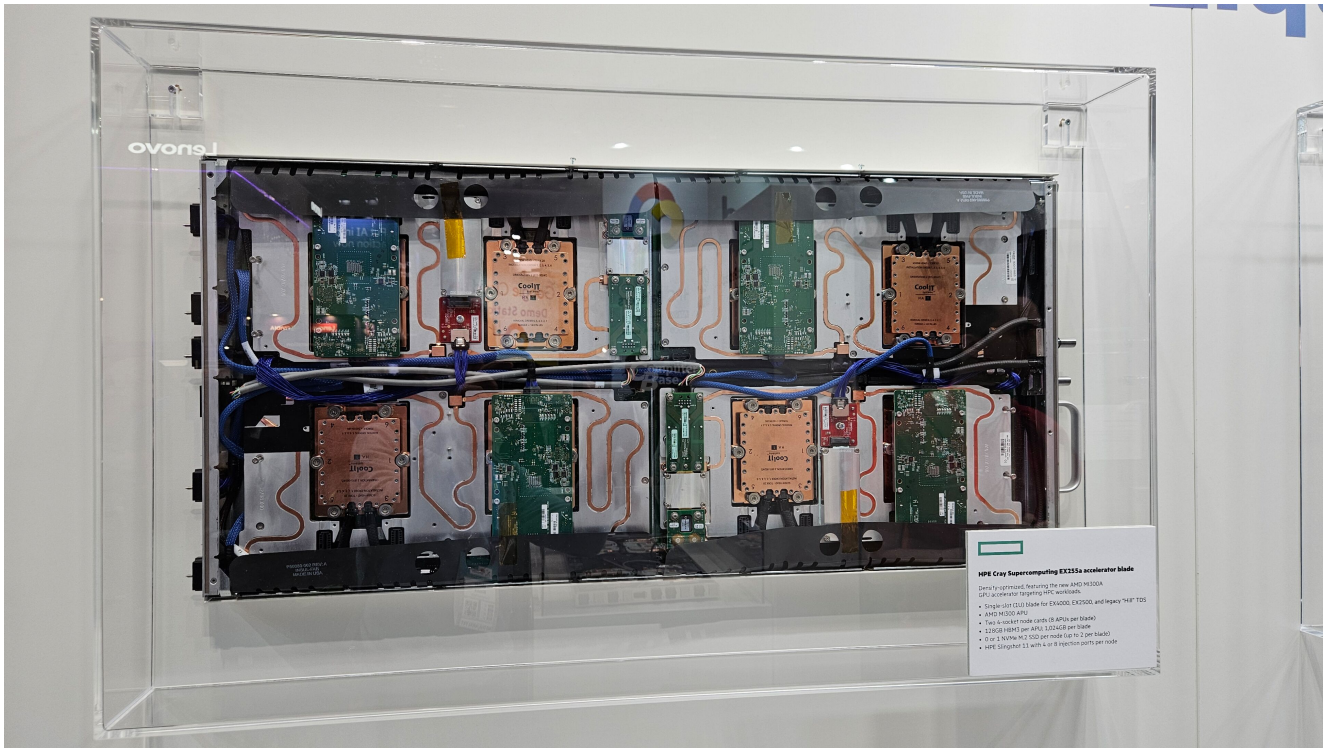
El Capitan - current #1 on Top 500 chart

Supercomputer?



The entire machine is networked 'blades' sitting in 'racks'

Supercomputer?



Each 'blade' consists of multiple CPU/GPUs that do the computation

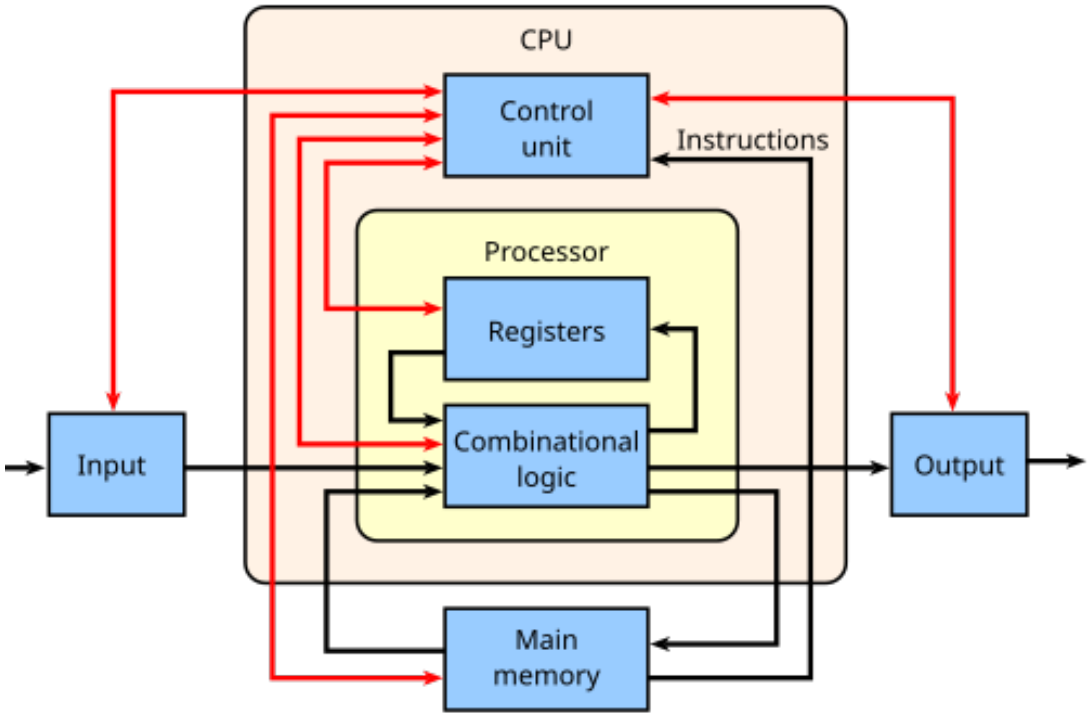
Top 500

Machine	HPL	HPCG
Fugaku	442.01 PFLOPs	16.01 PFLOPs
Frontier	1,353.00 PFLOPs	14.05 PFLOPs
Aurora	1,012.00 PFLOPs	5.61 PFLOPs
LUMI	379.70 PFLOPs	4.59 PFLOPs
Alps	434.90 PFLOPs	3.57 PFLOPs

Top 500 Machines for HPCG with HPL peak FLOPs

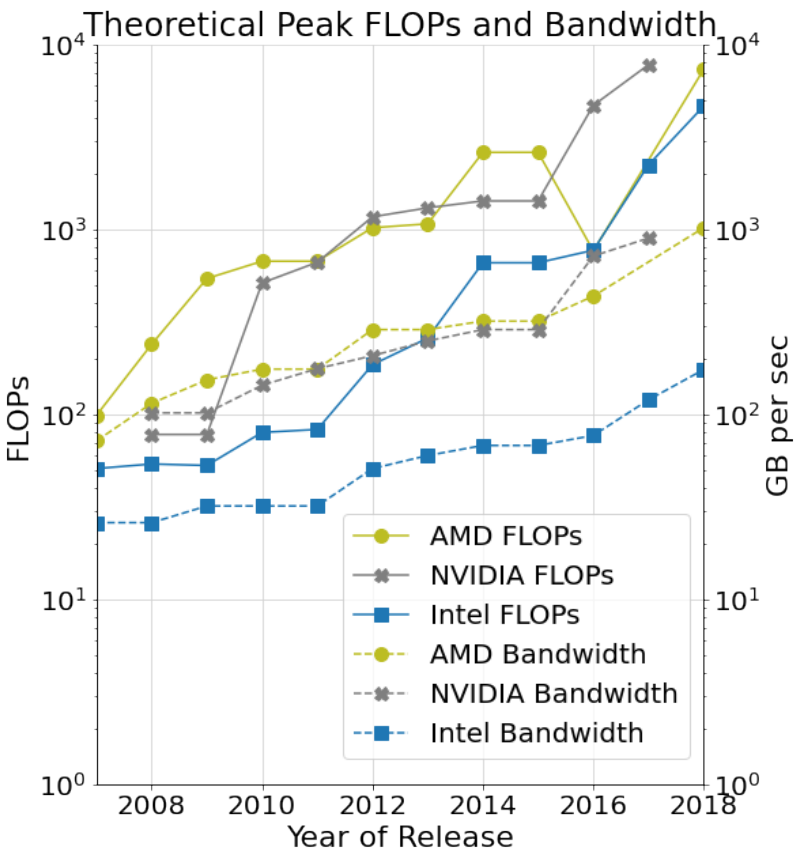
Difficult to realize peak FLOPs with CG on modern machines

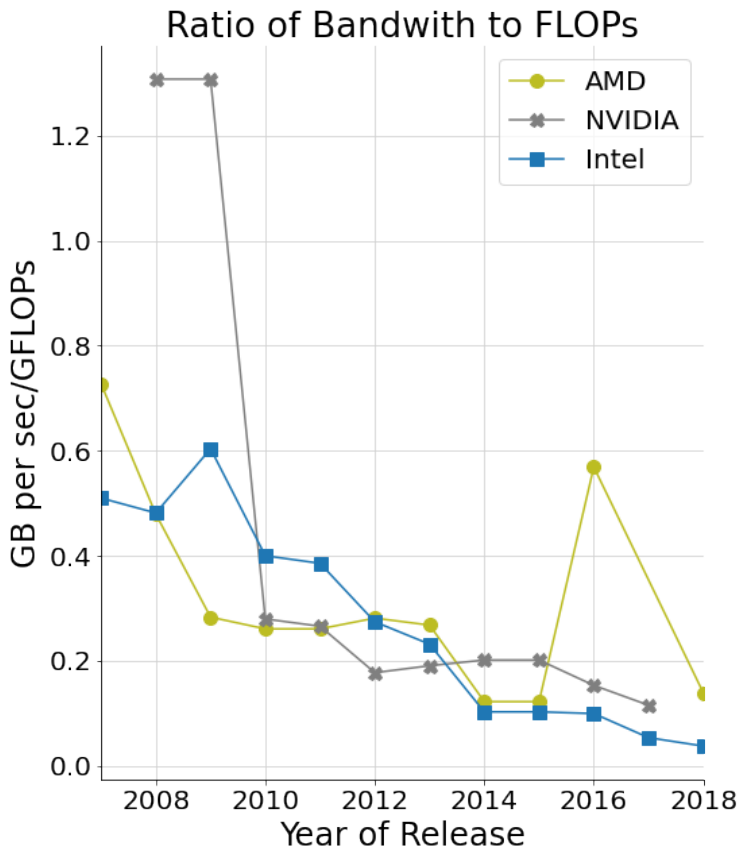
Two Constraints



Two big constraints - moving data to the processor & processing the data

Modern Hardware



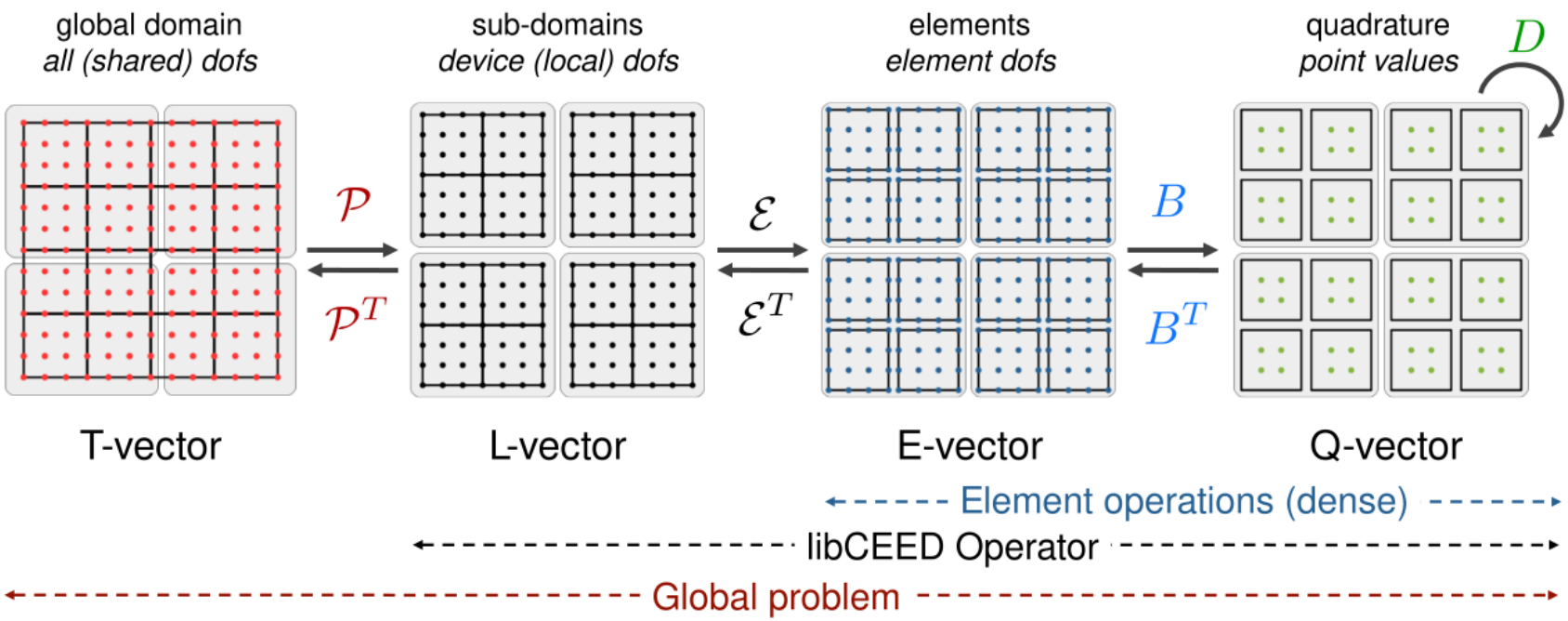


Memory bandwidth is improving slower than FLOPs

Mirrors difference between Top 500 HPL vs HPCG benchmarks

Matrix-Free Operators from libCEED

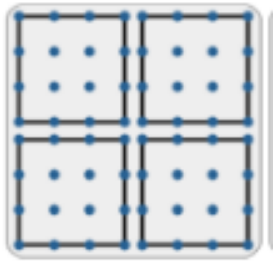
$$A = \mathcal{P}^T \mathcal{E}^T B^T D B \mathcal{E} \mathcal{P}$$



libCEED provides matrix-free operator evaluation on various hardware

Matrix-free operators apply these steps instead of populating a matrix

Matrix-Free?



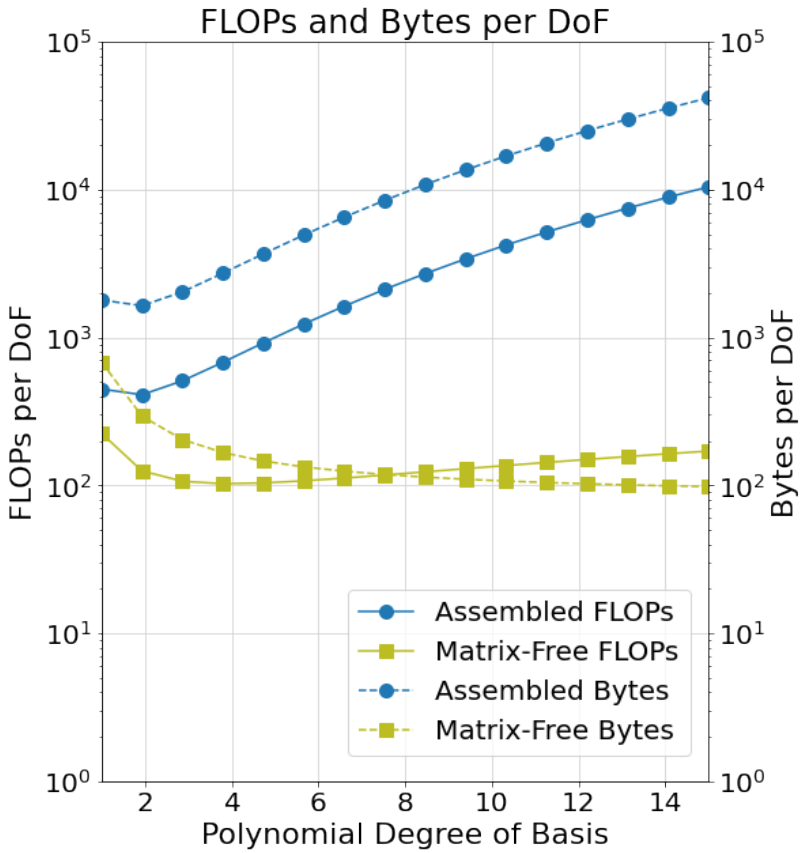
Basis of each element roughly of size $(n\,1D\ DoFs)^2$ ($(n\,1D\ DoFs)^3$ in 3D)

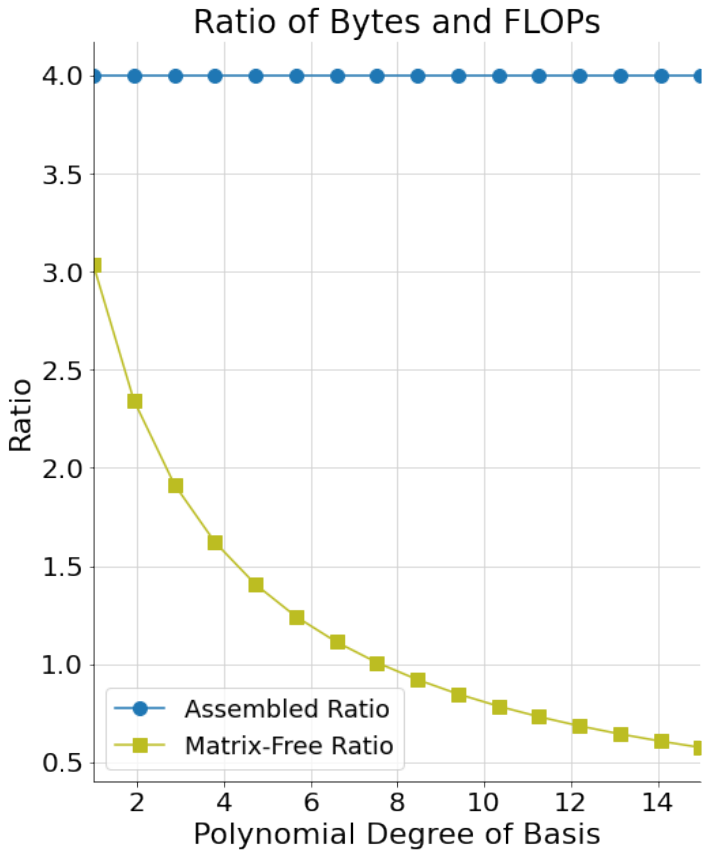
4^2 (or 4^3) in the picture

For 'tensor product' elements this drops to $(n\,1D\ DoFs)$ or 4

Calculation is more complicated factoring in all data movement

Benefits of Matrix-Free

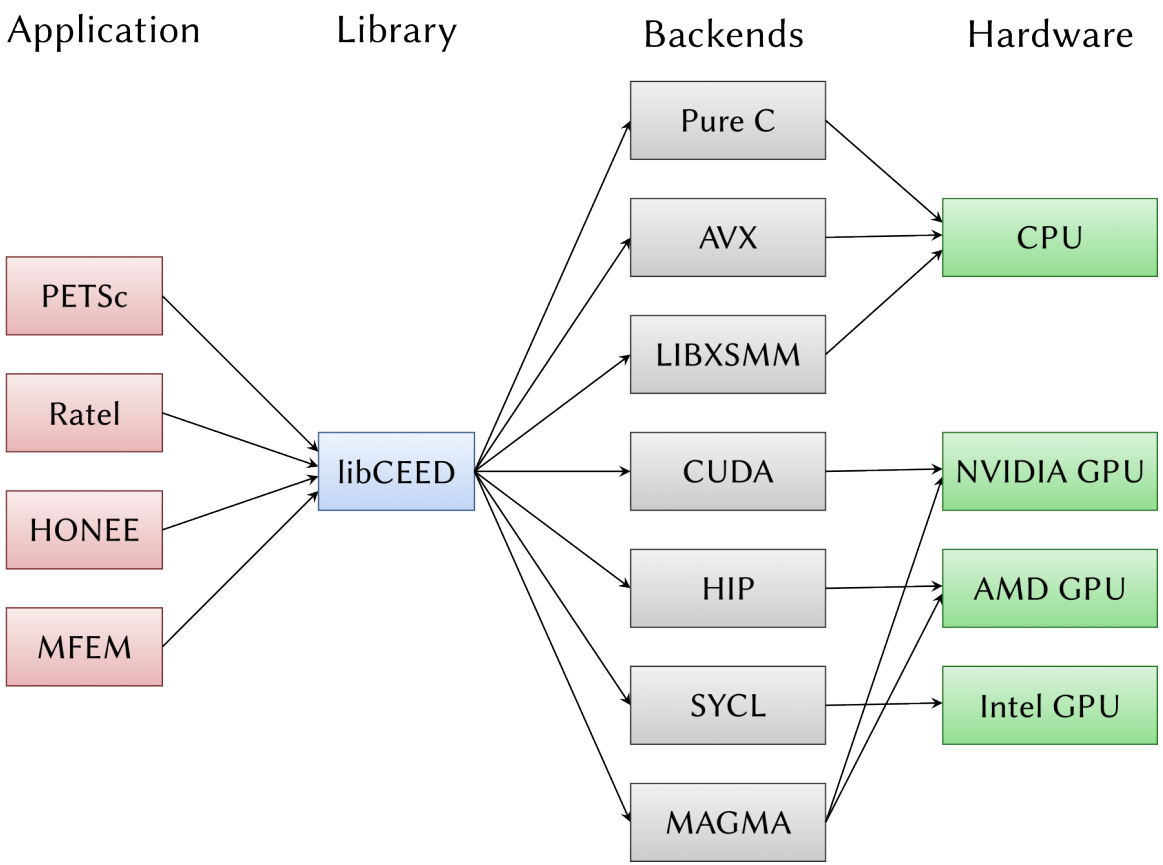




Requirements for matrix-vector product with sparse matrix vs matrix-free for screened Poisson $\nabla^2 u - \alpha^2 u = f$ in 3D

Matrix-free representations using tensor product bases better match modern hardware

Performance Portability



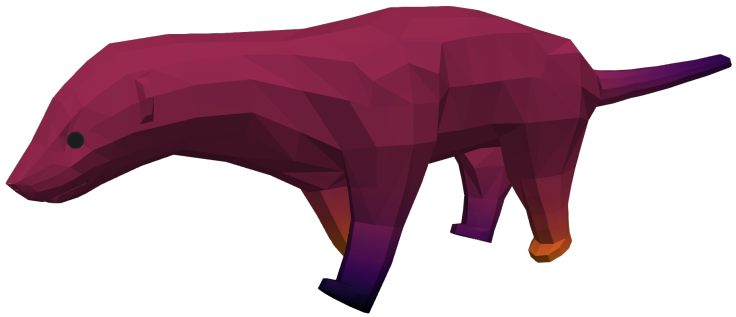
libCEED's design naturally allows multiple hardware implementations

Design Implications

Using matrix-free operators drives design decisions

- Direct solvers are out (require assembled matrix)
- Iterative solvers are in (Krylov methods, etc)
- High order = high accuracy & bad condition numbers
- Preconditioning is needed for fast convergence

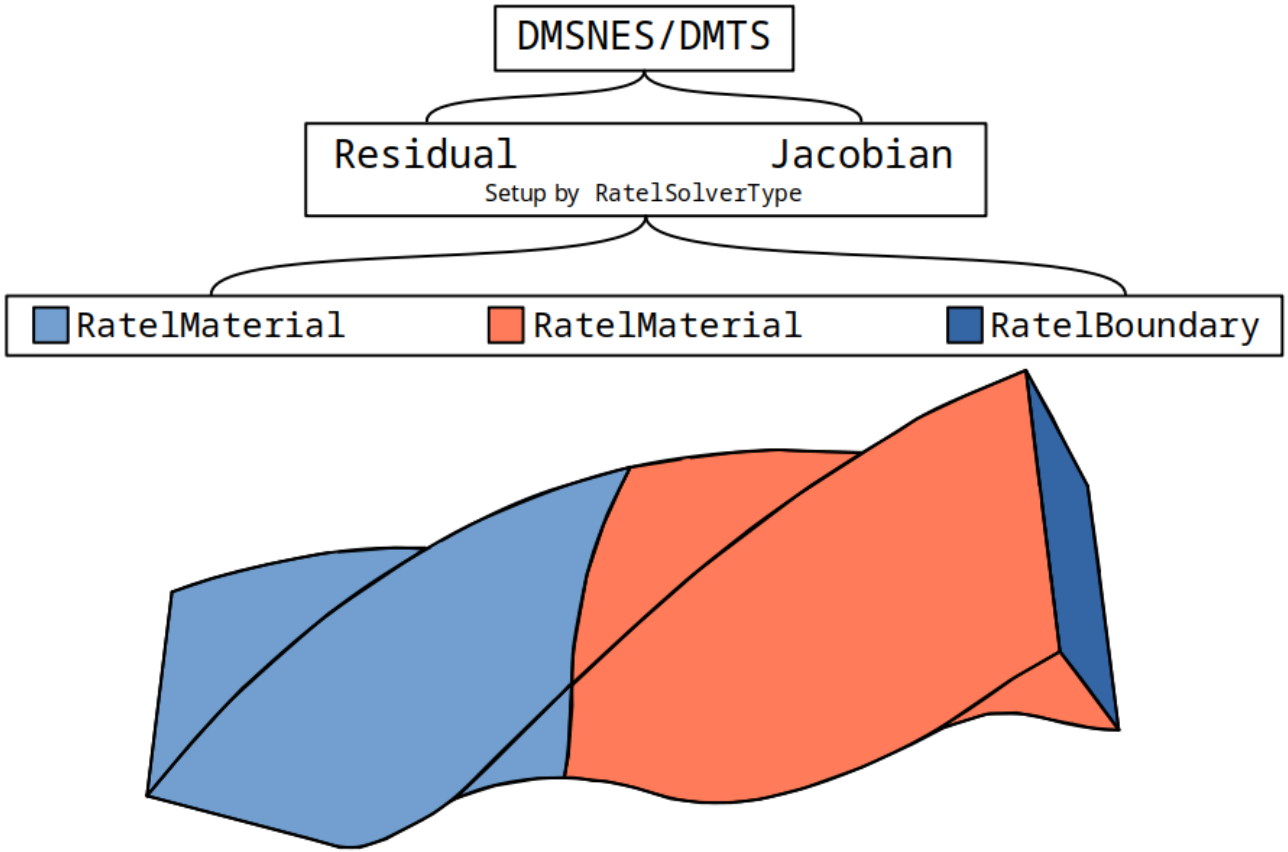
Ratel Team



Ratel Repo: https://gitlab.com/micromorph/ratel

Developers: Zach R. Atkins, Jed Brown, Fabio Di Gioacchino, Leila Ghaffari, Zach Irwin, Rezgar Shakeri, Ren Stengel, Jeremy L Thompson

Basic Design

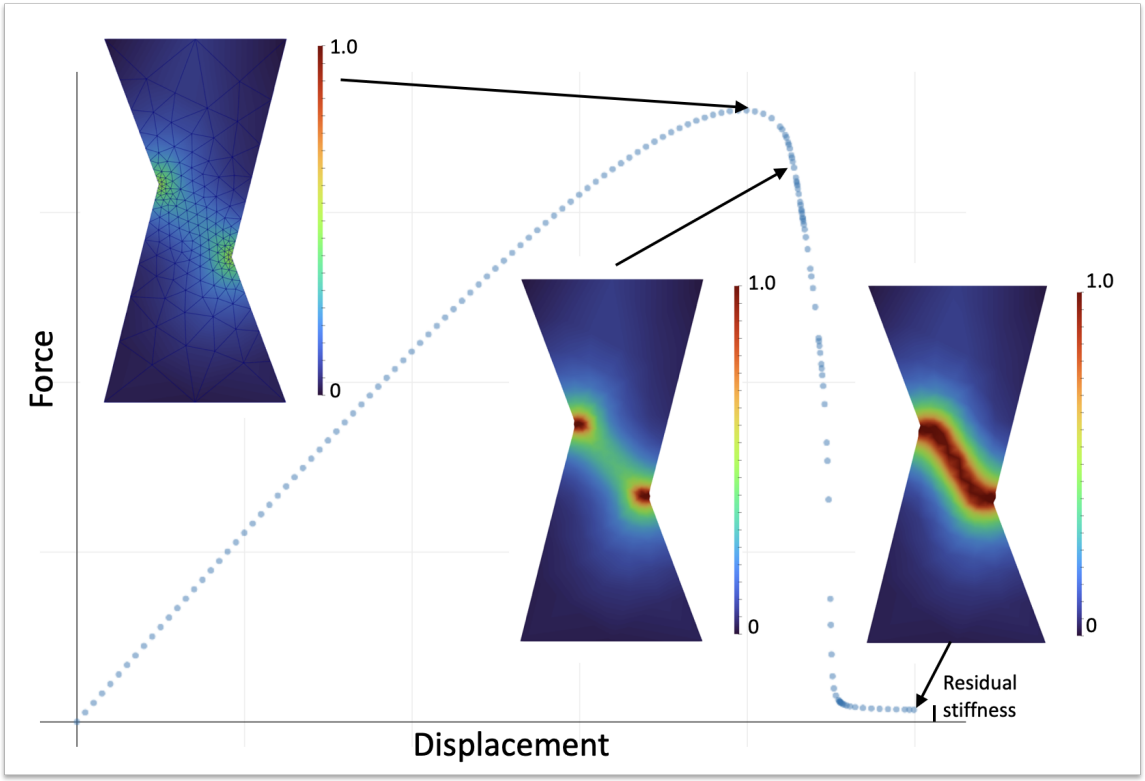


Each material region sets up part of the non-linear and linear equations

Too General, Need Specifics

Ok, lets look at some specific simulations

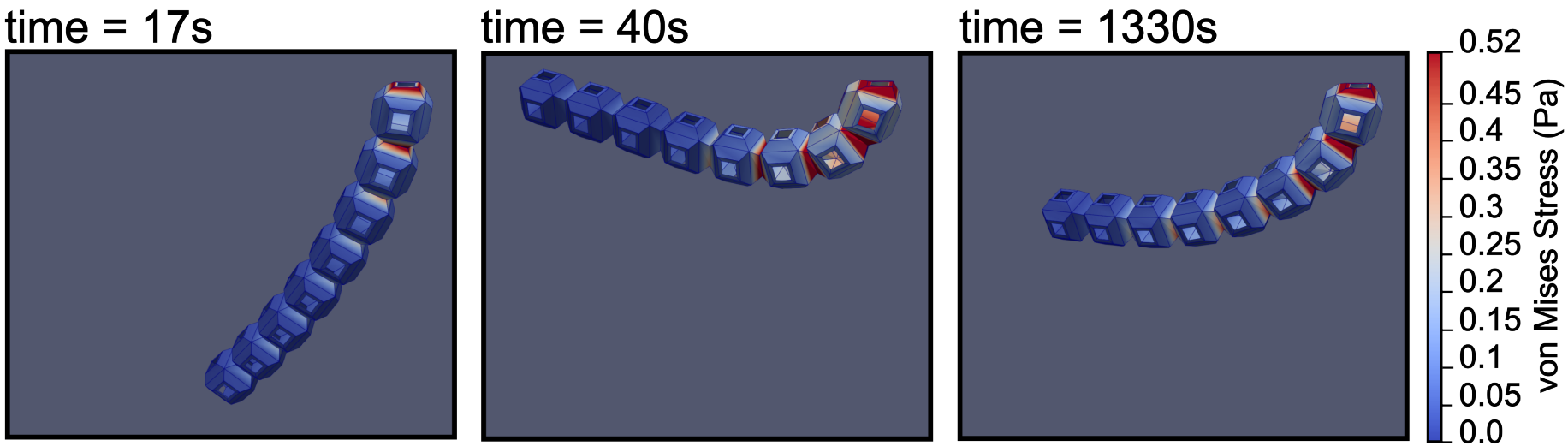
Example - Linear Damage




```
$ bin/ratel-quasistatic -options_file examples/ymls/ex02-
    quasistatic-elasticity-linear-damage-compressiveshear-
    AT2-face-forces.yml
```

Quasistatic simulation of compressive shear for generic brittle material

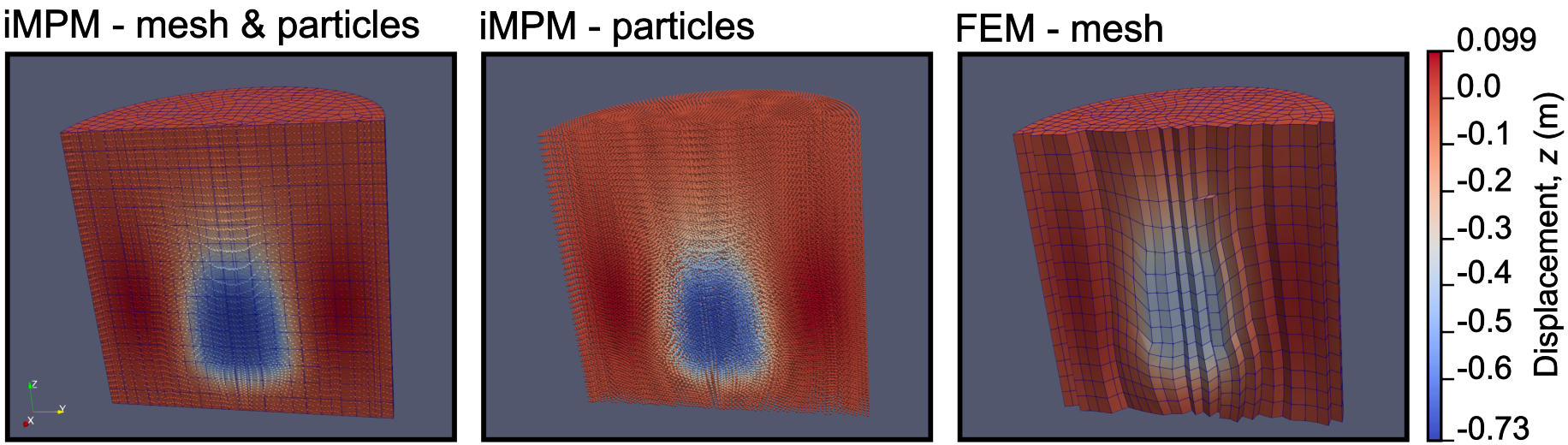
Example - Dynamic Pendulum




```
$ bin/ratel-dynamic -options_file examples/ymls/ex03-dynamic
    -elasticity-schwarz-pendulum-enzyme.yml
```

Dynamic simulation of Neo-Hookean Schwarz-P "pendulum" with Enzyme

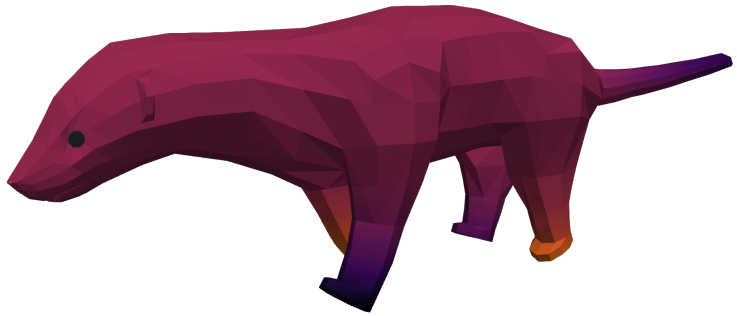
Example - MPM Sinker




```
$ bin/ratel-quasistatic -options_file examples/ymls/ex02-
    quasistatic-elasticity-mpm-neo-hookean-damage-current-
    sinker-cylinder.yml
```

FEM and iMPM simulations of dense sinker in near-incompressible "foam"

Questions?



libCEED Repo: https://github.com/CEED/libCEED
Ratel Repo: https://gitlab.com/micromorph/ratel

Grant: Predictive Science Academic Alliance Program (DE-NA0003962)

Memory Bandwidth vs FLOPs - Tradeoffs on Modern Hardware

Jeremy L Thompson

University of Colorado Boulder

jeremy@jeremylt.org

30 March 2026