Imperial College London

Flexible, Scalable Mesh and Data Management using PETSc DMPlex

M. Lange[1] M. Knepley[2] G. Gorman[1]

[1]AMCG, Imperial College London
[2]Computation Institute, University of Chicago

April 23, 2015

Unstructured Mesh Management

Mesh management

- Many tasks are common across applications:
  Mesh input, partitioning, checkpointing, . . .
- File I/O can become severe bottleneck!

Mesh file formats

- Range of mesh generators and formats
  Gmsh, Cubit, Triangle, ExodusII, CGNS, SILO, . . .
- No universally accepted format
  - Applications often “roll their own”
  - No interoperability between codes

Unstructured Mesh Management

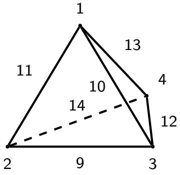
Finding the right level of abstraction

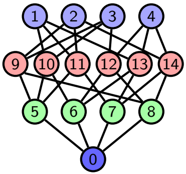
- Abstract mesh topology interface
  - Provided by a widely used library
  - Extensible support for multiple formats
  - Single point for extension and optimisation
  - Many applications inherit capabilities
- Mesh management optimisations
  - Scalable read/write routines
  - Parallel partitioning and load-balancing
  - Mesh renumbering techniques
  - *Parallel mesh adaptivity*

DMPlex: Mesh topology abstraction

DMPlex - PETSc's unstructured mesh API[1]

- Abstract mesh connectivity
  - Directed Acyclic Graph (DAG)[2]
  - Dimensionless access
  - Topology separate from discretisation
  - Pre-allocate data structures
- Enables new preconditioners
  - FieldSplit
  - Geometric Multigrid





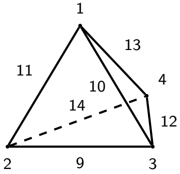
[1] M. Knepley and D. Karpeev. Mesh Algorithms for PDE with Sieve I: Mesh Distribution. *Sci. Program.*, 17(3):215–230, August 2009

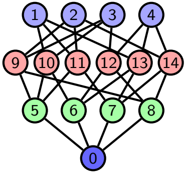
[2] Anders Logg. Efficient representation of computational meshes. *International Journal of Computational Science and Engineering*, 4:283–295, 2009

DMPlex: Mesh topology abstraction

DMPlex - PETSc's unstructured mesh API[1]

- Input: ExodusII, Gmsh, CGNS, Fluent-Case
- Output: VTK, HDF5 + Xdmf
  - Visualizable checkpoints
- Parallel distribution
  - Partitioners: Chaco, Metis/ParMetis
  - Automated halo exchange via PetscSF
- Mesh renumbering
  - Reverse Cuthill-McGee (RCM)





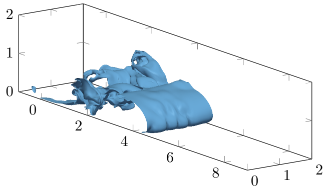
[1] M. Knepley and D. Karpeev. Mesh Algorithms for PDE with Sieve I: Mesh Distribution. *Sci. Program.*, 17(3):215–230, August 2009

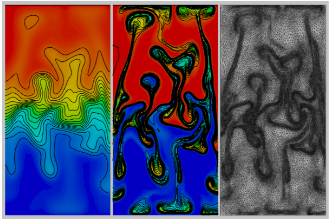
Fluidity-DMPlex Integration

Fluidity

- Unstructured finite element code
- Anisotropic mesh adaptivity
- Uses PETSc as linear solver engine
- Applications:
  - CFD, geophysical flows, ocean modelling, reservoir modelling, mining, nuclear safety, renewable energies, etc.

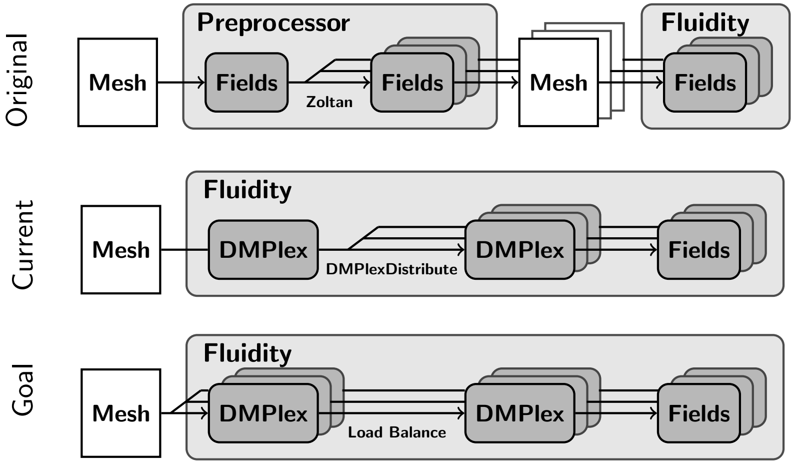
Bottleneck: Parallel pre-processing[1]





[1] X. Guo, M. Lange, G. Gorman, L. Mitchell, and M. Weiland. Developing a scalable hybrid MPI/OpenMP unstructured finite element model. *Computers & Fluids*, 110(0):227 – 234, 2015. ParCFD 2013

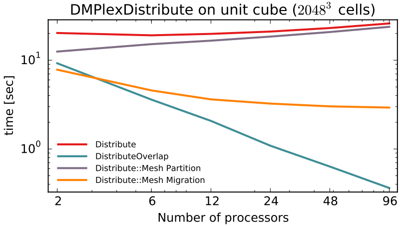
Fluidity - DMPlex Integration

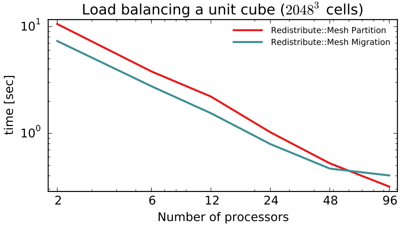


Fluidity - DMPlex Integration

DMPlexDistribute

- Before:
  - One-to-many
  - Single-level overlap
  - Overlap is expensive
- After:
  - Generic mesh migration
  - Parallel N-level overlap
  - All-to-all via ParMetis
- Available to other codes
  - Firedrake, Moose, ...

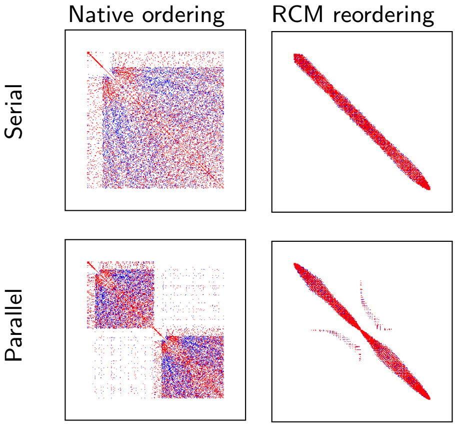




Fluidity-DMPlex Integration

Mesh reordering

- Fluidity Halos
  - Separate L1/L2 regions
  - "Trailing receives"
  - Requires permutation
- DMPlex provides RCM
  - Generated locally
- Fields inherit reordering
  - Better cache coherency



Imperial College
London

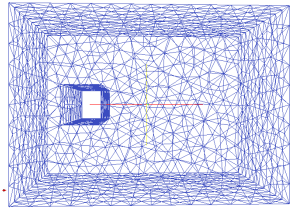
Benchmark

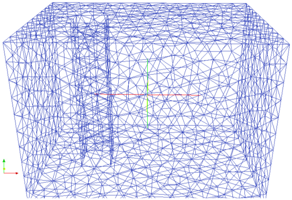
Archer

- Cray XC30
- 4920 nodes (118,080 cores)
- 12-core E5-2697 (Ivy Bridge)

Simulation

- Flow past a square cylinder
- 3D mesh, generated with Gmsh





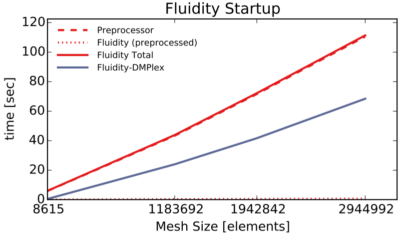
AMCG

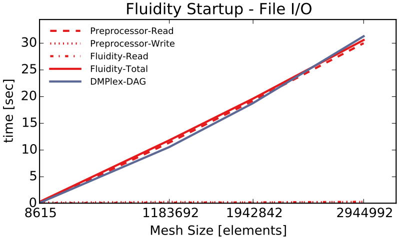
Imperial College London

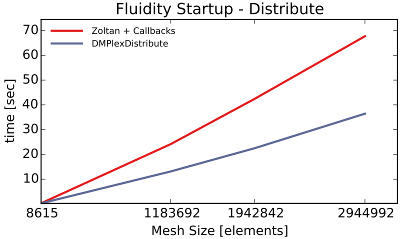
Results - Simulation Startup

Startup on 4 nodes

- Runtime distribution wins
- Fast topology distribution
- No clear I/O gains
  - Gmsh does not scale

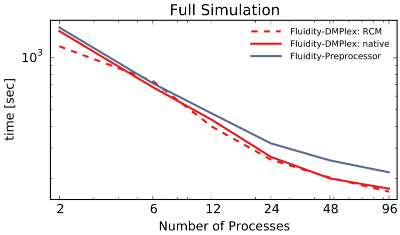






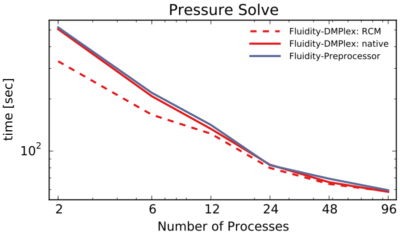
AMCG

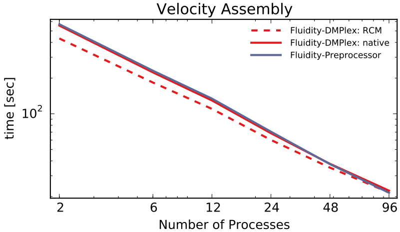
Results - Simulation Performance



Performance

- Mesh with ∼2 mio elements
- Preprocessor + 10 timesteps
- RCM brings improvements
  - Pressure solve
  - Velocity assembly





Discussion and Future Work

DMPlex mesh management for Fluidity

- No need to preprocess
- Increased interoperability:
  - ExodusII, CGNS, Fluent-Case
- Performance benefits
  - Fast runtime mesh distribution
  - Optional RCM renumbering

Future work

- DMPlex-based checkpointing in Fluidity
- Scalable parallel mesh reads with DMPlex
- Anisotropic mesh adaptivity via DMPlex