1. The Role of Mesoscale Eddies in the Meridional Overturning Circulation: Overview

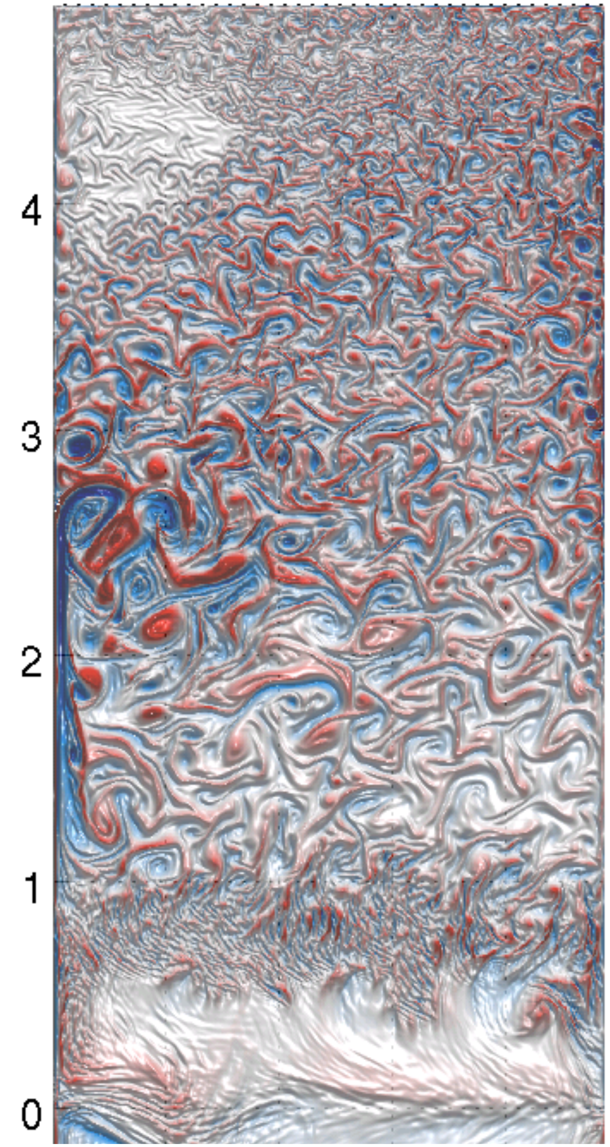
PI: Paola Cessi, Scripps Institution of Oceanography, UCSD

- Our project examines the role of mesoscale flows in the maintenance of the oceanic main thermocline, of the deep stratification, and in the heat transport.

- Mesoscale flows are not resolved by climate models that are integrated over long times.

- We study the **climatological response** of eddy-resolving ocean models to changes in external parameters, this requires integration times of several centuries.

- Future studies will focus on regional processes (i.e., eddying boundary currents) and their connection to the large-scale circulation. These will require high resolutions and non-hydrostatic dynamics.



2. Current HPC Requirements

- Architectures:
  - Currently: Cray XT4 (NERSC), BlueGene/P (ALCF)
  - Previously: Cray X1E (ORNL), BlueGene and IBM SP4 (SDSC)

- Compute/memory load:
  - 1-7 million core-hrs/year (1-2 million at NERSC)
  - 1024-2048 cores
  - ~250 wallclock hrs per production run
  - 0.25-.5 million core-hrs per production run
  - need <0.5 GB memory per core (<512 GB total)

- Data read/written:
  - ~100 GB per run
  - checkpoint 2-4 GB
  - move 0.5-1 GB from NERSC every couple of days
  - require ~100 GB online storage, 2-4 TB offline storage

2. Current HPC Requirements

- Necessary software, services or infrastructure: MPI, netCDF

- Current primary codes and their methods or algorithms
  - Use the MITgcm (externally maintained, public GCM)
  - Solves incompressible Navier-Stokes equations by explicit/implicit time-stepping
  - Parallelized via 2D (horizontal) domain decomposition
  - Inter-tile communication via MPI (OpenMP sublayer has fallen into disrepair)
  - All implicit steps are tile-local except for 2D global elliptic problem for pressure (3D for non-hydrostatic mode). Elliptic problem solved using pre-conditioned conjugate gradient solver.
  - Current experiments: $1792 \times 448 \times 20$ = 16 million grid points and require ~16 million time-steps.

2. Current HPC Requirements

- Known limitations/obstacles/bottlenecks
  - **Latency of global reductions:** >50% of computational time is spent in elliptic solver and 30%–50% of that time is spent on global reductions (less of a problem on BlueGene/P than CrayXT4)
  - **Memory bandwidth:** each step requires a lot of highly strided access to much more memory than can fit into CPU caches
  - **System stability:** long runs have recently meant that the system crashes several of times per run, requiring backtracking or other mucking about. This makes it difficult to complete runs in a timely manner

3. HPC Usage and Methods for the Next 3-5 Years

- Upcoming changes to codes/methods/approaches
  - Will need to fix OpenMP layer
  - Possibly replace the elliptic solver
  - (These changes will have to either be relatively easy or implemented by the MITgcm team: we are just end-users.)

- Changes to Compute/memory load
  - **Higher resolution runs:** up to 5K processors, probably factor 5 increase in total memory, but not a significant increase in memory per node.
  - **Tracer/Lagrangian transport studies:** Ensembles, more memory.

- Changes to Data read/written
  - At most a factor 5 increase in storage requirements

3. HPC Usage and Methods for the Next 3-5 Years

- New infrastructure needs:
  - High memory bandwidth and/or have large caches
  - Low latency on collective operations (like BlueGenes)
  - More FPUs wouldn’t hurt …
  - Better fault tolerance/less downtime

- Anticipated limitations/obstacles/bottlenecks on 10K-1000K PE system.
  - Memory bandwidth and cache size.
  - Programming model/data locality.

4. Summary

- What new science results might be afforded by improvements in NERSC computing hardware, software and services?

  - **Higher resolutions:** Would allow the assessment of the contribution of submesoscale flows and boundary processes on the large-scale circulation. New parameterizations.

  - **Large-scale, non-hydrostatic simulations:** Convective parameterizations in ocean models are currently crude. Experience with atmospheric models shows that proper representation of convection is crucial to obtaining realistic circulations; there are hints this may be true in the ocean as well.

  - Parameter sensitivity studies using eddy-resolving coupled models.

  - Tracer transport simulations: Ensembles, more memory, ...