RMI

HIRLAM All Staff Meeting/ALADIN Workshop
16–19 April 2018, Toulouse

The mass-based non-hydrostatic dynamics dwarf

Daan Degrauwe and Fabrice Voitus

Content

- Context and motivation
- The Mass-based Multigrid/Krylov Solver (MMKS) dwarf
- Preliminar results
- Future plans

For quite some time now, we are concerned by the limitations of our spectral dynamics:

- scalability on massively parallel machines
- stability at very high resolutions (steep slopes)

When considering non-spectral methods, we try to keep as much as possible intact:

- Semi-implicit timestepping
- Semi-Lagrangian advection
- non-staggered A-grid
- mass-based vertical coordinate

Context and motivation

- Concerning the use of the A-grid:

  PhD of Steven Caluwaerts (2016) shows that this should be okay.

- Concerning the mass-based vertical coordinate:

  Recent work on the dynamics equations by Fabrice Voitus:

  - (symmetry in lower boundary condition between implicit and explicit part by using a modified vertical velocity 'W');
  - elimination to D instead of to d
    - $\Rightarrow$ no 'C_2 constraint'
    - $\Rightarrow$ only one Helmholtz equation to solve;
  - formulation of equations with orography treated implicitly.

- Thanks to these developments, the essence of a semi-implicit, mass-based, non-hydrostatic model becomes solving the following Helmholtz equation:

$$\left[\mathbf{I} - \delta t^2 c_*^2 \nabla^2 \mathbf{B}_D^* \mathsf{m}^2\right] D = D^{\bullet\bullet}$$

Dwarf definition

- Thanks to these developments, the essence of a semi-implicit, mass-based, non-hydrostatic model becomes solving the following Helmholtz equation:

$$\left[\mathbf{I} - \delta t^2 c_*^2 \nabla^2 \mathbf{B}_D^* \mathsf{m}^2\right] D = D^{\bullet\bullet}$$

- The central question then becomes:

How do non-spectral methods perform when solving this type of Helmholtz problem?

- Thanks to these developments, the essence of a semi-implicit, mass-based, non-hydrostatic model becomes solving the following Helmholtz equation:

$$\left[\mathbf{I} - \delta t^2 c_*^2 \nabla^2 \mathbf{B}_D^* \mathsf{m}^2\right] D = D^{\bullet\bullet}$$

- The central question then becomes:

  How do non-spectral methods perform when solving this type of Helmholtz problem?

- To answer this question, the decision was made to develop the Mass-based, Multigrid/Krylov Solver (MMKS) dwarf:
  - standalone program to focus on a key archetype problem
  - simpler than a toy-model (no time integration, advection, diffusion, etc.)
  - technically closer to the full ALADIN/HIRLAM model (3D, MPI-distributed)

Dwarf definition

What makes these equations unique w.r.t. e.g. UM or ICON:

- keeping the mass-based vertical coordinate guarantees hydrostatic balance. This is much harder to achieve with a height-based coordinate.
- (for the time being), the reference state is kept very basic:
  - at rest
  - hydrostatically balanced
  - isothermal
  - dry

  The only difference with our current dynamical core is that the reference state *can* account for orography.
- in some sense, this dwarf proposes a way in between
  - (a) The very strict (horizontally homogeneous!) spectral method
  - (b) Using the actual atmospheric state as the reference state

Three versions of the dwarf will be considered:

Sally, the spectral dwarf

Dwarf incarnations

Three versions of the dwarf will be considered:

Sally, the spectral dwarf

Kristof, the Krylov dwarf

Dwarf incarnations

Three versions of the dwarf will be considered:

Sally, the spectral dwarf

Kristof, the Krylov dwarf

Mike, the Multigrid dwarf

- We all know Sally!
- The spectral dwarf is in fact what we use in ALADIN/HIRLAM.
- The Helmholtz equation is solved as follows:
  1. Transform the RHS to spectral space
  2. Divide every spectral coefficient by $\left(1+\frac{\delta t^2 c_*^2 \lambda_l}{k_x^2+k_y^2}\right)$
  3. Transform back to gridpoint space
- The communications for the transforms are quite heavy
  $\Rightarrow$ all-to-all communications

- Krylov methods are iterative methods to solve (sparse) linear systems.
- For instance, the Conjugate Gradient algorithm solves the system $\mathbf{Ax} = \mathbf{b}$ as follows:

1. Initialize $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$; $\mathbf{p}_0 = \mathbf{r}_0$
2. Iterate over $j = 0, \ldots$ until convergence, taking following steps:

$$\alpha_j = \frac{\mathbf{r}_j^T \mathbf{r}_j}{\mathbf{p}_j^T \mathbf{A} \mathbf{p}_j} \tag{1}$$

$$\mathbf{x}_{j+1} = \mathbf{x}_j + \alpha_j \mathbf{p}_j \tag{2}$$

$$\mathbf{r}_{j+1} = \mathbf{r}_j - \alpha_j \mathbf{A} \mathbf{p}_j \tag{3}$$

$$\beta_j = \frac{\mathbf{r}_{j+1}^T \mathbf{r}_{j+1}}{\mathbf{r}_j^T \mathbf{r}_j} \tag{4}$$

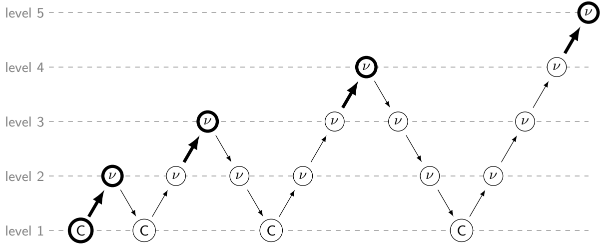
$$\mathbf{p}_{j+1} = \mathbf{r}_{j+1} + \beta \mathbf{p}_j \tag{5}$$

Steps in the Krylov algorithm involving communications are:

- Evaluation of $\mathbf{Ap}$, which involves taking derivatives
  - $\Rightarrow$ halo exchange between neighbouring processors
- Scalar products
  - $\Rightarrow$ global reduction

- Iterative solvers usually have difficulties with the large scales, while small scales converge more quickly.
- Multigrid methods exploit the fact that large scales at one resolution, are actually small scales at a coarser resolution.
- Moreover, the problem becomes much cheaper at coarser resolutions.

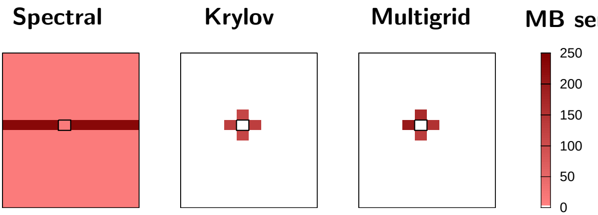
The Full Multigrid Method (Fulton, 1986) uses simple relaxations (ν) at subsequent resolutions to arrive at a solution:



- At the coarsest resolution, one can use a direct solver, a spectral solver, or a Krylov solver
- No global communications are required
- Halo-exchanges are required for the relaxation steps and for the interpolations.

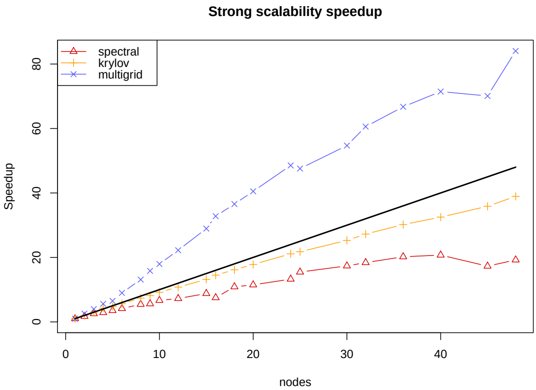
Communication volume

Difference in communications is already visible from a small test on 192 tasks:



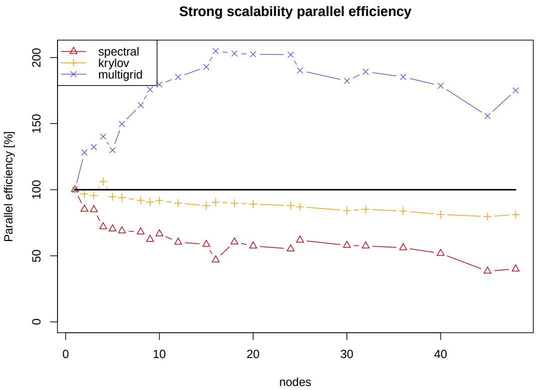
Total 5071 MB | Total 489 MB | Total 665 MB

Performing tests with a fixed grid ($1536 \times 2304 \times 90$), on an increasing number of processors on RMI's hpc:

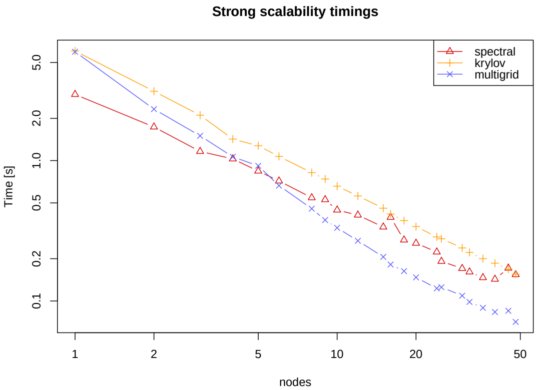


Strong scalability tests

Performing tests with a fixed grid ($1536 \times 2304 \times 90$), on an increasing number of processors on RMI's hpc:



Performing tests with a fixed grid ($1536 \times 2304 \times 90$), on an increasing number of processors on RMI's hpc:



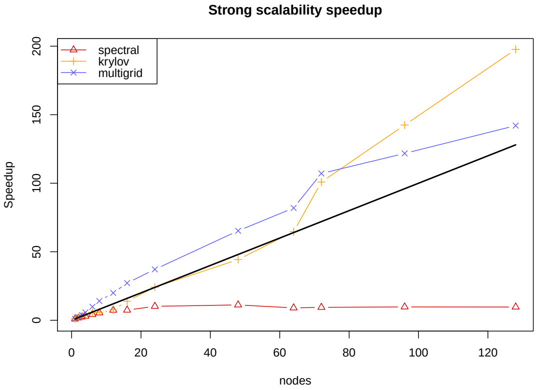
Careful when jumping to conclusions!

Strong scalability tests

Important remarks on these results:

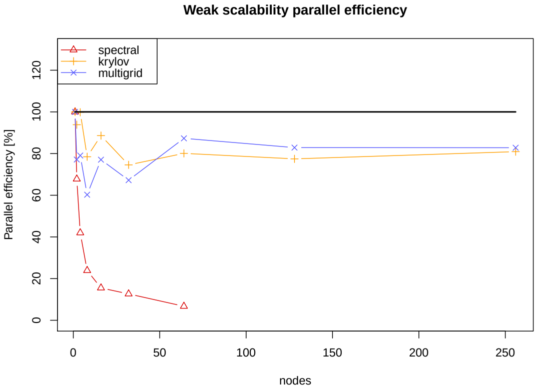
- Number of iterations was fixed; for other weather conditions, convergence may not be complete!
- Settings for spectral dwarf may not be optimal (e.g. no vertical distribution)
- Only one field is transformed in spectral dwarf;
  in the ALADIN/HIRLAM model, several fields (and their derivatives) need to be transformed.
- Gridpoint solvers implementation is not fully optimized

Running with a $1728 \times 1728 \times 90$ grid on ECMWF's cca:



The Multigrid dwarf performance seems to saturate at higher node counts. The reason is that the coarsest grid isn't large enough to properly distribute between the nodes.

Letting the problem size grow with the number of procs:



Conclusions and future work

- The MMKS dwarf provides a tool to test the practical impact (scalability!) of recent theoretical developments on the dynamics equations
- Preliminar results are very promising, but ...

Conclusions and future work

- The MMKS dwarf provides a tool to test the practical impact (scalability!) of recent theoretical developments on the dynamics equations
- Preliminar results are very promising, but . . .

. . . a lot of work remains to be done!

- Test robustness of non-spectral solvers under various meteorological conditions
- More efficient/robust solvers (e.g. preconditioning)
- Tests with implicit treatment of orography; effect on stability
- Plug in full ALADIN/HIRLAM model; study effect on accuracy
- Integration with Atlas
- . . .

Thank you