Work Plan on High-Performance Computing for HIRLAM/ALADIN models

25 May 2011

HIRLAM/ALADIN dynamics/physics expert group

(file: Final_version_WP.tex)

1. Introduction

Due to uncertainties and sometimes conflicting views about long terms options related to dynamics regarding High-Performance Computers, a working group on dynamics/physics for ALADIN/HARMONIE was formed to examine this specific issue and propose agreed actions. The PAC ALADIN asked this working group to produce a work plan about short-term actions and long-terms programs in this area.

Beside this separation between short and long term, a twofold separation for working directions has been suggested:

1) How to run the "current system" optimally on HPCs, i.e. without questionning the current general strategy of the models (SL, Spectral, SI,...)

2) Investigate what should be the necessary changes to the current strategy to sustain good performances on HPCs on the longer term, possibly taking into acount the possibility of bigger changes in computer architectures. However, the scalability cannot be the only ultimate criterion, hence it is emphasized that such changes should not lead to a model dynamics of worse accuracy/quality than the initial system.

The working group thinks that these guidelines are appropriate and therefore proposes the actions listed below in sections 2 and 3.

Agreement about the current situation inside the working group

To begin with, the working group would like to take the opportunity for putting their point of view about the current situation:

(a) The problems of foreseeable future performance at short- and mid-term are believed to be less severe in the dynamics than for other parts of the system (assimilation, surface, physics...). This position is suggested by several profiling studies at Météo-France, HIRLAM, ECMWF and also experience from other model communities, as COSMO, for instance (however, for this latter, the strategy for the dynamical core is very different). The main bottle-necks for the current forecast model part are the I/O, the semi-Lagrangian halo (in case of strong scalability constraints) and the imbalances in the physics. However, for the I/O aspect, the problem can be significantly alleviated by getting the largest part of I/O tasks outside of the integration model itself, through the implementation of an "I/O server" on dedicated processors. For the input, the total file (i.e. not-frame files) strategy is to become an handicap for scalability.

(b) Studies in HIRLAM/ALADIN and IFS suggest that the current situation is still safe in the mid-term of four years at least (for architectures up to "some" thousands of processors). The dynamics is not believed to become a severe problem at this term. The AROME model with typical NWP problems still scales reasonably at 4096 task/CPUS, with no I/O, or in a lesser extent, with I/O activated at each hour. The problem could becomes severe beyond 10000 processors, for real NWP-size problems, not before.

(c) On the mid-term, poorly parallelisable parts (e.g. transpositions) , or parts creating load imbalance (e.g. management of SL halo, specificity of E zone) would first become a handicap in the current state of the dynamics.

(d) On an even longer term, the global current strategy (SI, SL, spectral...) lacking of "locality" from the computational point of view, may need to be revised.

(e) The scalability, when examined on a problem of *given size* with respect to increasingly big CPU numbers does not reflects fairly the reality of NWP. Bigger computers will likely be exploited with bigger size problems. In other words, it is not expected in a near future that the growing curve of CPU numbers will cross (or approach) the growing curve of grid-points number; i.e. it is not expected in a foreseeable future that the performance of each individual CPU will become so weak that each CPU will need to be affected to only few grid-points, in order to be able to produce the forecast in the allocated time-slot. Hence, it seems more reasonable to explore "weak scalability" and some "intermediate scalability" rather than "strong scalability" itself, at least for mid-term extrapolations.

(f) It is also somehow assumed that the network bandwidth of future computers will be improving with the increasing numbers of CPU, allowing at least no lost of scalability when the same amount of information per CPU needs to be communicated. (This relatively week assumption ensures the scalability of local methods requiring only exchanges between neighboring points.)

(g) Even in quite "strong scalability" AROME experiments, with 56 points in each of the 1024 tasks/CPUs on ECWMF's IBM-P6 machine (which, once again, in our opinion is not representative of a reasonable NWP problem), the SL version of AROME was not beaten by its Eulerian counterpart in term of elapse time for a given forecast. This seems to indicate that situations where a higher intrinsic performance is associated with less scalable algorithm, could be viable on a longer term than intuitively thought (depending on the communication performances)

2. Work plan with the current general strategy (spectral, SI, SL)

Context

The first –purely technical– question is to determine how the efficiency of the current forecast model varies when the number of CPUs increases and the relative performance of each CPU increases, is stable, or decreases. Besides, studies in HIRLAM Group suggest that vertical Finite Elements for the Euler Equation system are easier to implement in height-based coordinates. Although the vertical discretization is not at the heart of the scalability problem, the link with computer aspects exists through a possible increase in efficiency, linked itlself to a better robustness or stability – or a better accuracy at a given cost. Hence the vertical discretization is also at the heart of the efficiency problem. This illustrates that the guidelines for the evolution of the dynamical kernel can not only involve computer aspects. Computer aspects and other aspects are not always easily separable.

Work plan on the short term

(a) Submission of projects for accessing to computer time on advanced machines, e.g. the Curie II computer at GENCI (90.000 CPU). This should be restricted only to computers with architectures reasonably reflecting the middle-term evolutions on which the current code could be used without heavy modification: no mixed CPU-GPU architectures, nore use of SHMEM library or UPC (Unified Parallel c) language approaches, which are for time being too far from what can be reasonnably expected at middle-term. The aim is to test the applications in forecast mode only, and typical of future configurations.

(b) Experimentations and discussion of the results. This includes experiments with removing "bricks" in order to better estimate the evolution of the relative efficiency of the purely adiabatic kernel, and other bricks (surface package, physics...). Also this may include experiments with adiabatic model options (including comparisons with Eulerian option). The performances should be examined from the point of view of the absolute number of processes but also the number of grid-points per process.

(c) Besides, in order to begin to examine the aspects linked to the global strategy, this task should also contain actions to encourage comparisons with other dynamical cores built on different strategies (COSMO, UM, WRF,...). These models would be compared on the same machines, and mainly in adiabatic mode, in order to compare the efficiency of dynamical cores only, but for other strategies (explicit, Eulerian, ...). Indeed, due to a non-uniform fields decomposition in different models any such comparison will be necessary in a way not contaminated by differing programing strategies. This task requires help from the consortium management.

Work plan on the long term

(a) Implement, test and validate (as an additional functionality) formulations of the Euler Equations system cast in height-based coordinates, with vertical finite-elements discretizations. A demonstration tool in vertical plane cartesian geometry is delivered and produces promising results in academic contexts. As mentionned above, achieving more precise and potentially more robust schemes, indirectly alleviates computationnal constraints since more efficiency is gained at a given level of quality of the results. This aspect is therefore linked to the ones examined here, indirectly.

(b) Explore ways to profit from new hardware or software features, namely, follow activity on mixed architectures (CPU-GPU), devoting some kernels like matrix multiplication or FFT to UPC, involving SHMEM-like approaches,...

(c) Profit from the object oriented programming (OOP) style of coding. This would in the ideal case modularize the code in a way that any encapsulated method could be easily altered by its alternative without harm for the rest of the system. In such a (ideal) system the model adaptation with respect to any HPC architecture will be just reduced to the selection of appropriate methods allowing the most optimal performance for a given platform. (The ongoing OOPS project is for the moment not addressing those issues of model dynamics and related data-flow. It can be however seen as the first and important step toward the full OOP meteorological model.)

3. Work plan with changes in the general strategy (spectral, SI, SL)

Context

For time being, it is agreed that exploring alternative strategies should be made in a conservative way, always maintaining the possibility to use the older relevant solutions, for safety and for step-by-step clean comparison. This means that alternative strategies should be added progressively as extension of the current one, in the most modular way as possible. A project in which an existing dynamical core is imported from outside in our NWP system is regarded as too risky so far from the NWP point of view, and therefore not retained here.

After extensive discussions, it is judged too early – or too speculative at this point – to plan evolutions toward strategies with explicit or minimally implicit time-schemes (in Eulerian mode of course). It was not possible to find a consensus position in which such evolutions would fit in a work plan even at long term. The minimal agreement position which could be reached is that before being able to insert this in a work plan, the clear necessity of such an evolution should be more firmly established, and then recognized inside the group. Consequently, among all the possible relevant changes of strategies, only the thematics of non-spectral and Eulerian approaches are retained here.

The relevance of the spectral strategy may be questioned for middle term future, from the point of view of the computer (non-local character...) but also from the point of view of the high-resolution NWP (very steep slopes and more "non-linear" flow...). The use of non-spectral strategies also raises the problems of horizontal staggering. Besides, the relevance of the semi-Lagrangian technique for high resolutions (with more irregular flows) is sometimes questioned, then justifying comparison studies with Eulerian technique.

The paths proposed here have three-fold utility : (i) address questions about the relevance of current techniques (spectral, semi-Lagrangian) for high-resolutions; (ii) draw a securized path toward techniques which are likely to be more easily scalable than the current ones; (iii) allow to explore new scientific avenues to improve the current model (e.g, inclusion of orography in the SI scheme, non-uniform SI reference states...).

Work plan on the short term

(a) Find a wide consensus on a long-term roadmap which restricts the changes of strategy to the two following elements : non-spectral approach and Eulerian discretisation, as discussed above. This roadmap should be delivered as a document.

(b) Compare the behaviour of Eulerian vs. semi-Lagrangian techniques as they are in the current code, for very-high resolution real-case flows.

(c) Address the problem of horizontal staggering for non-spectral methods: do we really need a staggered grid for non-spectral methods when semi-Lagrangian schemes with high-order horizontal discretizations are used ? The problems of non-staggered grids have been reported to occur in Eulerian context with low-order discretizations but the status in other cases in not certain.

(d) Study of possibilities to improve the current identified limiting factors in the efficiency of the current code. This seems to involve in priority the management of the SL halo, and the extension zone.

Working plan on the longer term

(a) Separately explore the switch to grid-point methods for limited parts of the dynamical core (computation of derivatives, SI solver). Remaining in the current staggering (A-grid) an horizontal finite volume (or elements) method should be implemented and tested as an alternative to the bi-Fourier spectral method for LAM derivatives. In order to remove only the "spectral" option while keeping the SI option, this requires the implementation of a 3D grid-point solver and test of the results/performances all the rest being kept identical. These studies should be led from the scientific as well as from the computer points of view (scalability...).

(b) Final examination of the issue of horizontal staggering. If needed, i.e. in case of difficulties with the finite-volume vorticity-divergence formulation an alternative to A-grid will have to be sought, for permitting a viable finite-volume formulation.

4. Additional remarks

Although this document only involves, by nature, LAM applications, it is recommended that whenever possible, the global application should stay in phase with innovative developments in the model dynamical core.