CY36T2 Phasing stay at Météo-France Toulouse 23/05/2010 – 04/06/2010
Carlos Geijo (AEMET, Madrid)

During this stay (my first at MF) the objective has been to contribute to the technical validation of the new version of ALADIN-ALARO-AROME (ALADIN for short), the CY36T2. This process takes place every time a new release of this limited area weather forecasting model is to be delivered, usually once or twice a year. The word "phasing" has historical reasons and refers to the fact that this activity is necessary to get coherent cycles or versions of different forecasting systems like IFS-ARPEGE or, as in this case, ARPEGE-ALADIN.

During the validation a list of tests is carried out. These tests consist of running the new software with prescribed initial and boundary conditions and a comprehensive set of configurations. For this particular case, the list contains 122 tests in single processor mode and 137 tests in multiprocessor mode. The runs are performed on a NEC SX9 vector supercomputer. The validation of each test is successful only after it has been confirmed that the calculations do not differ significantly from those used as reference.
There is a utility ("mitraillette") that provides a way to run all these tests with a minimum of scripting overhead.

The activity of developers and phasers during a validation period is high. It is always the case that some of the tests will terminate abnormally or unsuccessfully because of several possible different reasons: bugs, erroneous configurations, etc… It is the task of the phasers to help identify the cause and implement its removal. As a consequence of this process, several branches and sub-branches of the release code are generated very quickly (e.g. in the course of a day). This proliferation of versions is conveniently handled with the software configuration management system "ClearCase". In order to improve the coordination of the different people involved in the work, phasing meetings are held regularly (at least once a week) although more limited scope informal talks can take place at any time.

On my arrival Jean Maziejewski provided me with the badge necessary to access the excellent MF "MétéoPole" facilities here at Toulouse, and briefly gave me some useful information about practical aspects to keep in mind during my stay. Eric Escalière, in charge of user support at GAMP, provided me promptly with passwords and basic instructions for the computing system that I have been using in these two weeks. They two have always kindly been available in case of need. My stay has coincided in time with that of Ole Vignes and Oldrich Spaniel, who I have fruitfully interacted with and learnt much from them. My supervisor has been Olivier Riviere who has decisively contributed to make this stay advantageous for the phasing process.

My participation has focused on the check listed with code AR1T. This test targets AROME runs in single processor (tests 117 and 118) and multiprocessor (tests 136 and 137) modes. I come now on to comment on the results of these tests.

Tests AR1T (117-118)
These tests run on the so-called "AROME grand domain" (512x600x60) at a nominal horizontal resolution of 2.5 km and 60 secs time step, with non-hydrostatic dynamics and MF-AROME physics and SURFEX scheme. The integration is 4 hours long and the initial situation is taken from 2-Nov-2009 at 00UTC. No initialization. No PC scheme. In this version, the physics package includes a number of "extrafields" for aerosols and trace gases. Configuration 117 and 118 differ in the initialization of the message passing library, the former switches this initialization off.

The test returned a SIGSEGV (Segmentation Violation) that the trace back tool (Dr.Hook) located in arp/phys_dmn/mf_phys.F90. Investigation proceeded by: a) study of the flow in the routine and concomitant routines where the exception occurred, b) improving the feedback from the system by enabling a higher level of diagnostic messages, c) trying alternative namelist configurations d) introducing check points in the code.

The problem originated in an incorrect referencing to a pointer array (IPTR) in mf_phys.F90. These pointers are used in the interface to AROME physics (arp/phys_dmn/apl_arome.F90) for the tendencies of the "extrafields", which are accommodated in the memory space ZTENDGFL. Once identified the origin of the crash, the fix was simple to implement.

The problem seems to be related to the NEC compiler. Ole Vignes performed a quick test on ECMWF-IBM and did not find the exception, although it was not carefully checked that the two tests were indeed the same test. Yves Bouteloup did not find problems either when running the test on a Linux PC, athough again it is not clear whether he used exactly the same settings.

Once the run completed (118 also), the numerical results where carefully checked with the reference (in this case those obtained for the same case with the previous version 36t1). The comparison revealed significant differences (above 10%) in fields related to surface fluxes (SURFFLU.U, SURFFLU.V and THETAPWP_FLUX). Olivier Riviere suggested then to try an experiment with the new logical switch LDXFUMSE disabled. This switch is a new feature in the coupling between surface and air in the AROME physics package. As he had correctly guessed, after this change the results compared well.

The next step was the generation of a ClearCase view with the fix incorporated. This was taken as an excellent opportunity to learn some basics about ClearCase. Finally, the modification was merged with others and committed to the pre-cycle pack arp_public_CYT36T1_t2.02 the 3th of June.

Tests AR1T (136-137)

The basic setup for these tests is the same as for the previous ones but in this case the multiprocessor mode is used. Configuration 136 is for “pure” A-level parallelization. In this mode, the grid point computations are divided in zonal bands and each one is assigned to a different processor. In transform computations, each processor processes all the vertical levels for a given field. Configuration 137 is for “pure” B-level parallelization. In this mode, calculations in the grid point space are divided in zonal as well as meridional bands while transformations back and forth between direct and reciprocal space are distributed among processors in such a way that different processors can handle different levels of the same field. In addition to these “pure” modes it is possible to have “mixed” modes, i.e., B-level in grid point calculations and A-level for transform calculations and the other way round. Control of these configurations is performed via namelist NAMPAR0, in particular by the parameters NPRGPNS (number zonal bands), NPRGPEW (number meridonal bands), NPRTRW (number of processors to use during transform), NPRTRV (number of processors to use during transform in vertical direction).

Configuration 136 (4,1,4,1) run successfully once the fix for 117-118 was introduced. SX9 delivers one hour forecast in about 9 minutes, with a mean CPU time per 60 seconds time step of about 7 seconds.

Configuration 137 (2,2,2,2) however fails after a few time steps. In non-hydrostatic mode (LNHDYN=T) the system returns Floating-point exceptions in arp/adiab/gpxyb.F90. In hydrostatic mode (LNHDYN=F) the same error occurs in arp/phys_dmn/actqsat.F90. In both cases, few time steps before termination the calculation of semi-lagrangian trajectories by ald/adiab/elarmes.F90 produces departure points out of the domain, the trajectories cross the lower and upper boundaries because some unrealistic vertical velocity values are passed to the SL scheme. Further investigation shows that the Floating-point exceptions are produced (in the hydrostatic case at least) by corrupted values in the temperature fields.

Several attempts to understand and avoid this behaviour like: a) build of an executable with FFT code xla/external/fourier/fft992.F and pre-processor flag MATMUL instead of MATHKESIAN, b) trying alternative namelist configurations (LSLONDEM=F, NDLNPR=0) they all were unsuccessful.

It was checked that “mixed” parallelization modes like (2,2,4,1) work all right. Indeed, it seems that the problem arises only when the vertical direction parameter NPRTRV takes a value bigger than one ((2,2,1,4) gives the same error). The effort to fix the problem included the analysis of the sub-process arp/setup/suspeca.F90 which is responsible for distributing the spectral fields among processor after they have been read from FA files. However this analysis did not lead to conclusions when this phasing stay finished.