Improving ARPEGE/AROME performances

R. El Khatib, P. Marguinaud, F. Voitus, K. Yessad, G. Desroziers,
L. Berre (Météo-France),
T. Dalkiliç (Meteoroloji Genel Müdürlüğü), A. Stanešić (DHZ)
D. Degrauwe (IRM)

- Preparing for scalar architectures

- Fullpos redesign & improvements

- Model fields IO redesign

- Plans for assimilation

- Ideas & plans for the future

First steps (2010)

- Prepare a benchmark at high resolution for Météo-France models :

    - ARPEGE, IFS-like global model

    - AROME, limited area model, same code as ARPEGE

- Enable OpenMP in AROME

- First refactoring of IO in SURFEX (AROME surface scheme)

- First IO server prototype

Fullpos : how it used to work

- Fullpos is the software used by Météo-France and its partners to go from one geometry to another (coupling files, but also in 4DVAR today and OOPS tomorrow)

- Fullpos relies on existing parts of the model (spectral transforms, SL halos, global data flow)

- Because of the limitations of the model in the past, a geometry change used to be a two step process :

  - Start the model, interpolate horizontally, write to disk

  - Restart the model on horizontally interpolated data, interpolate vertically and post-process

Improving Fullpos

The spectral transforms became a modular package and this made possible the handling of several geometries simultaneously.
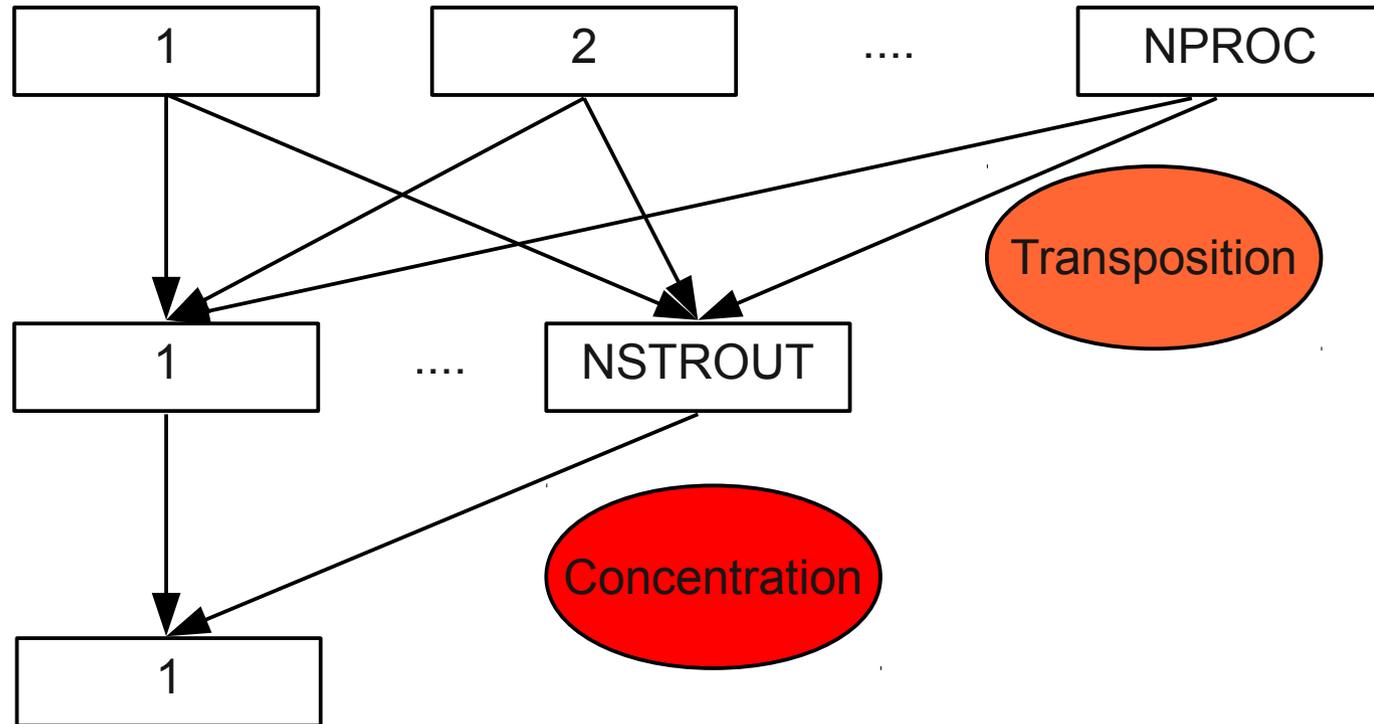
→ less IO, less setup (no need to start the code twice); almost twice faster on global to global, 25% better on global to limited area
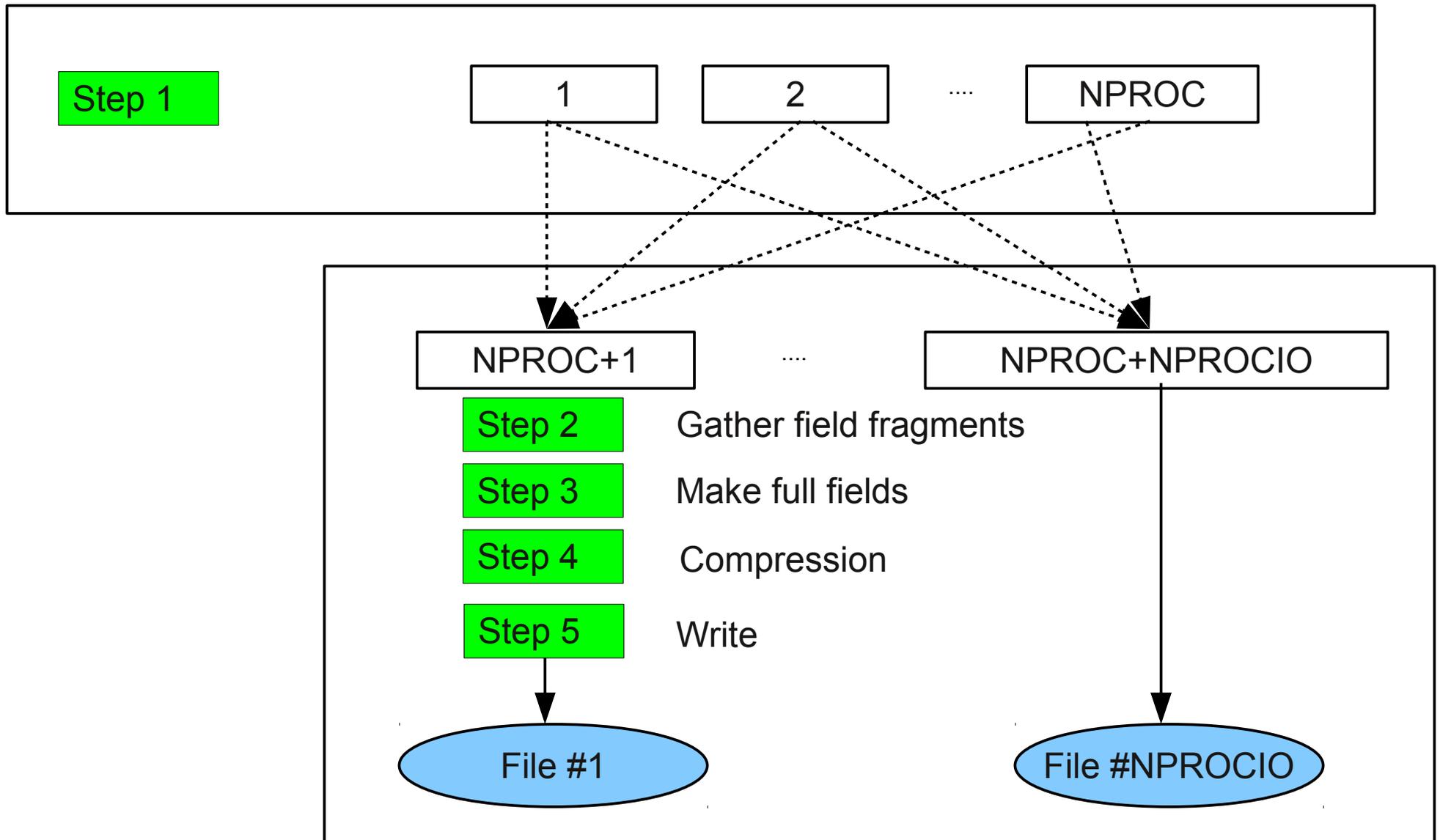
+ Boyd bi-periodicization

+ better handling of extension zone in LAM

Model IO : how it used to work

Using an IO server

Step 1

| 1 | 2 | .... | NPROC |

NPROC+1 .... NPROC+NPROCIO

Step 2 — Gather field fragments

Step 3 — Make full fields

Step 4 — Compression

Step 5 — Write

File #1

File #NPROCIO

Using an IO server

- IO cost in the model is now close to zero

  For AROME 1.3km L90, we use 270 (model) + 4 (IO) Ivybridge nodes

- Full production environment available (tools for retrieving produced model data + handling multiple files)

- Currently used by some HIRLAM partners and soon at Météo-France…

Scalability & data assimilation

- 4D-Var based on a 4D ensemble: 4DEnVar

  – Keeps benefits of 4D-Var (global analysis, same obs. operators, etc...)

  – Localization of the raw covariances made in model space

  – Minimization cost similar to 3D-Var

  – Parallelization properties: no more sequential calls to TL/AD models, but precomputed large ensemble of parallel short-range forecasts, and use of "4D vectors"

- 4DEnVar upcoming challenges:

  – Define a reasonable and efficient localization in space and time

  – Memory and IO issues to handle the large ensemble trajectories

- Other topics of interest :

  – Efficient coupling with ocean and wave models in LAM context

  – **Efficient transport scheme....XXXXXXXXXXXXXXXXX**

The future...

- Follow actively developments made at ECMWF with co-arrays on the core of the model (spectral transforms and semi-lagrangian)

- Become familiar with benchmark extrapolations techniques (on processors and IB fabrics)

- Design a performance simulator to evaluate the performance of key model components on very large numbers of processors using simplified code.