RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

METEO FRANCE

Getting the Météo-France weather forecasting models ready on NEC SX-Aurora TSUBASA

NEC User Group 20-22 June 2023

Ryad El Khatib - Météo-France
With warm thanks to CHMI

Overview

Motivations

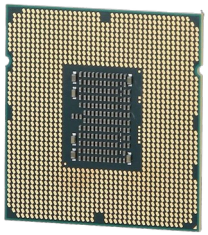
Presentations of the models
Context of work
Source code status and its roadmap

Porting and optimization highlights
Performances : NEC vs AMD
Speedup from double precision to simple precision

Summary and outlook

Motivations

- Météo-France intends to deliver a RAPS benchmark at the end of this year, prior to the ITT by mid-2024



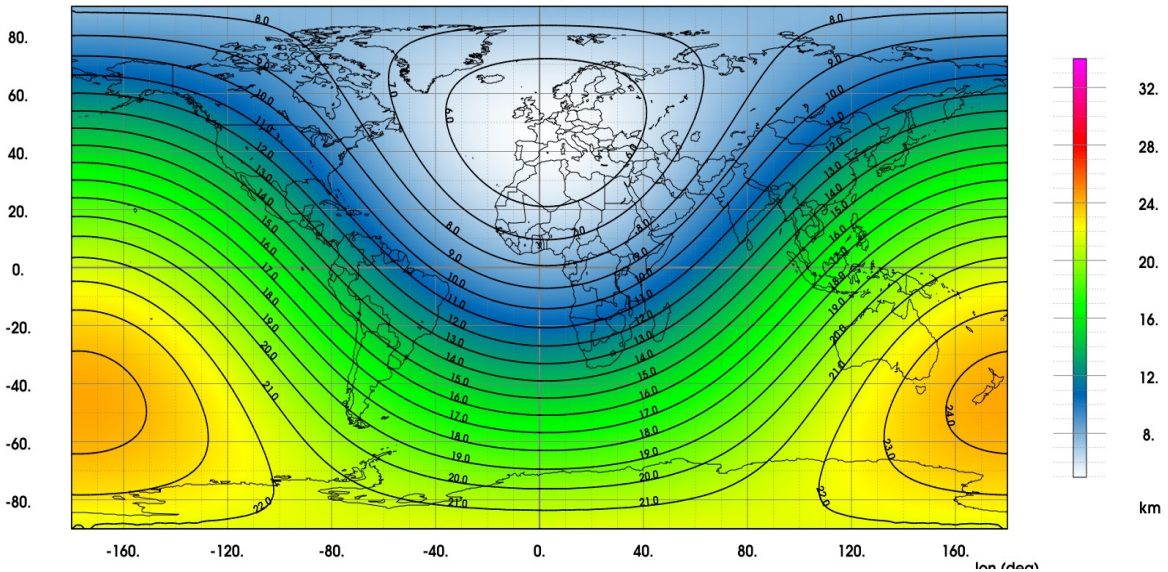
?

- The source code of the numerical models are being deeply refactored to enable them to run on GPUs, but that refactoring must not penalize the other architectures

- NEC vector architecture has always proved to be a competitive solution, therefore we don’t want it to be out of the game

Quick overview of the models : ARPEGE (global)

Systems	Characteristics (since June 2022)
ARPEGE *Deterministic*	Tl1798c2.2 L105 (5km on W Europe) 4DVar (6h cycle): Tl224c1L105 & Tl499c1L105 5 forecasts per day up to 114h
ARPEGE-EDA (AEARP)	Tl499c1 L105 ; 50 members 4D-Var (6h cycle): Tl224c1 L105 Background covariances averaged on 12h and updated every 6h
ARPEGE-EPS (PEARP)	Tl1798c2.2 L90 (5km on W Europe) ; 34+1 members ; 4x102h 35 EDA members and singular vectors Perturbed parameters, 2 convection schemes

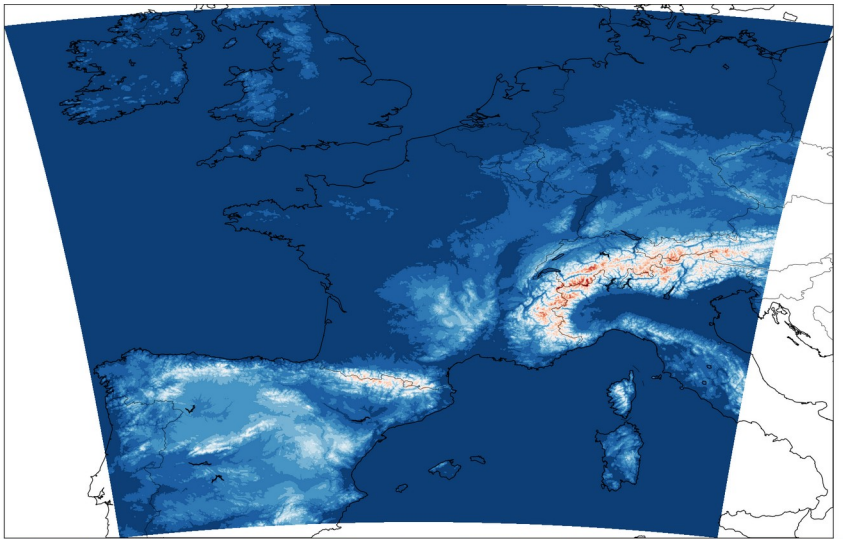


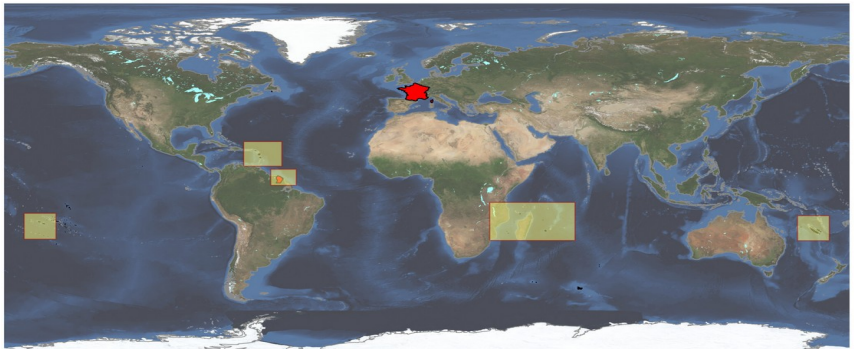
Horizontal resolution
ARPEGE/PEARP
5km < Dx < 24km

(Courtesy of François Bouyssel)

Quick overview of the models : AROME (limited area)

Systems	Characteristics
AROME-France *Deterministic*	1.3km L90 (from 5m to 10hPa) 3DVar (1h cycle) 8 forecasts per day up to 51h
AROME-France Nowcasting	1.3km L90 3DVar (no cycling – 10' cut-off) 24 forecasts per day up to 6h
AROME-IFS	1.3km L90– Dynamical adaptation of IFS (altitude) and Arome-Fr (surface) 4 forecasts per day up to 51h
AROME-EPS (PEARO)	1.3km L90 - 16+1 members Four times per day up to 51h Initial perturbations from AROME-EDA and boundary conditions from PEARP
AROME-EDA (AEARO)	3.25km L90 ; 25 members 3DVar (3h cycle)
AROME Overseas (5 domains)	1.3km L90 – Dynamical adaptation of IFS (altitude) and Arpege (surface) 4 forecasts per day up to 51h
AROME-EPS Overseas (5 domains)	2.5km L90 – 15 members Same initial conditions AROME Overseas Lateral conditions from PEARP 2 forecasts per day up to 51h
AROME Commercial	2.5km L90 – several domains Dynamical adaptation Arpege





(Courtesy of François Bouyssel)

Context of work

- Access granted by CHMI to their system
  - 48 vector hosts containing 384 vector engine cards of type VE 20B
  - fully non-blocking high-speed interconnect
  - total of 18 Terabytes of HBM2 high-speed memory
  - Compiler version 4.0.0 + mpi version 2.22.0

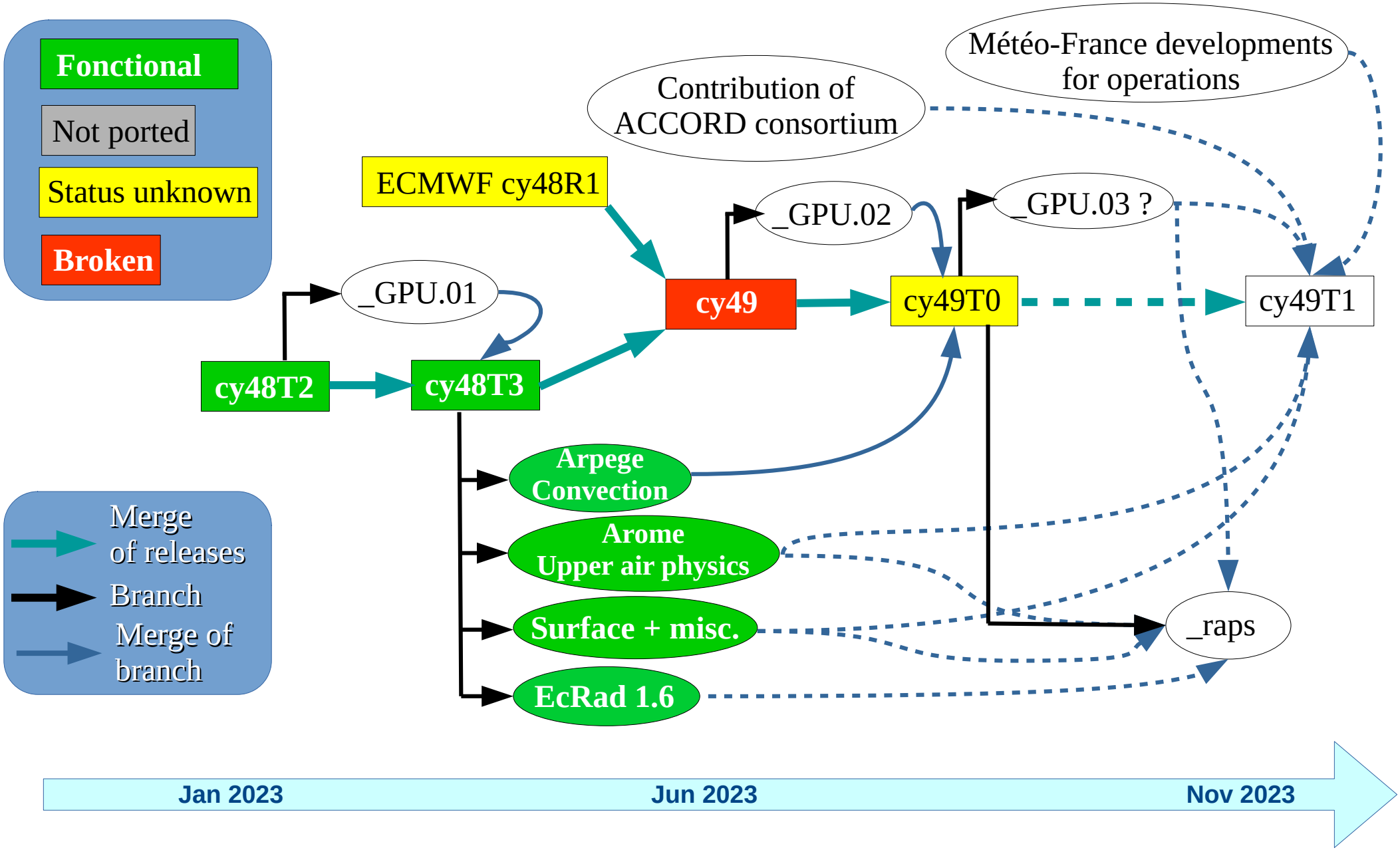
- **8 nodes** at maximum (due to the configuration of the scheduler)
  – limiting, but also an opportunity not to spill resources

- **I/Os disabled** in the tests : later they will be run on the scalar host of the nodes (thanks to an I/O server)
- **Post-processing disabled** because no plan (yet) to port the post-processing on GPUs

- X86-64 reference is Météo-France supercomputer (AMD Rome)

- Forecasting models should run in **simple precision**

Source code status and its roadmap



Porting & optimizations highlights

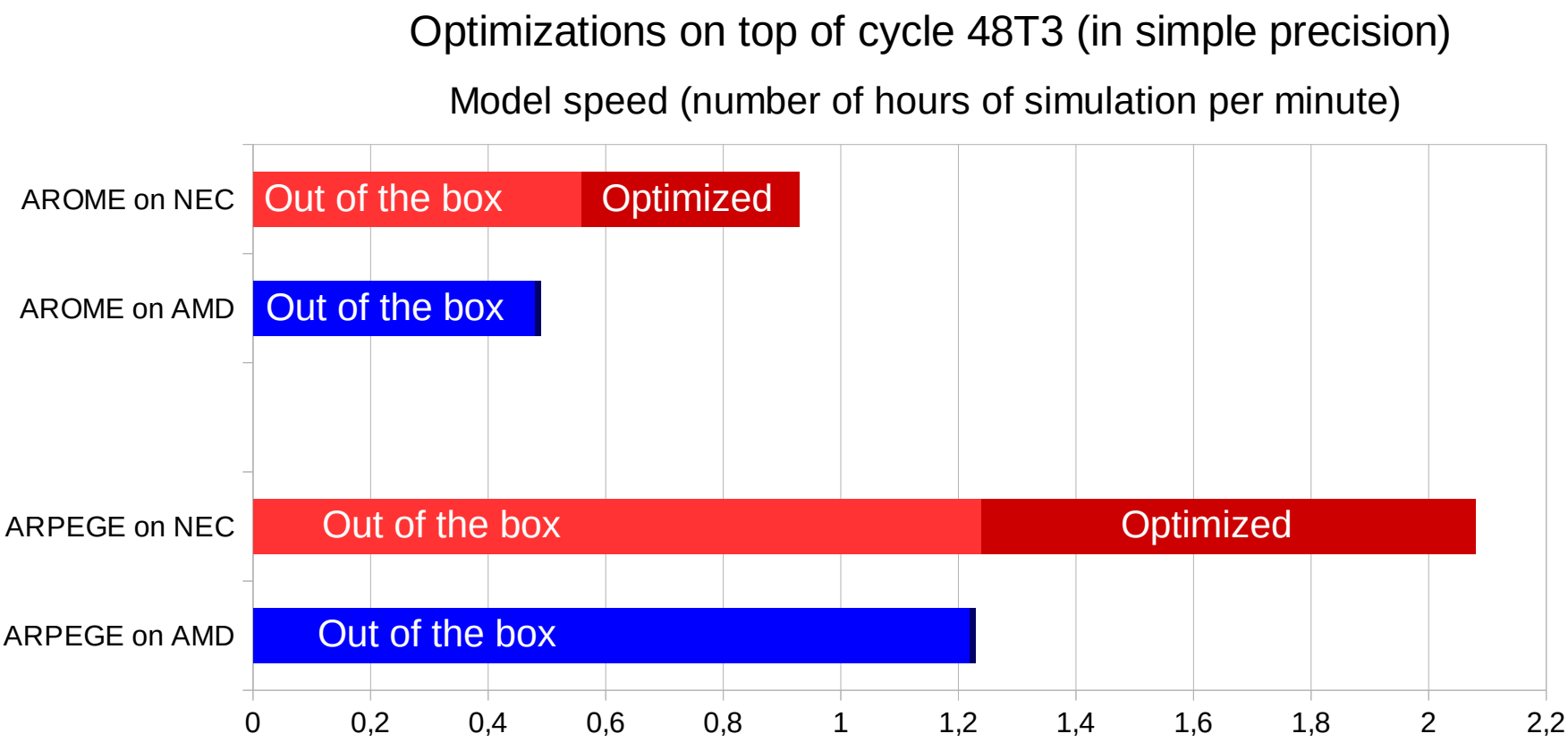
- In general :
  - ✔ Systematic search to try and vectorize expensive non-vector loop
  - ✔ Noticeable speedup here and there by reducing overhead of subroutines calls (in-lining)
  - ✔ Remove array syntax as much as possible

- Spectral transforms :
  - ✔ The traditional Legendre transforms algorithm is faster than the « Fast Legendre Transforms » algorithm (for the model Arpege)
  - ✔ The FFTW package has been ported. it brings a marginal performance improvement (for the model Arome)

- Found optimizations from NEC for DWD (2020) in the latest version of ecRad ☺ ; together with additional optimizations =>
  - ✔ ≈ 30 % speedup for Arpege
  - ✔ ≈ 12 % speedup for Arome

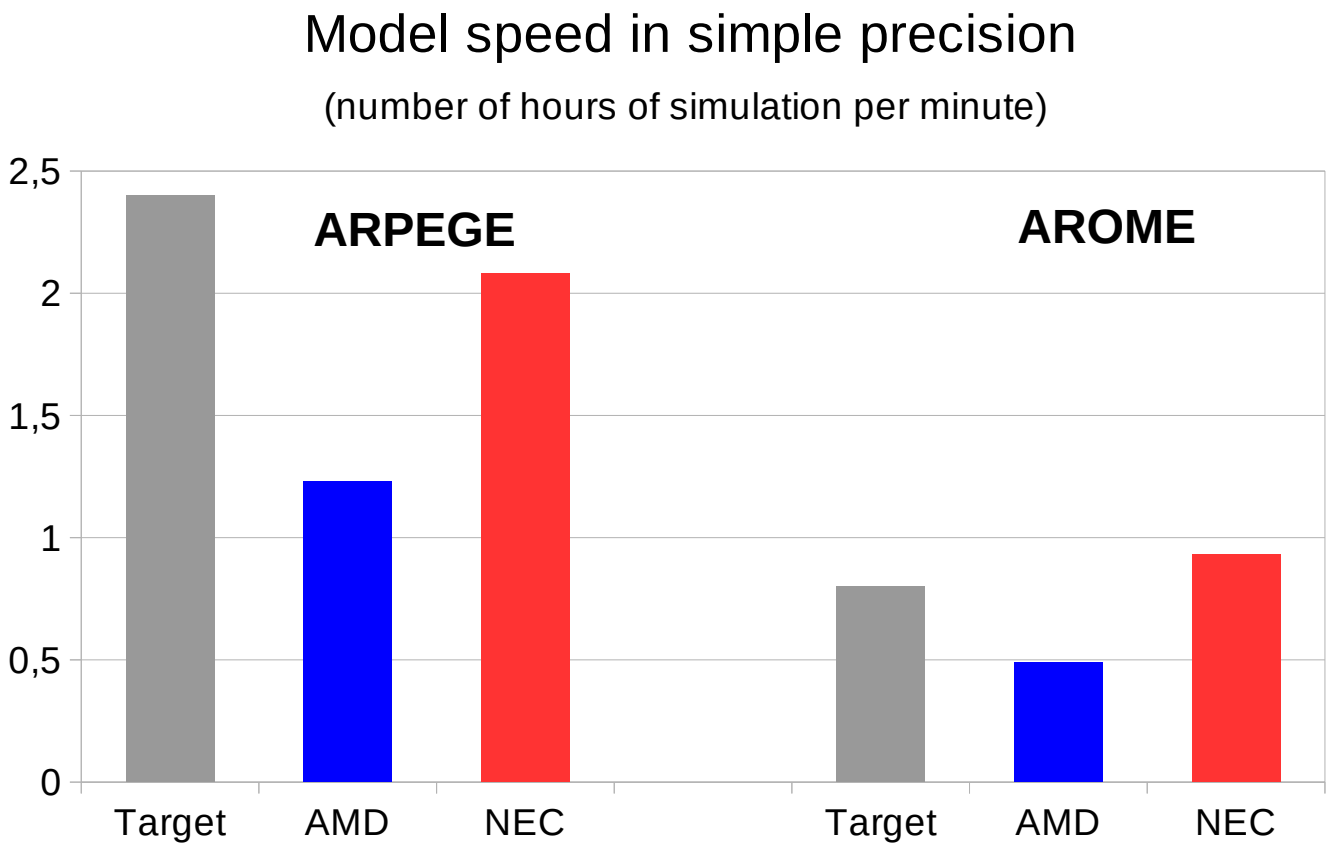
Overall benefits of the optimizations NEC vs AMD

Evaluations on 8 nodes for each architecture



The optimization work benefits to NEC almost exclusively !

Performance against the target : NEC vs AMD



Evaluations on 8 nodes for each architecture

Operational target :

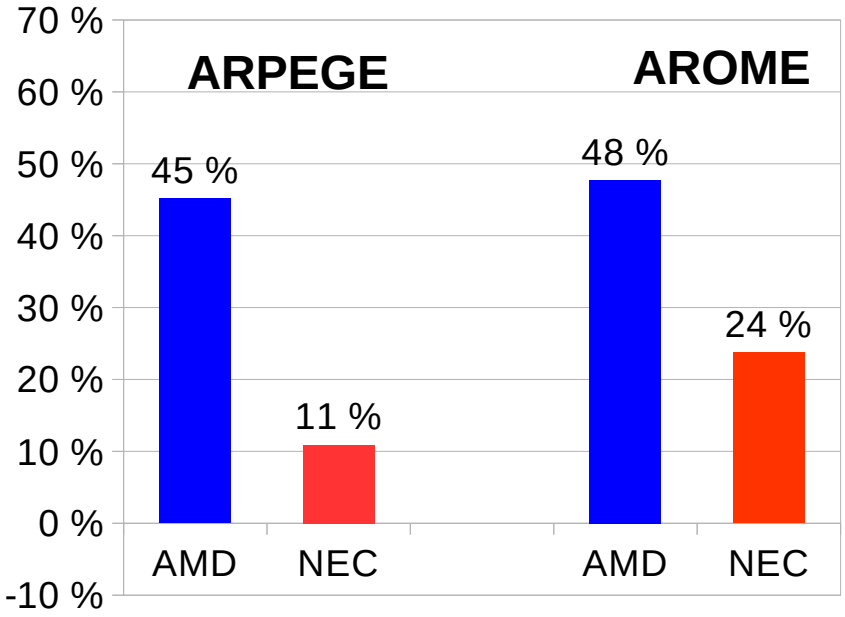
Arpege :
10 min / 24 hours
of forecast

Arome :
30 min / 24 hours
of forecast

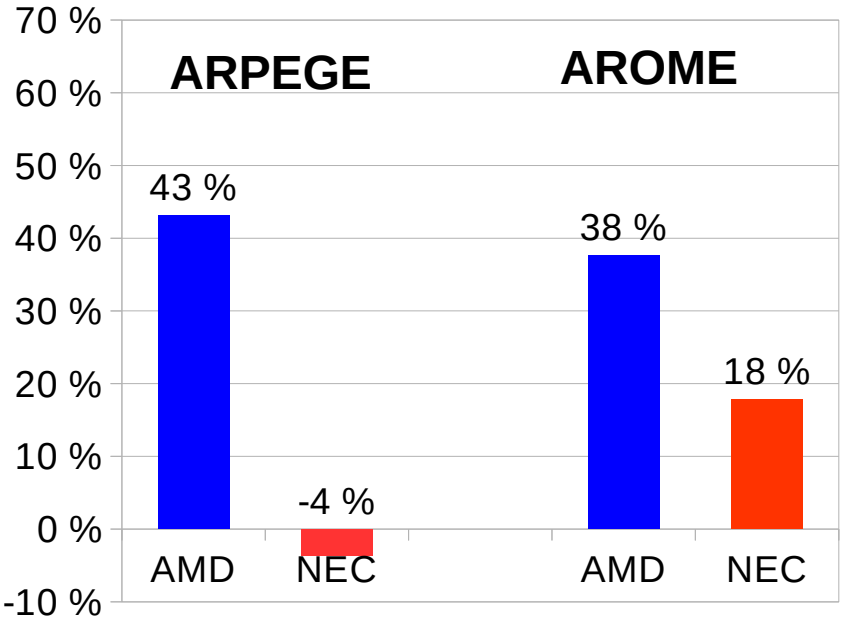
Speedup from double to simple precision NEC vs AMD

Evaluations on 8 nodes

"maximum" speedup from double to simple precision



"Fair play" speedup from double to simple precision



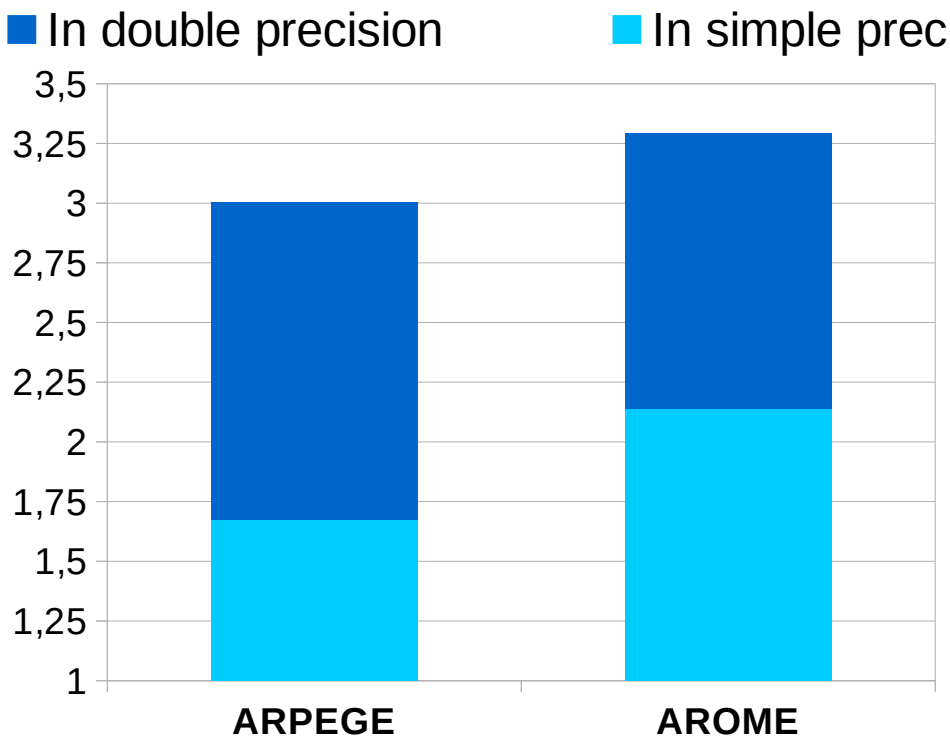
« maximum » = resources are optimized (at least we try to) : increase NPROMA, etc …
« fair play » = only the executable is different

The directors' corner ;-)

Directors needs synthetic graphics and numbers,
and ask how much they need to pay

Ratio of "System Billing Units"*, or

how many nodes AMD Rome for one node VE 20B to run operations



* Not considering the energy consumption

Conclusion : summary and outlook

- The code refactoring needed to run on GPUs doesn’t harm the vector architecture (up to now)
- Arpege is slightly less performing than Arome (but more optimizations seems possible)
- Acceleration from double to simple precision is poor (on VE 20B compared to AMD Rome)
- TODO :
  - Fix cycle 49 !
  - Assimilation model (« OOPS ») : portability test (ongoing) then actual size benchmarking
  - Test higher resolutions (Arpege Tc1800, Arome 750 m)

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

METEO FRANCE

Thank you for your attention !