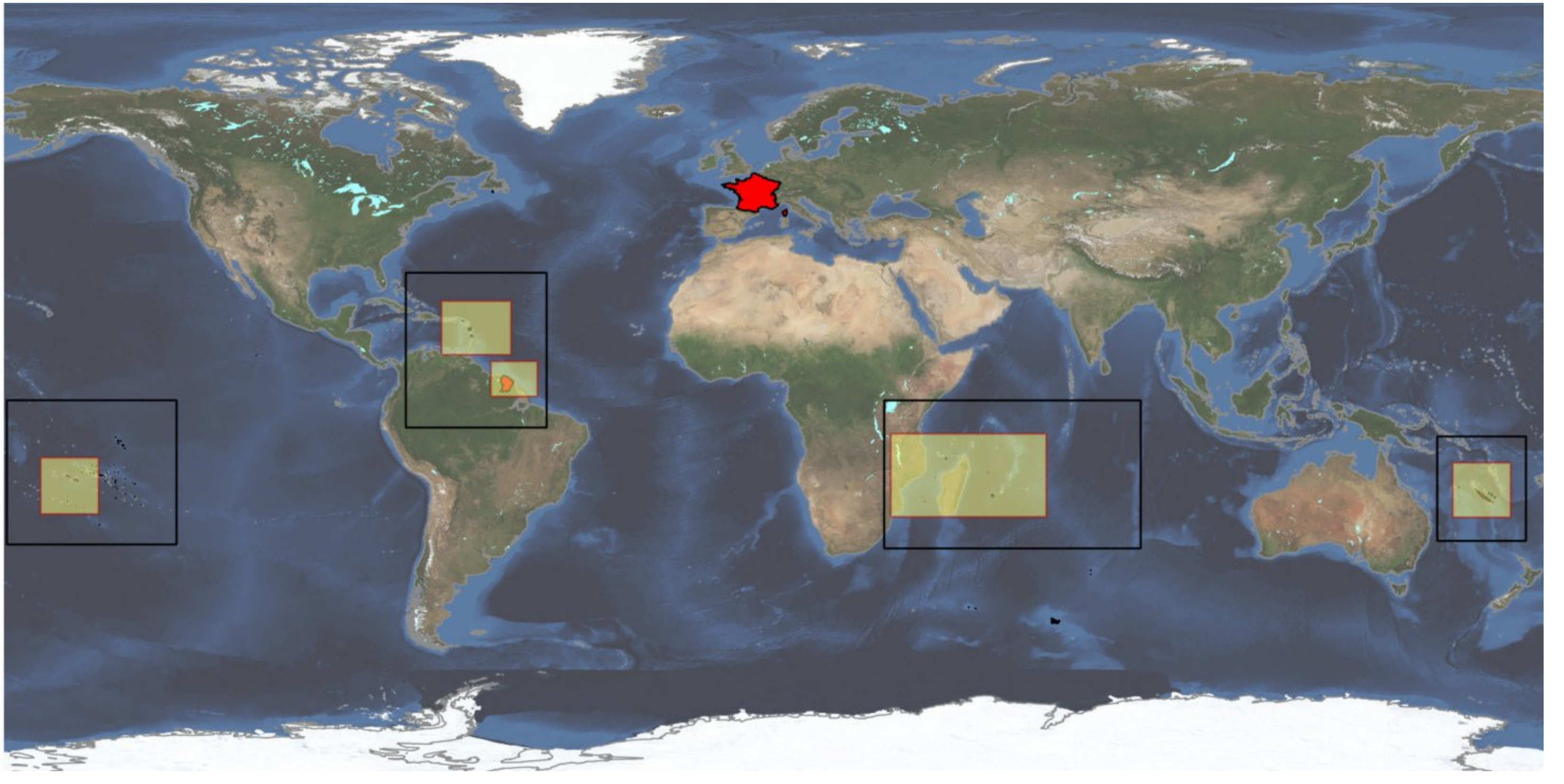
Running ARPEGE-NH at 2.5km

Ludovic Auger, Philippe Marguinaud (CNRM/GMAP)

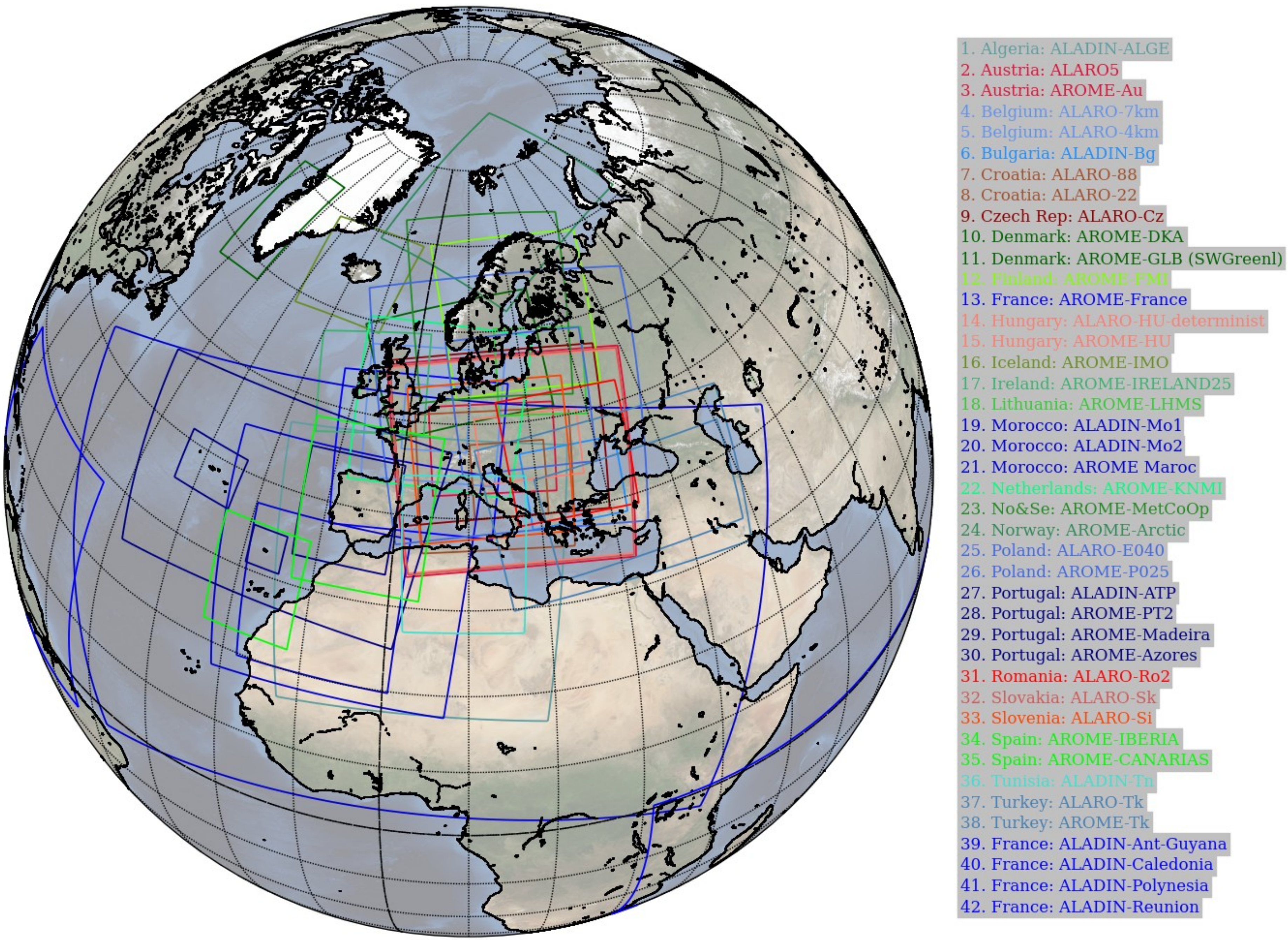
- NH dynamics in AROME and ARPEGE
- ARPEGE-NH & the DYAMOND project
- Profiling ARPEGE-NH at 2.5km
- Perspectives on our future cluster
- A few plots at 2.5km
- Conclusions & perspectives

Running NH models at Météo-France



France & several overseas domains,
for nearly 15 years + many other test domains

Other AROME domains



ARPEGE & AROME NH dynamics

- Very same dynamical core & code

  → Possible to test, validate and debug on small domains or small planets

- Spectral, semi-lagrangian, semi-implicit

  → Very long time-steps ; eg 1.3km ↔ $\Delta t=50s$

- Semi-implicit & predictor/corrector scheme

  → Very stable

AROME & ARPEGE-NH equations

$$\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t} + \frac{RT}{p}\nabla p + \frac{1}{m}\frac{\partial p}{\partial \eta}\nabla\phi = \boldsymbol{\mathcal{V}}$$

$$\boxed{\begin{aligned}&\frac{\mathrm{d}d\!l}{\mathrm{d}t} + g^2\frac{p}{mR_\mathrm{d}T}\frac{\partial}{\partial\eta}\left[\frac{1}{m}\frac{\partial(p-\pi)}{\partial\eta}\right] - g\frac{p}{mR_\mathrm{d}T}\frac{\partial\mathbf{V}}{\partial\eta}\cdot\nabla w \\ &\qquad -(\mathsf{X}-d\!l)\left[\frac{R_\mathrm{d}}{R}d\!l + \left(1-\frac{R_\mathrm{d}}{R}\right)\mathsf{X}\right] - \dot{\mathsf{X}} = -g\frac{p}{mR_\mathrm{d}T}\frac{\partial\mathcal{W}}{\partial\eta}\end{aligned}}$$

$$\frac{\mathrm{d}T}{\mathrm{d}t} + \frac{RT}{C_v}D_3 = \frac{Q}{C_v}$$

$$\boxed{\frac{\mathrm{d}\hat{q}}{\mathrm{d}t} + \frac{C_p}{C_v}D_3 + \frac{\dot{\pi}}{\pi} = \frac{Q}{C_vT}}$$

$$\frac{\partial\pi_\mathrm{S}}{\partial t} + \int_0^1 \nabla\cdot m\mathbf{V}\,\mathrm{d}\eta = 0$$

Additional equations for vertical velocity and true pressure

- Hydro-static equations + 2 new equations for new NH terms :

$$d\!l = -g\frac{p}{mR_\mathrm{d}T}\frac{\partial w}{\partial\eta} + \frac{p}{mRT}\nabla\phi\cdot\frac{\partial\mathbf{V}}{\partial\eta}$$ = vertical divergence

$$\hat{q} = \ln(p/\pi)$$ = reduced true pressure

- The NH formulation extends the hydro-static system
- Stability with a long time step is ensured by the NH implicit system

Predictor/corrector

$$\frac{X_F^n - X_O^t}{\Delta t} = \frac{R(X_F^{n-1}) - R(X_O^t)}{2} + \frac{L(X_F^n) - L(X_O^t)}{2}$$

X^n is an estimation of $X^{t+\Delta t}$ at the iteration n

$R=$ residual $=$ full nonlinear model (M) minus linear part (L)

$X_O($ resp $X_F)$ is the value of the variable at the origin (resp final) point

- Computation at stage (n-1) of X^{n-1} is used to estimate the non-linear residual at time t + Δt.
- The solution of the iterative algorithm tends towards the perfect non-linear implicit scheme.
- The configuration ARPEGE-NH T8000 uses a single iteration (no instabilities observed during a 50 day forecast).

NH time step

- Hydrostatic time-step:
  - Inverse transforms
  - Physics
  - Semi-Lagrangian
  - Direct transforms
  - Semi-implicit
- Predictor/corrector adds the following steps:
  - Inverse transforms
  - Direct transforms
  - Semi-implicit

+ More fields involved in spectral transforms and semi-Lagragian

→ The weight of spectral transforms much higher in NH than in hydrostatic

The DYAMOND project

DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains

- Intercomparison of atmospheric models : ICON, NICAM, MPAS, NASA, FV3, SAM, UM, ARPEGE-NH
- Not EU funded
- 40 day climate run (August September 2016), grid mesh < 5km, no deep convection
- Use ECMWF SST forcing
- Microphysics, radiation, small-scale turbulence, topography
- Output every 3h, total ouput ~ 50Tb

ARPEGE-NH for DYAMOND

- Run the 40 day simulation on our Broadwell cluster
- Run performance tests on ECMWF cluster
- Use only 2x4 cores out of 40, because we have only 64Gb/node, and max job size < 300 nodes
- Use double precision
- 280 nodes (forecast) + 20 nodes (IO)
- Can deliver about 2 hours of forecast per hour of elapsed time

ARPEGE-NH for DYAMOND

- 2.5 km, 75 levels, Δt=100s, linear grid
- Usual ARPEGE physics, without deep convection, without SURFEX
- NH dynamics using one predictor/corrector iteration
- Use 1km GTOPO30 orography, max slope is 49° (38° for AROME); more chaotic orography
- Max height is 8250m
- First level at 20m

Profiling

- Test from 200 to 1000 Broadwell nodes (on ECMWF cluster)
- Only a few time-steps

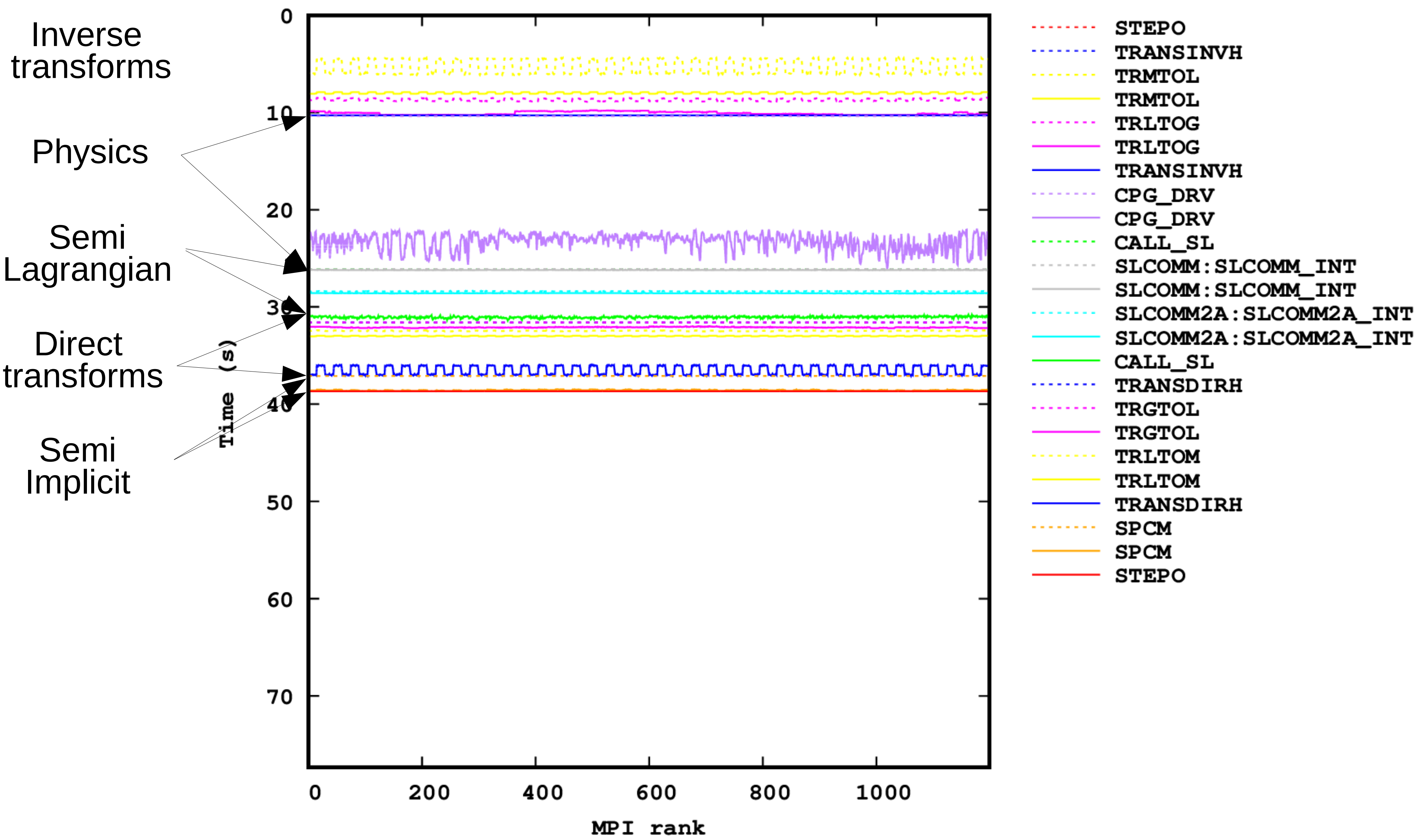
  Insert barriers (using DrHook) in critical routines:

```
SUBROUTINE CALL_SL
IF (LHOOK) CALL DR_HOOK ('CALL_SL', 0, …)      ──►   mpi_barrier ();
                                                     c_drhook_start ("CALL_SL");
...

IF (LHOOK) CALL DR_HOOK ('CALL_SL', 1, …)
END SUBROUTINE                                 ──►   c_drhook_end ("CALL_SL");
                                                     mpi_barrier ();
```

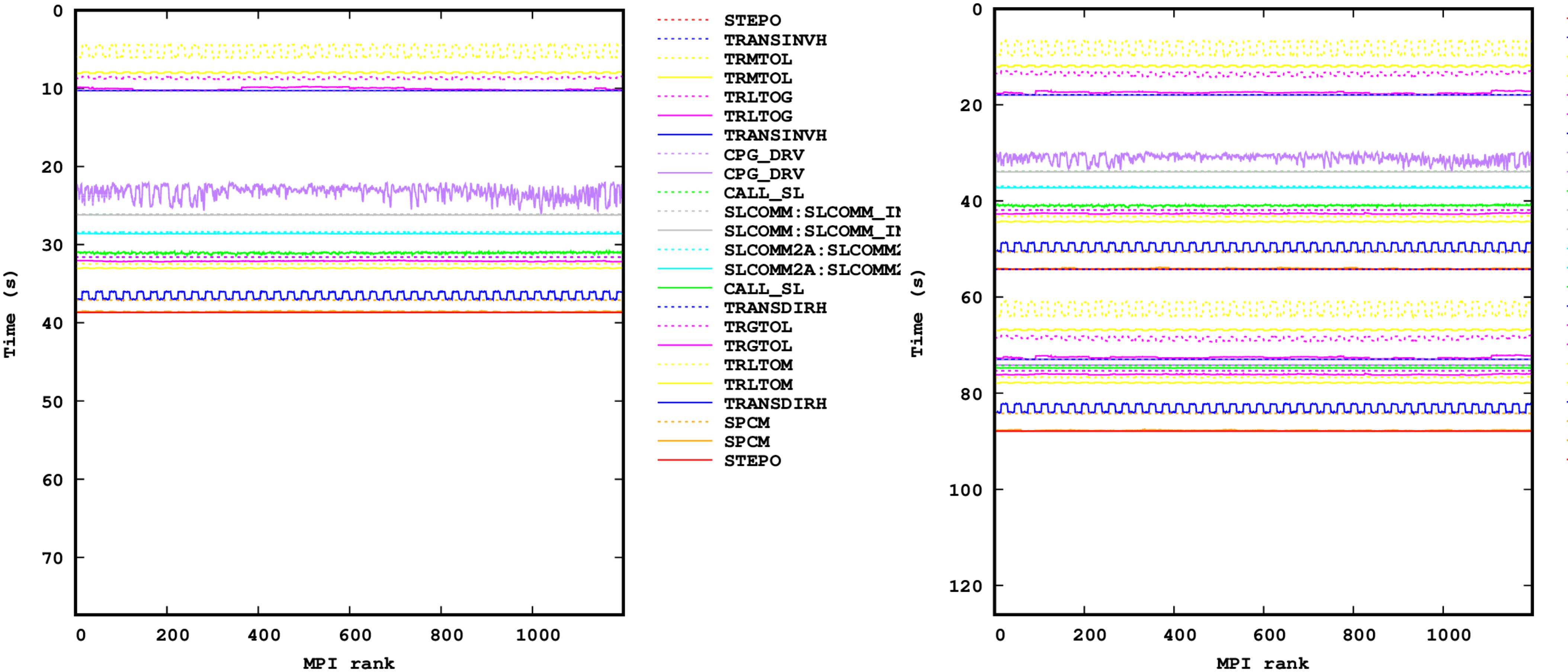
Record the times when a routine is entered and left, after and before the barriers
→ Highlight the real cost of communication and load imbalance
→ Can study individual time steps

Profiling
(single hydrostatic time-step)



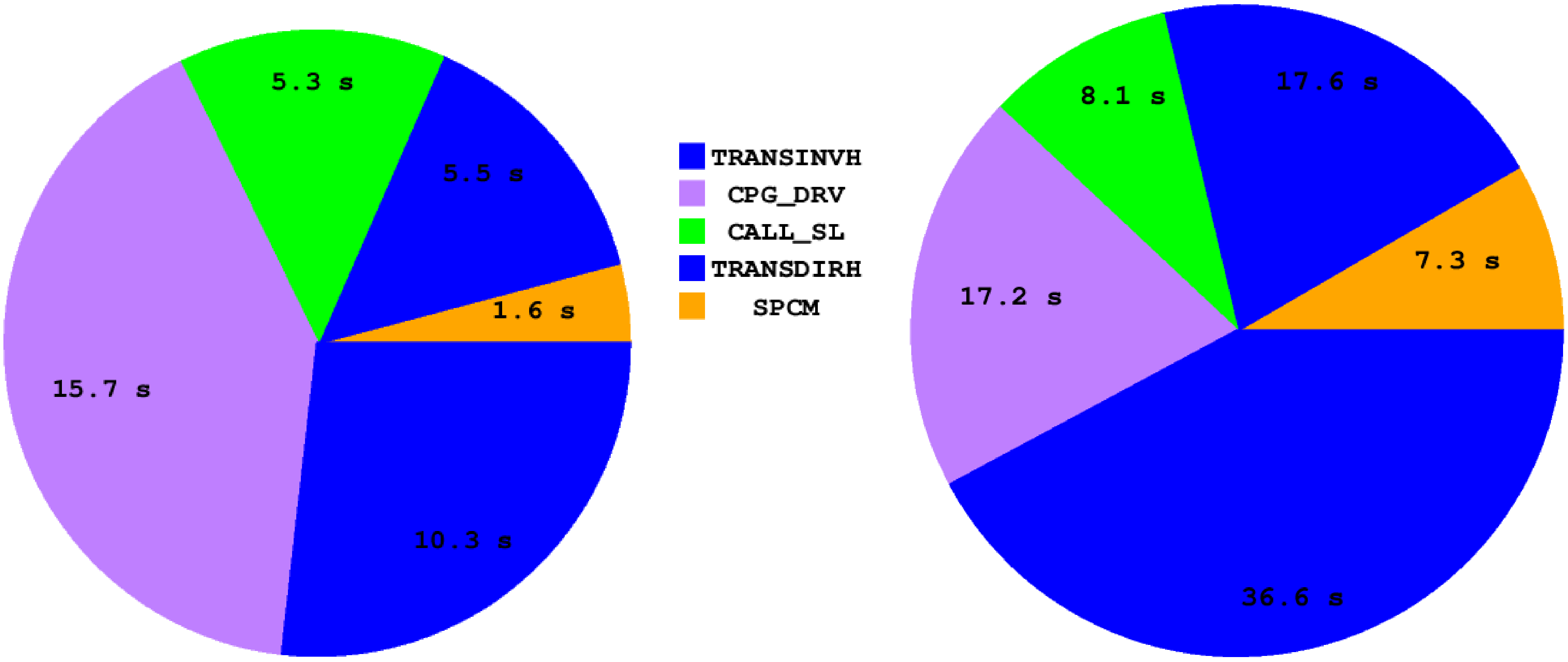
ARPEGE H vs NH

(200 Broadwell nodes)

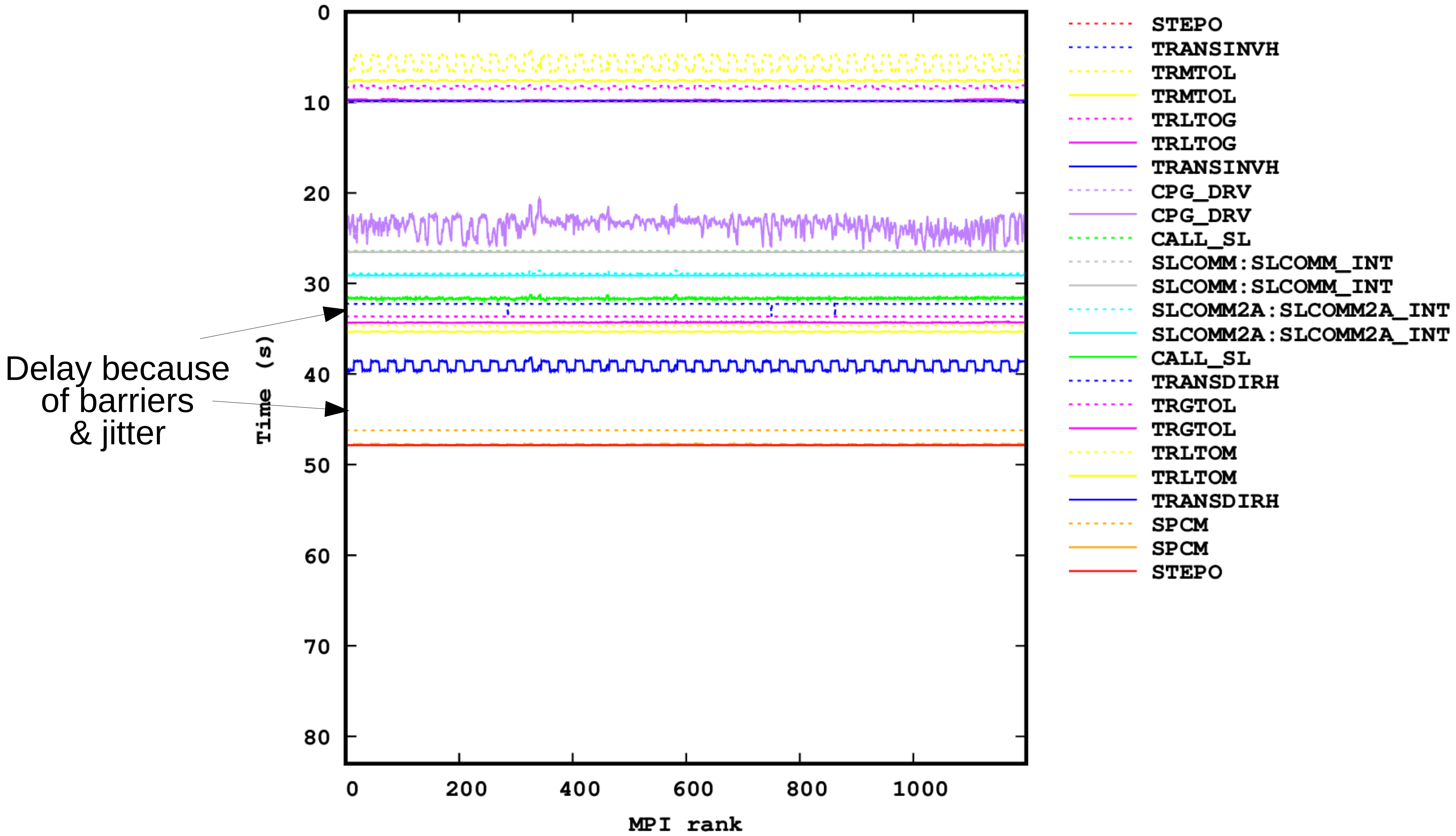


More steps : Inverse transforms, direct transforms, semi-implicit + more fields

ARPEGE H vs NH

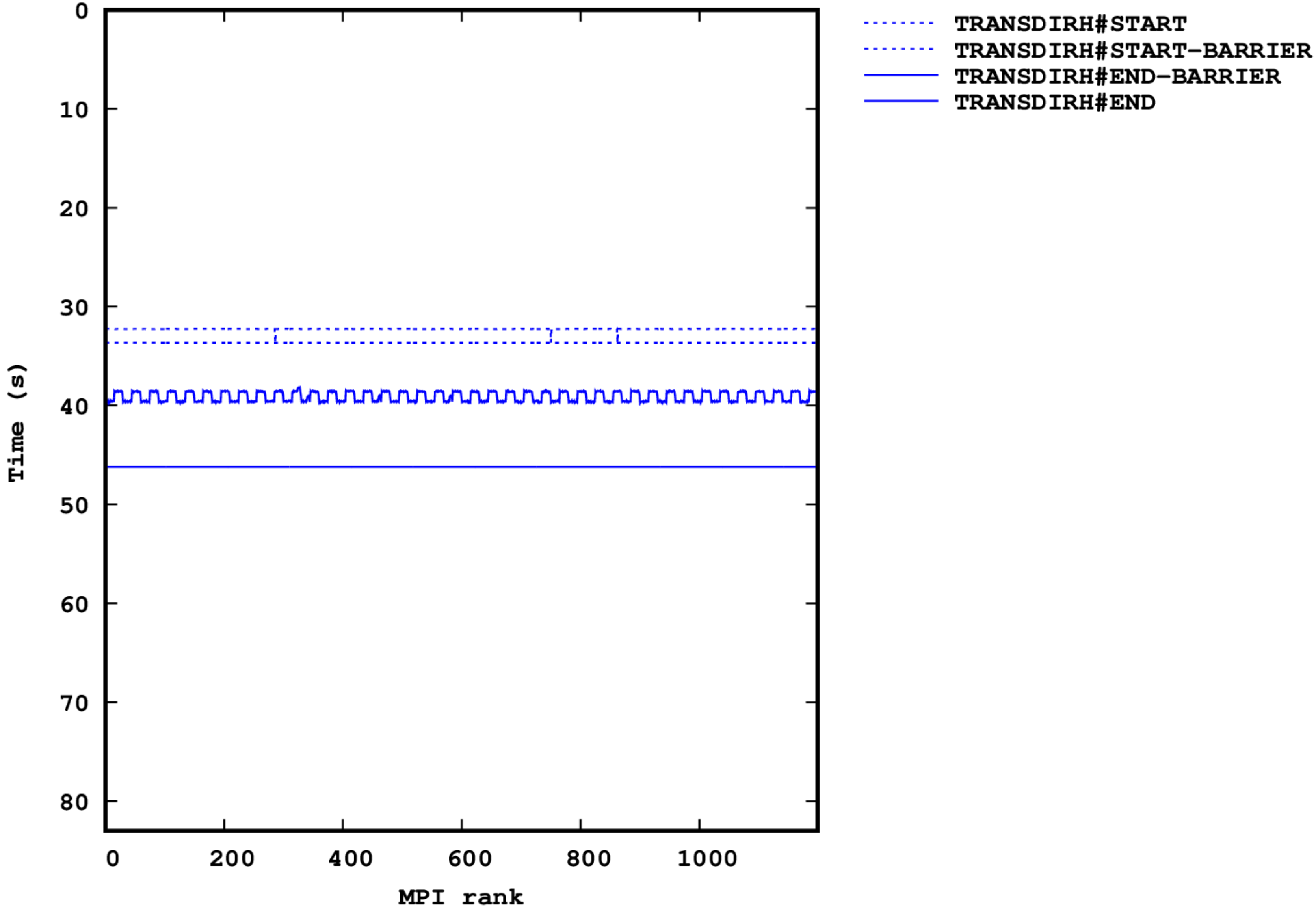


Jitter

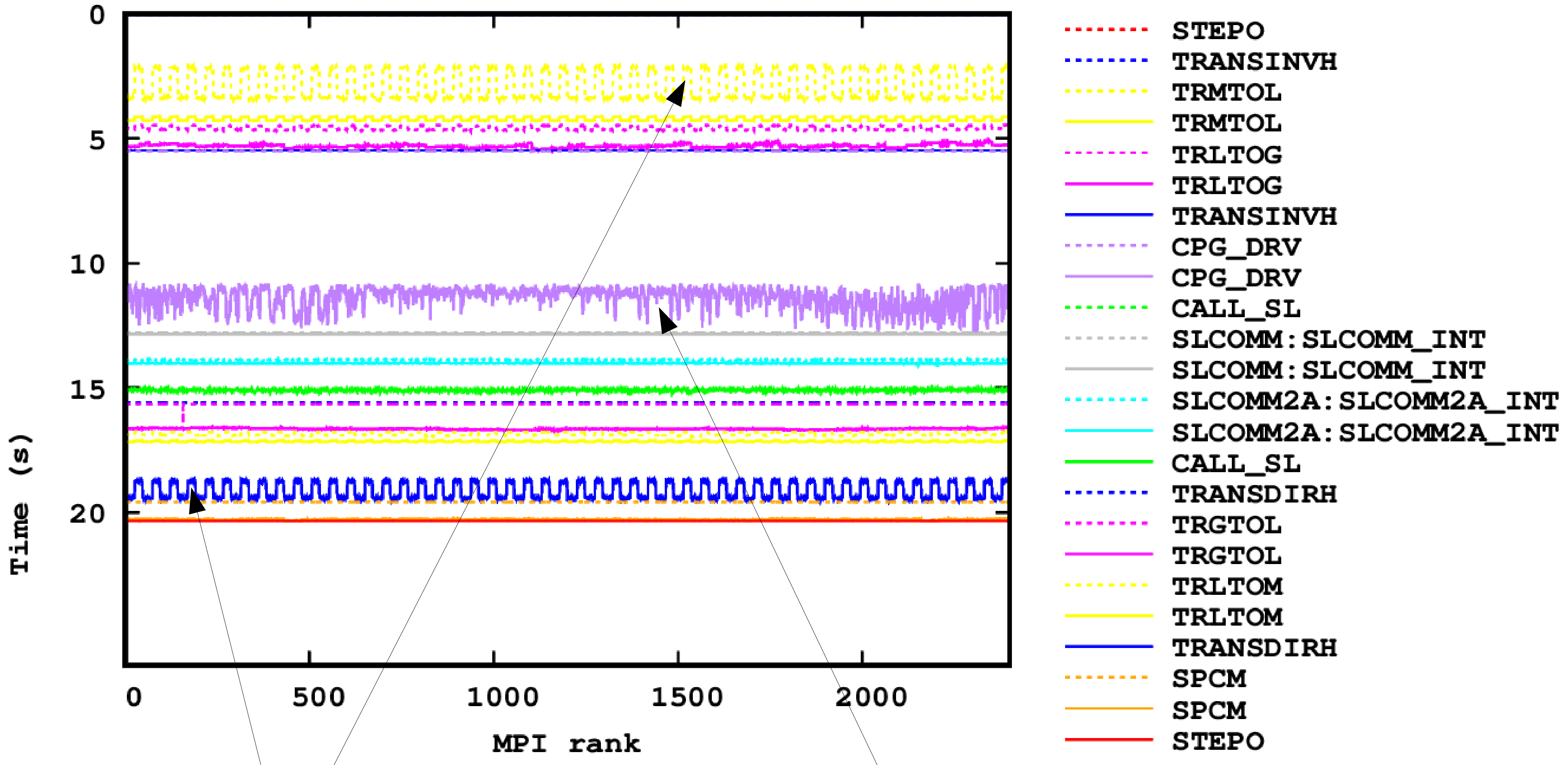


Delay because of barriers & jitter
Time (s)
MPI rank
0
10
20
30
40
50
60
70
80
0
200
400
600
800
1000
STEPO
TRANSINVH
TRMTOL
TRMTOL
TRLTOG
TRLTOG
TRANSINVH
CPG_DRV
CPG_DRV
CALL_SL
SLCOMM:SLCOMM_INT
SLCOMM:SLCOMM_INT
SLCOMM2A:SLCOMM2A_INT
SLCOMM2A:SLCOMM2A_INT
CALL_SL
TRANSDIRH
TRGTOL
TRGTOL
TRLTOM
TRLTOM
TRANSDIRH
SPCM
SPCM
STEPO

Jitter


TRANSDIRH#START
TRANSDIRH#START-BARRIER
TRANSDIRH#END-BARRIER
TRANSDIRH#END
Time (s)
0
10
20
30
40
50
60
70
80
0
200
400
600
800
1000
MPI rank

Load imbalance



Predictable
(Legendre transforms)

Unpredictable
(Physics)

Météo-France next HPC system

- Today : 2 x 1800 dual 20 cores Broadwell processor nodes
- In 2020 : budget x 2 → x 5 expected on computational power

Question :

Would it be possible in 2020 to run a global NH short range forecast ?

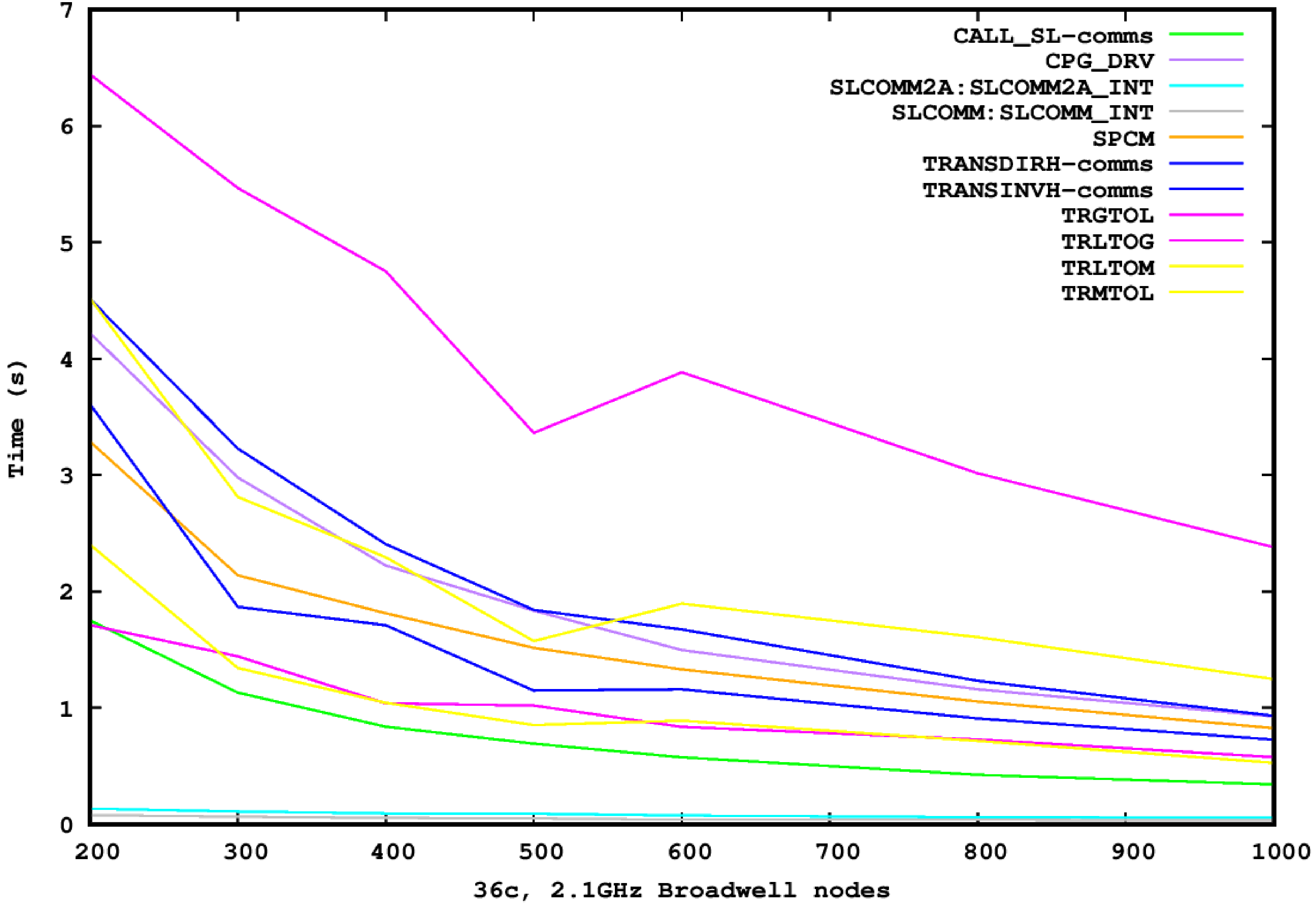
What will our next machine be like ?

Assume the increase in power (x 5) is equally split between:

- More powerful nodes (x 2.23)
- More nodes (x 2.23)

→ 3600 nodes/cluster with processors 2.23 more powerful than the 20 core 2.2GHz Broadwell

Tests on ECMWF cluster

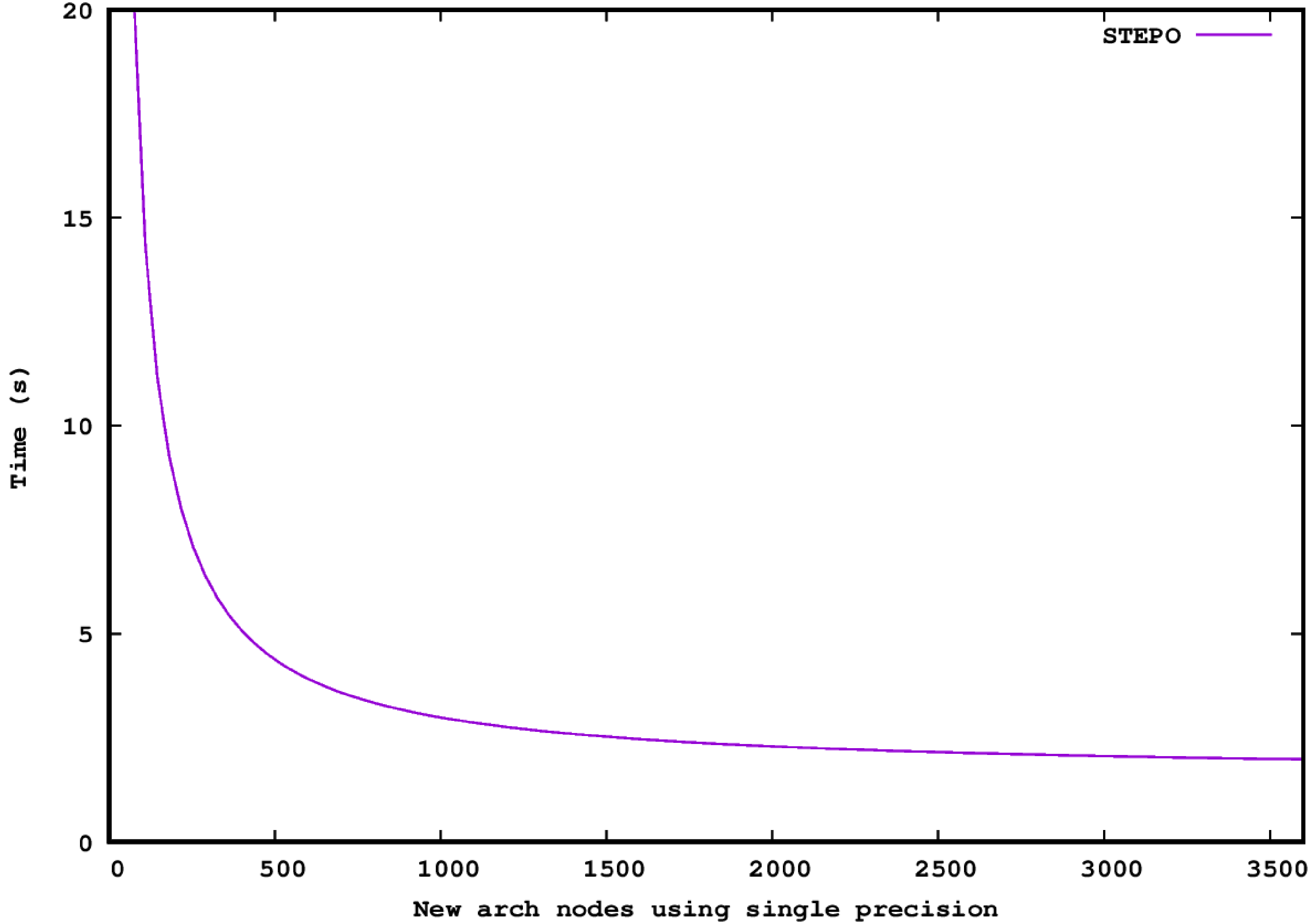


Extrapolating

- ECMWF : 36 cores, 2.1GHz → Météo-France 40 cores, 2.2 GHz ~ + 7%
- Reduce times in calculation routines by 2.23
- Keep times in communication routines (conservative)
- Extrapolate using Amdahl law
- Apply a gain of 45% using single precision

Cost of a time step

(New machine, single precision)



A short range global 2.5km forecast ?

Today, hours of forecast/hour:

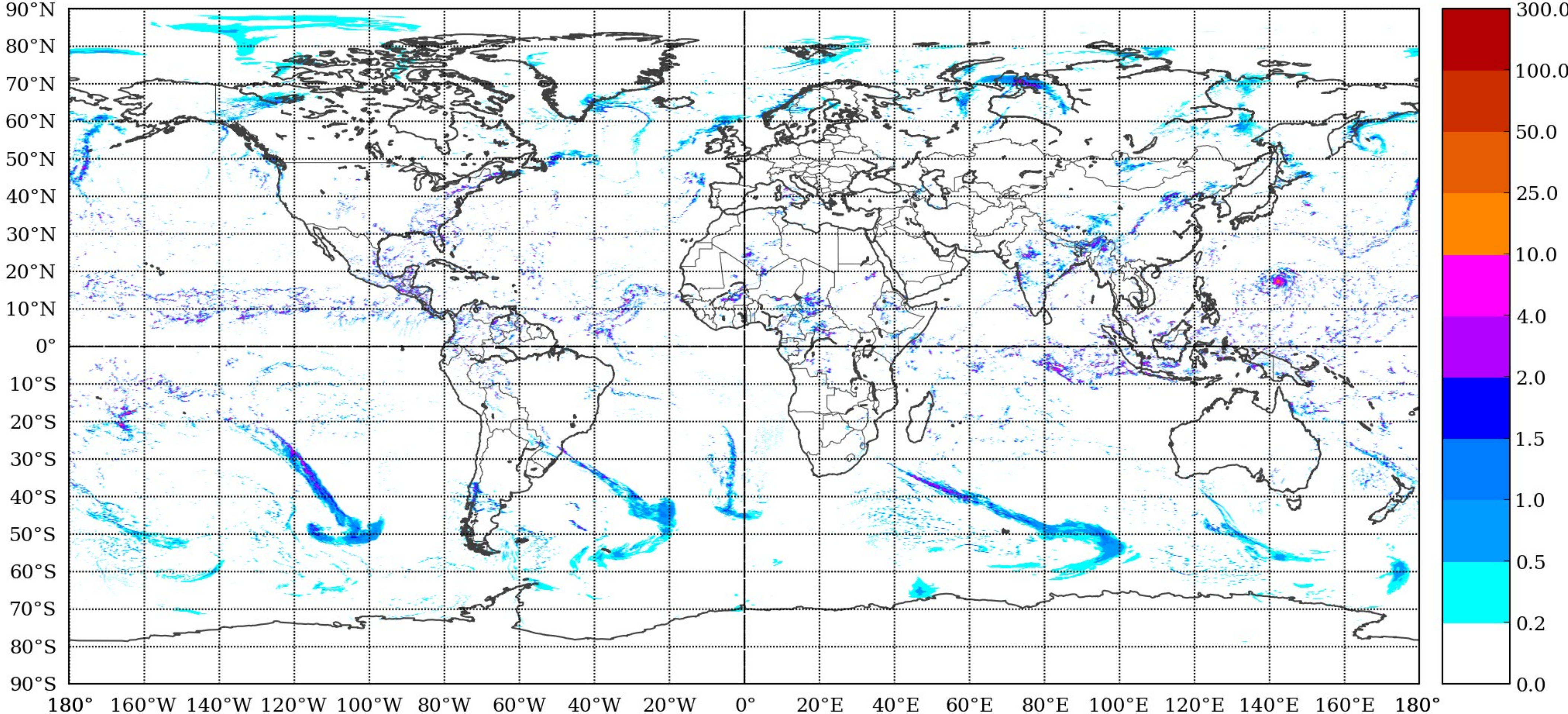
- ARPEGE 150h/h
- AROME 60h/h

$\Delta t = 100s$

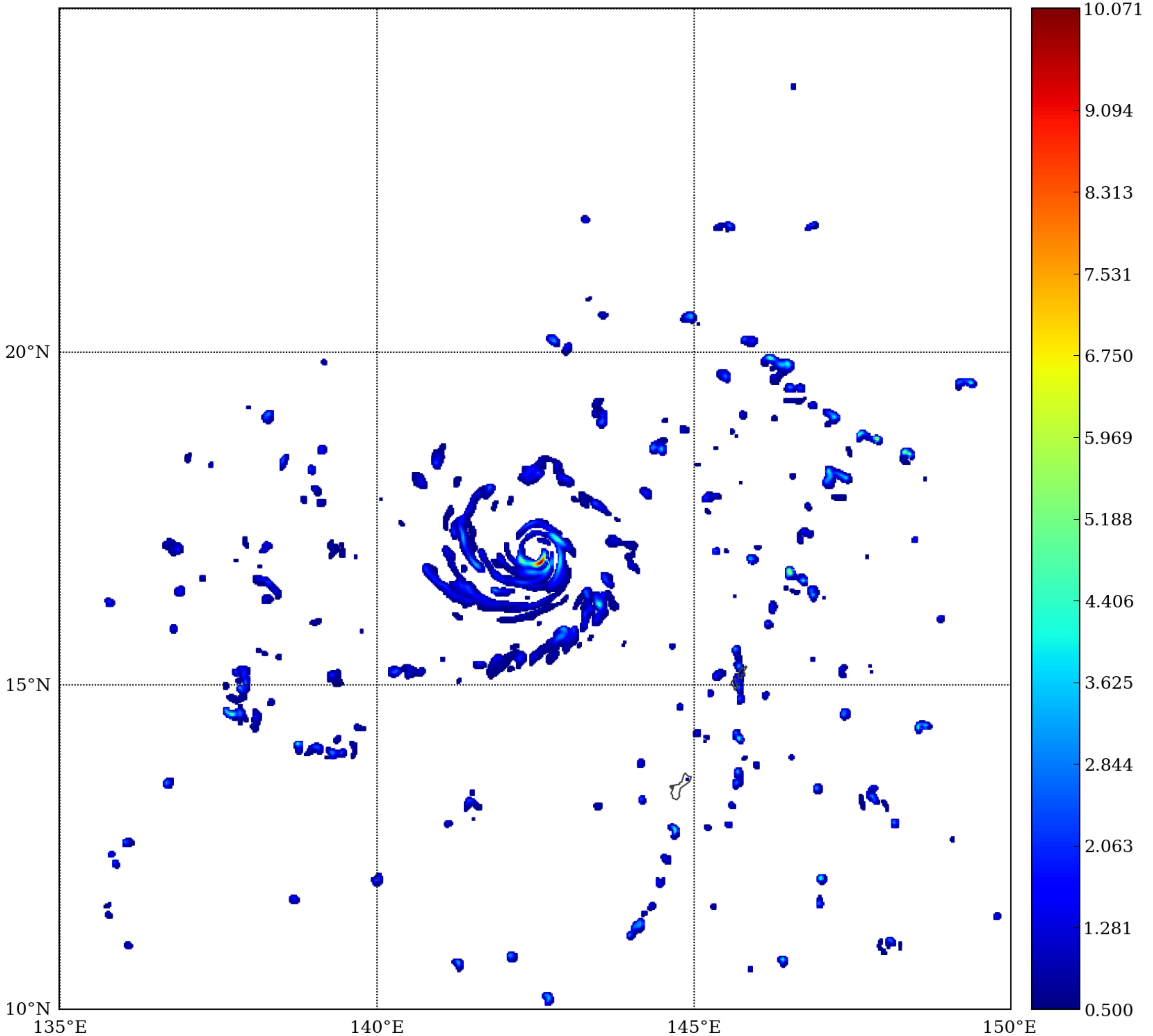
- 500 nodes : 5s elapsed time → 20h/h
- 1000 nodes : 3s elapsed time → 30h/h

A 24h hours forecast in 45 minutes looks feasible

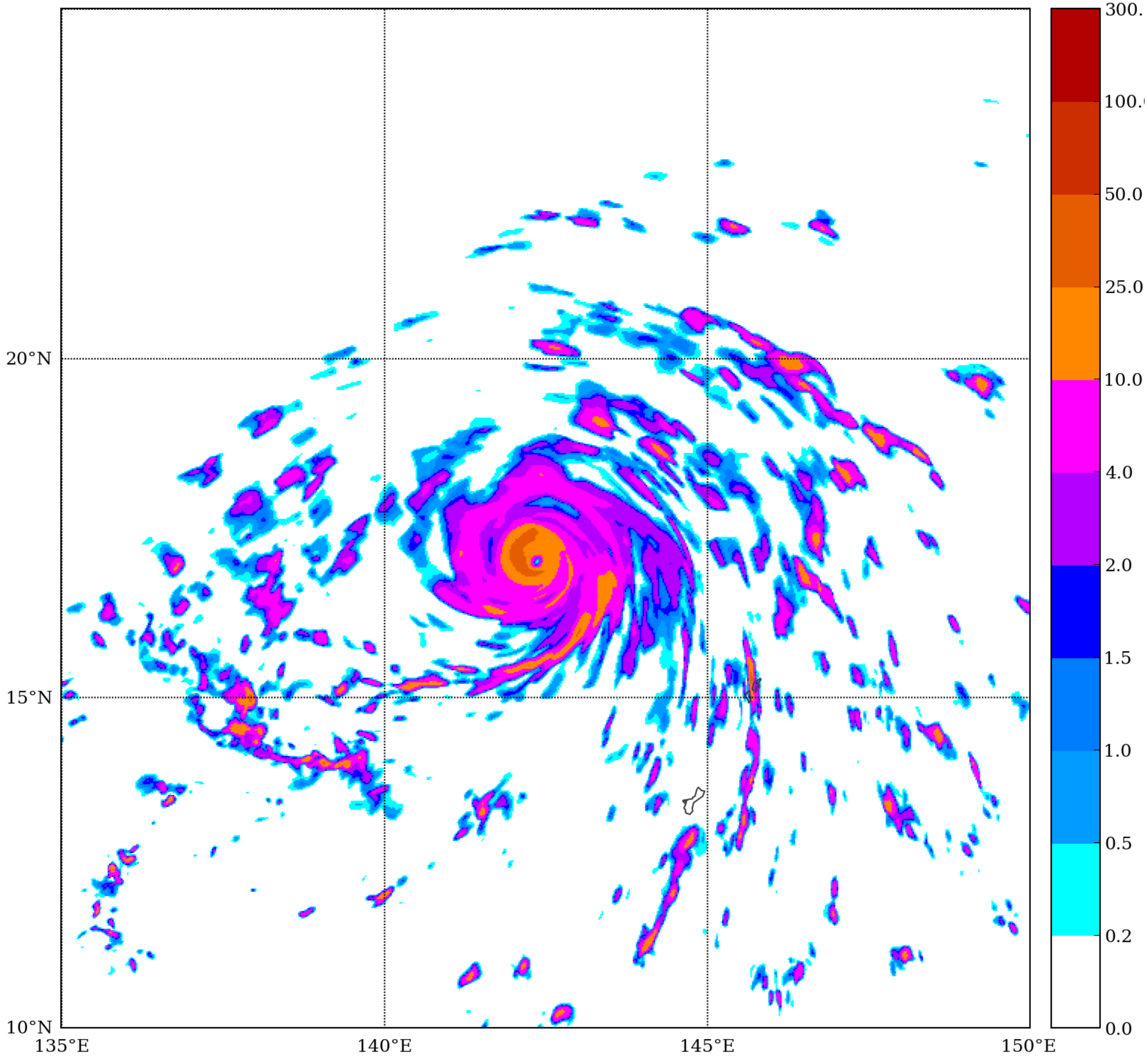
Global precipitation field at 2.5km



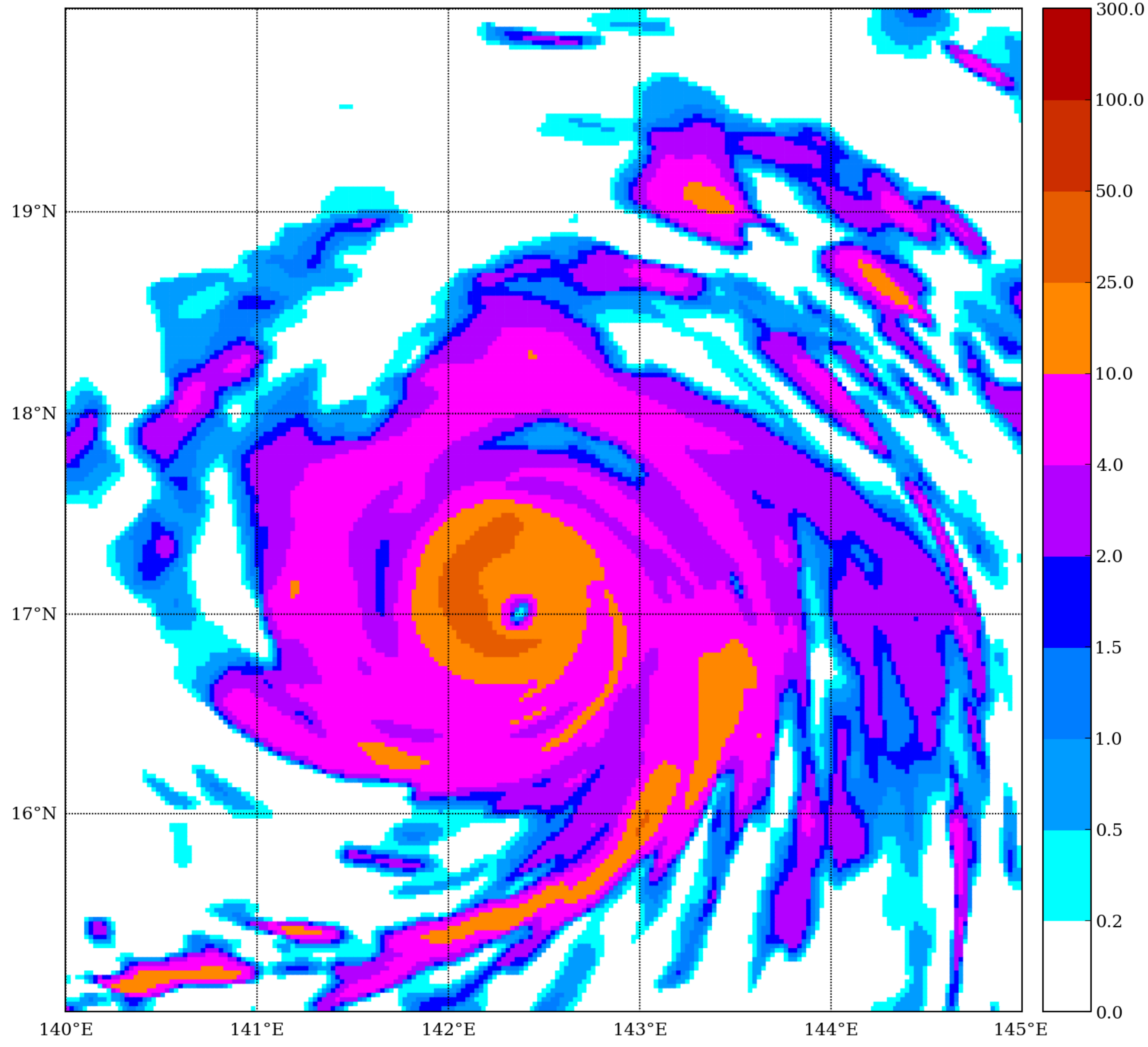
Precipitation field at 2.5km (zoom)



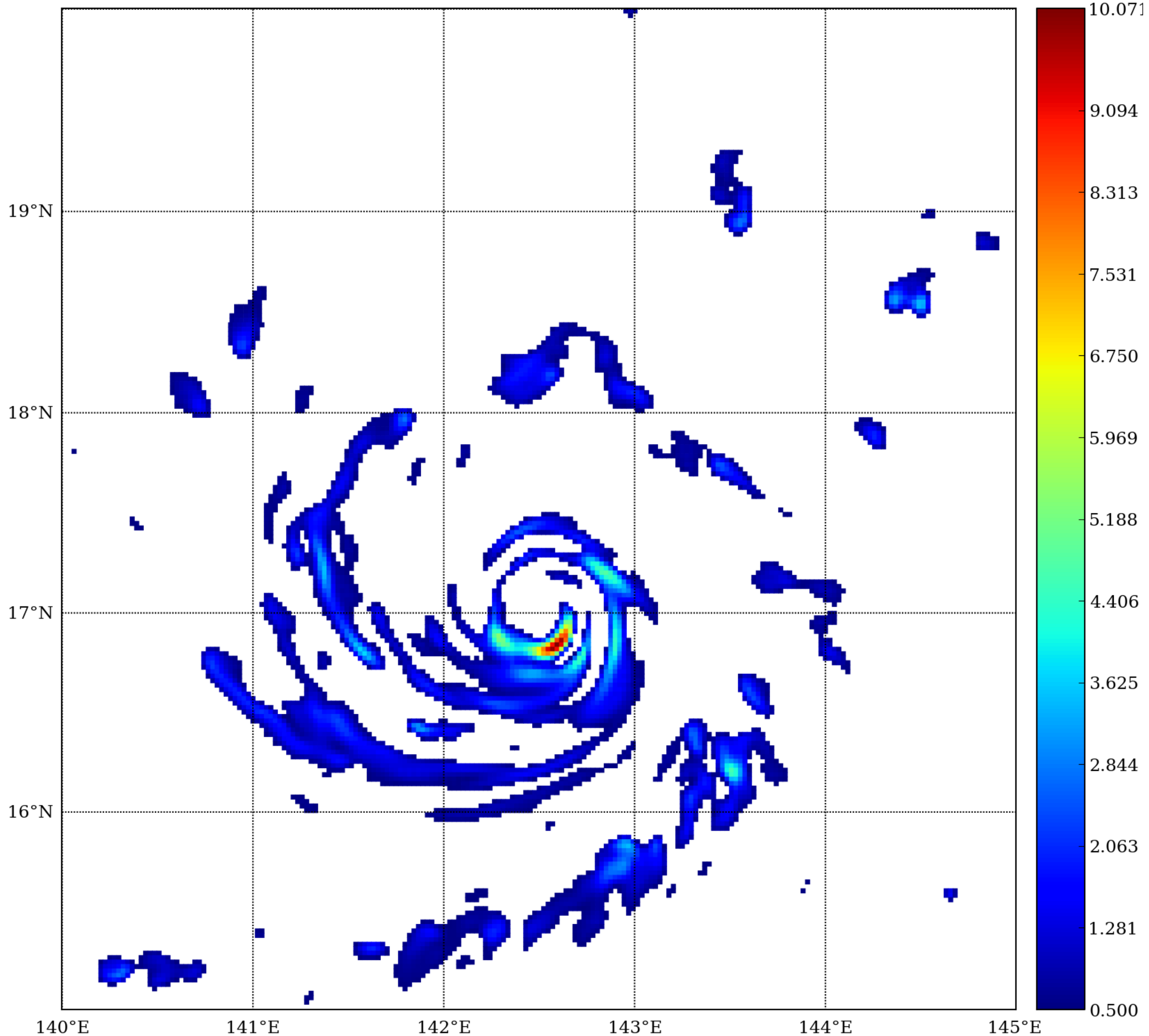
Precipitation field at 2.5km (zoom)



Precipitation field at 2.5km (zoom)



Vertical velocity



Conclusion & perspectives

- Local formulations for derivatives are currently tested with LAM model AROME using an iterative solver for implicit system: this would reduce communications but at the price of using a grid-point solver (Thomas Burgot).
- Local formulations could be also tested with the global geometry, thereby avoiding the costly Legendre transforms (Pierre Bénard).
- New options for the NH kernel are also tested (quasi-elastic formulation, new formulation for vertical velocity equations). Quasi-elastic seems to be less stable (Fabrice Voitus).

Conclusions & perspectives

- Using single prec. will lead to a 45% saving in elapsed time
- Try AROME physics
- Make SURFEX work (PGD & PREP)
- 2.5km ARPEGE NH will certainly be used by researchers on our new system.
- Our dynamics will probably meet performance requirements for operations for the next decade