

Outcomes of Météo-France RAPS 2024 & Overview of the Benchmark 2025

Ryad El Khatib / Météo-France
5th ACCORD ASW
Zalakarosn 31 March - 4 April 2025

Plan

Outcomes of the RAPS 2024 exercise

- Recall of the aims and content
- Various results

Overview of the Benchmark 2025

- Purpose and challenges
- Content
- Timeline

R.A.P.S. : « Real Applications for Parallel Systems »

- A simplified benchmark to let all the concerned industry partners to be prepared
- Proposed to all industry partners in exchange of the signature of a Non-disclosure Agreement.
- Enable open discussions with industry partner – something not possible during the official benchmark procedure
- Valid from February 2024 to November 2025

RAPS content (recall)

AROME forecast without post-processing

- Cubic grid (750 m resolution), truncation 624 - 674
- 120 vertical levels
- In single precision
- For CPU chips only
- Performance required : 26 min/24h [needs 110 nodes on Belenos]

ARPEGE forecast without post-processing

- Cubic grid (truncation T1799)
- 120 vertical levels
- In single precision
- One source-code for CPU chips, another one for Nvidia GPU chips
- Performance required : 30 min/102h
[CPU version needs 189 nodes on belenos]

General feedback on the RAPS

OBJECTIVE OBSERVATIONS :

- **8 companies** have provided a benchmark report last summer
- 3 other companies started late or didn't report
- **10 compilers** of different origins have been tested.
- CPU chips manufactured by **Intel**, **AMD**, **ARM**, and **NEC** Vector Engines have run the applications – but not always at high resolution or in the requested speed, by lack of means.
- **Nvidia GPU chips** have run ARPEGE.

SUBJECTIVE OBSERVATIONS :

- Much differences in human resources or computing resources investments from one company to another
- Certain companies collaborate with one another, others go it alone
- Certain companies interacted a lot with us, others not.
- There is a wish to concurrence Nvidia in the domain of GPU.

A few detailed results of the RAPS

SOFTWARE

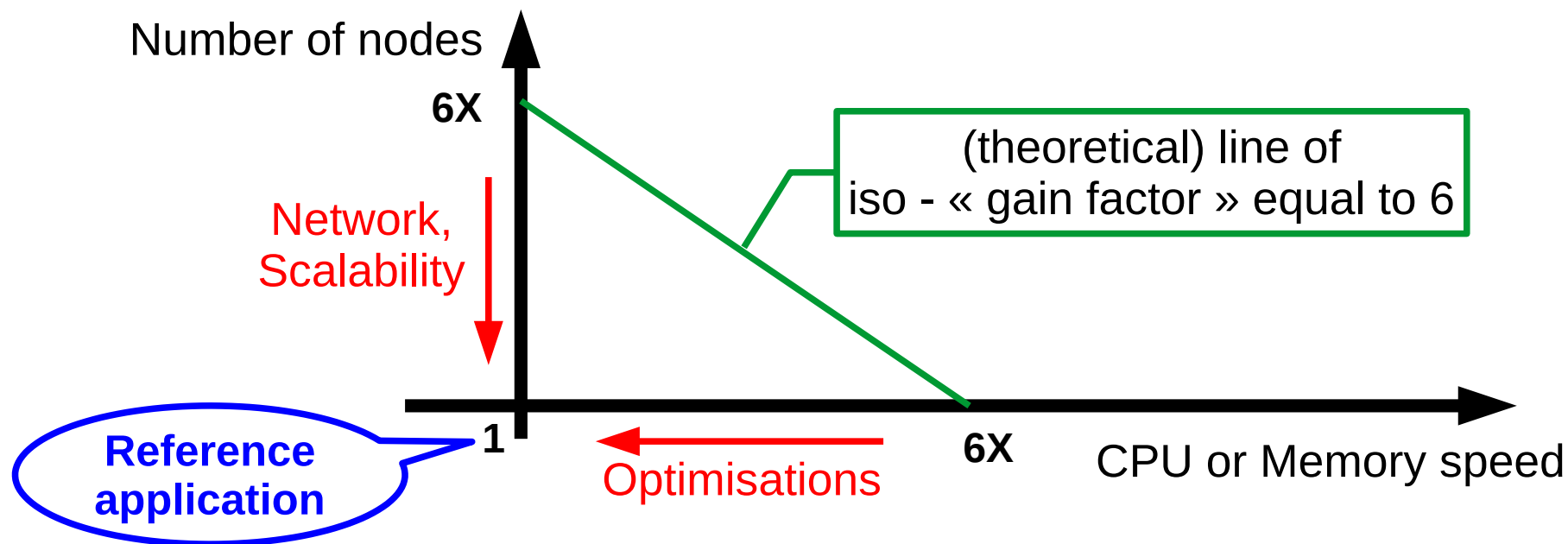
- 30 portability fixes for compilers
- A few compiler bugs have been reported and fixed or bypassed
- 2 runtime failures (depending whether the compiler is permissive) have been fixed
- One company has provided a few code optimizations

HARDWARE

- The forecasting models on CPU are seen memory-cheap
- There is a rather good balance in the dependency of the models against the main hardware components (chip, memory, network)
- The model ARPEGE on GPU (in the state of the art at that time) is too slow and too memory-expensive to compete with its CPU version.

Purpose of the benchmark

Calibrate with application performances measurements
the relative power of the future machine compared to the one of today.
This « relative power » is synthetized as a single number : the « gain factor »



Reference applications should at the same time :

- ***Represent what will run on the next machine***
- ***Be efficient on the current machine***

Challenges

We want CPUs for NWP, GPU for NWP, GPU for AI, but ...

- How to benchmark a NWP model on GPU if it cannot compete with CPU ?
 - How to estimate the performance of a GPU against a CPU ?
 - What if AI applications prevail over NWP on GPU in a few years ?
 - Are we experienced enough to benchmark AI applications today ?
(performance, network, IO specifications)
 - How to deal with the differentiation between GPU for HPC and GPU for AI ?
 - How to stimulate or maintain the competition between GPU vendors ?
- Etc ...

Strategy for now :

- A portability benchmark for NWP on GPU (ARPEGE)
+ a performance benchmark for GPU with AI applications
- Renewal of HPC in two phases :
 - 1) A large part in CPU + a smaller part in GPU
 - 2) Purchase order to extend the clusters, proportions CPU/GPU to be defined

Benchmark content for CPU

AROME forecasting model without post-processing (*like in the RAPS*)

- Requested elapse time : 26 min/24h. [*Belenos : 100 nodes*]*

ARPEGE forecasting model without post-processing (*like in the RAPS*)

- Requested elapse time : 30 min/102h. [*Belenos : 165 nodes*]*

Hybrid 4DVAR ARPEGE (under OOPS)

- Same geometry as the forecasting model
- In double précision
- 2 minimisations en troncation T499, cubic grid
- Same observations (amounts and kinds) than in operations today
- 200 members
- Requested elapse time : 35 min. [*Belenos : 44 min/224 nodes*]

* better than in the RAPS thanks to optimizations and parameters tuning

Benchmark content for GPU

ARPEGE forecast on GPU (*state of the art*) for portability testing

- Same high geometry than the version for CPU
- Source-code ready for Nvidia GPU chips
- Requested elapse time : 45 min/24h.*

AIFS AI application (learning)

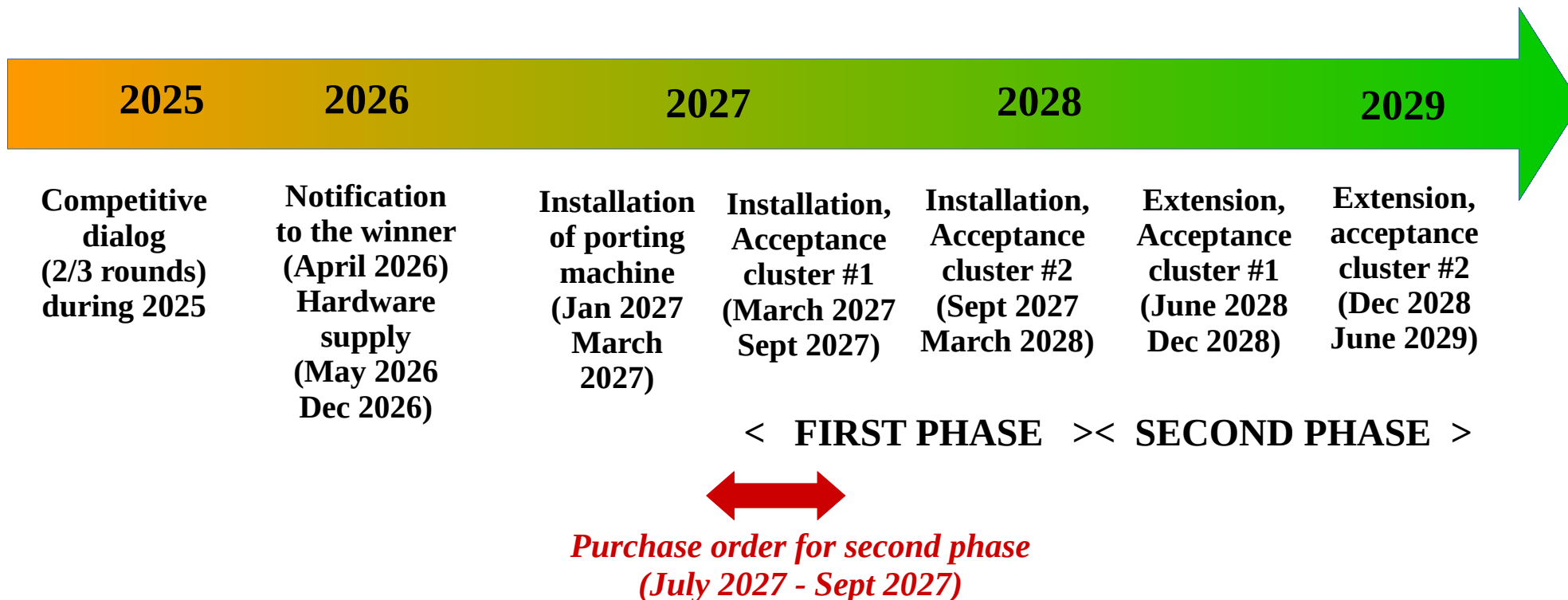
- Based on ECMWF AIFS model
- With a relatively small dataset (175 GB)
- Designed to test intra-node GPU performance and communications

Py4Cast AI application (learning)

- Based on Météo-France Py4Cast model
- With a large dataset (15 TB)
- Designed to test multi-node GPU performance and I/O

** Large value because it must not be a difficult target*

Timeline of the HPC renewal



Thank you for your attention