

# Nouvel accroissement de la puissance de calcul de Météo-France

Le nouveau système de calcul intensif de Météo-France est entré en production opérationnelle début février 2021 : il est composé de 2 supercalculateurs quasi identiques, d'une puissance de 10,7 Pétaflops chacun, installés dans 2 salles de calcul distinctes, en région Toulousaine. Le fournisseur de la solution est la société ATOS, dans le cadre d'un contrat de service d'une durée ferme de 4 ans, prolongeable de manière optionnelle pendant 12 mois supplémentaires.

Ce choix technologique a été opéré à l'issue d'un dialogue compétitif de 18 mois sur lequel les principaux industriels s'étaient positionnés. Les critères de choix reposaient principalement sur la performance de la machine mesurée sur un jeu de codes représentatif de la production de Météo-France (essentiellement des modules d'ARPEGE et d'AROME), sur les éléments de résilience et de robustesse prévus dans l'architecture, ainsi que sur l'efficacité énergétique de la

configuration. Le dialogue compétitif a ainsi permis d'auditionner les industriels à deux reprises (une fois à Toulouse, une seconde fois aux États-Unis, en marge du salon *Super-computing* où toutes les entreprises majeures en matière de calcul intensif étaient présentes). Cette seconde audition, initialement non prévue, s'est *in fine* révélée nécessaire au vu de la grande instabilité du marché des composants informatiques, tant en matières de plannings de sorties de nouvelle génération de processeurs ou de composants mémoires, que de coût des équipements eux-mêmes. Jadis, un fournisseur de supercalculateur (à l'instar de *Cray*, *Fujitsu* ou *NEC*) maîtrisait en quasi-totalité toute la ligne de production dont il avait besoin pour construire la machine qu'il proposait. Il lui était donc relativement facile de définir le calendrier cible d'apparition sur le marché d'une nouvelle solution, ainsi que son coût. Aujourd'hui, ce temps est révolu : chaque industriel proposant un supercalculateur fait appel à un très grand nombre de partenaires, y compris pour des aspects clés comme la fourniture du processeur, du réseau interne ou des compilateurs.



Photo 1 : calculateur Taranis

La maîtrise, par le vendeur, des calendriers et des coûts est donc très délicate et souffre, parfois, de revirements inattendus. Ce fut notamment le cas entre mi-2018 et le printemps 2019, où certaines lignes de produits ont subi des retards considérables, remettant totalement en cause la solution initialement proposée par le candidat.

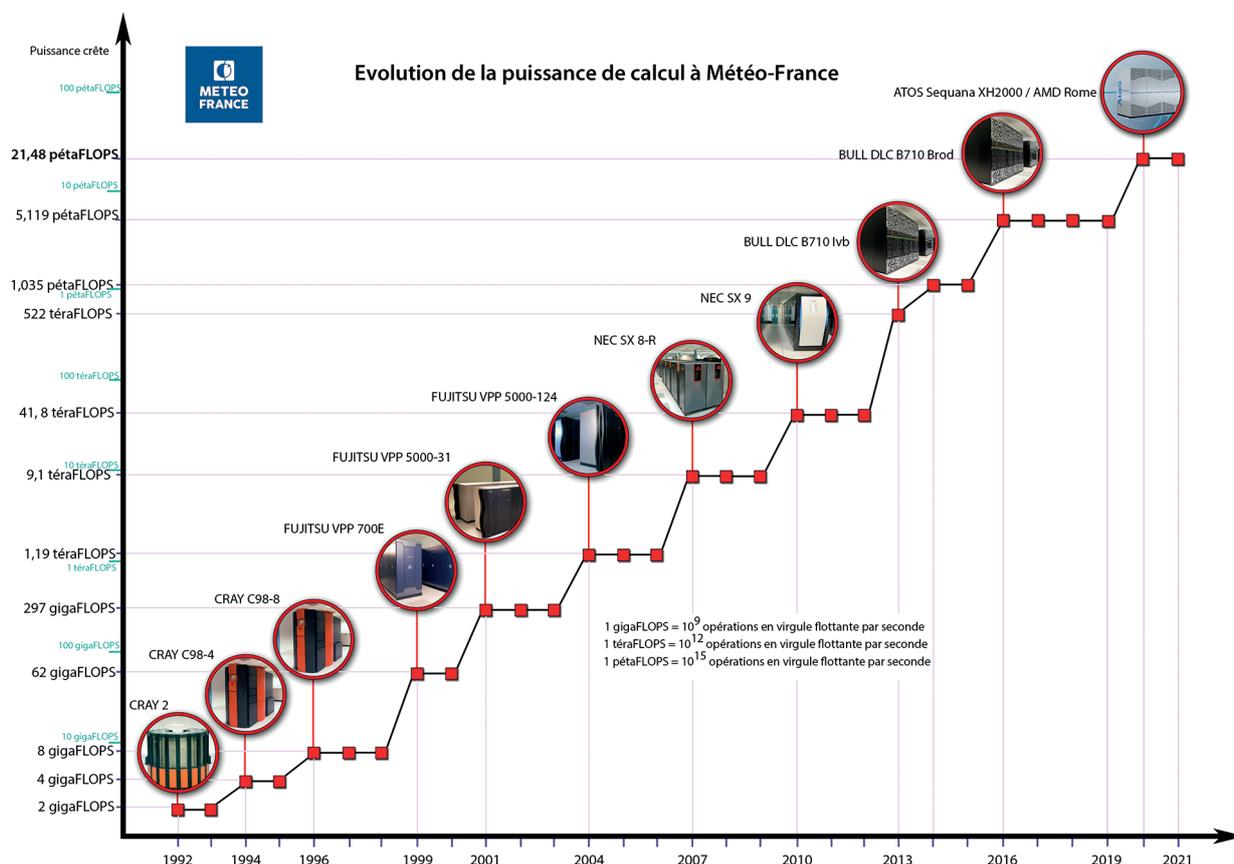
*In fine*, après une compétition intense et très serrée, le marché de fournitures de la nouvelle configuration de calcul intensif a été notifié à la société ATOS. Il s'agissait-là d'une première, au sens où c'était la première fois à Météo-France que le titulaire du marché « Calcul » sortant était reconduit à l'occasion de l'appel d'offre suivant. La solution retenue repose sur la plateforme *Sequana XH2000*, dont la fabrication est réalisée à Angers par BULL (filiale HPC du groupe ATOS). Le composant de base de

la configuration est le processeur AMD "Epyc Rome". Ce CPU d'architecture x86 est à la pointe de la technologie aujourd'hui disponible (finesse de gravure de 7 nanomètres pour une fréquence d'horloge à 2,25 GHz). Les différents nœuds de calcul sont reliés entre eux par un réseau interne fourni par la société israélienne Mellanox, et les principaux outils de compilation utilisés proviennent de la suite logicielle d'INTEL. Enfin, l'essentiel du stockage disque directement attaché à la machine est issu du fournisseur américain Netapp. Chaque nœud de calcul dispose de 128 cœurs se partageant une mémoire de 256 Go. La performance d'accès à la mémoire est toutefois non uniforme, et dépend de la localisation du cœur au sein du nœud. Au total, chaque calculateur comprend 295 296 cœurs de calcul : afin de tirer pleinement bénéfice du potentiel de ces machines, il faut donc que les équipes de recherche et de développement réalisent des prouesses en matière de parallélisation des codes et

d'équilibrage de charge. Afin que cette puissance de calcul considérable soit utilisée de manière optimale, un ambitieux programme de formation a été mis en place début 2020 : près de 150 utilisateurs issus de Météo-France et de ses partenaires (Mercator-Océan, Cerfacs, Shom, Ineris) ont été formés à la bonne utilisation des nouveaux supercalculateurs.

La complexité de l'architecture retenue, conjuguée au manque de maturité de certains composants et à des exigences spécifiques de Météo-France (afin notamment de viser une empreinte environnementale minimale) ont sérieusement compliqué et allongé la période d'installation et de mise au point du calculateur. Et ce, d'autant plus que la majeure partie du travail devait se faire dans des conditions particulières, puisque nous étions alors au plus fort de la crise sanitaire de l'année 2020, ce qui ne facilitait pas les échanges entre les équipes. Toutefois, grâce à un travail acharné, à la fois coté ATOS

Image 1 : Courbe d'évolution de la puissance de calcul à Météo-France



et coté Météo-France, la stabilisation du fonctionnement des 2 calculateurs a pu être atteinte en fin d'année 2020.

Une des difficultés majeures de cette période a été d'atteindre un niveau élevé de robustesse des calculateurs parallèlement à la solution de refroidissement retenue : les nœuds de calcul des supercalculateurs sont maintenus à une température stable via un système de refroidissement liquide direct (mélange eau/glycol) à température relativement élevée (48 à 58°C). Pour ce faire, des échangeurs sont situés en bas de chacune des armoires de calcul, et assurent le couplage entre la production d'eau venant de la logistique du site, et le mélange « eau/glycol » qui va circuler en moyenne à 53°C dans les composants de calcul eux-mêmes. L'eau qui alimente ces échangeurs arrive ainsi à une température de 39°C dans la machine et en ressort à 45°C. Ce régime d'eau a un double avantage :

- il permet d'éviter l'utilisation onéreuse de groupes froids (qui sont indispensables, non seulement en cas de refroidissement air, mais aussi dans les solutions antérieures de refroidissement à eau où nous étions à des régimes de 15/20°C). Avec une cible à 39°C, des refroidisseurs adiabatiques sont suffisants, ils sont nettement plus sobres et moins onéreux que des groupes froids.
- il rend aussi possible, en période hivernale, la récupération directe de la chaleur fatale issue de la machine pour chauffer les bâtiments.

Ces contraintes de fonctionnement, sur lesquelles Météo-France avait challengé les compétiteurs, afin de limiter son empreinte carbone et de contenir les coûts d'électricité sur la période d'exploitation de la machine, se sont révélées délicates à mettre en œuvre par ATOS. Météo-France était le premier client pour lequel un tel engagement avait été pris par cet indus-

triel (ce que nous avons découvert en cours d'année 2020), d'où des incidents multiples en début d'installation et des difficultés de réglages complexes à résoudre. Pour autant, début 2021, après un reconditionnement sur site d'une partie des composants, un mode opératoire permettant de conjuguer efficacité énergétique et robustesse technologique a pu voir le jour. Cette étape décisive a permis d'opérer la bascule de production sur les nouveaux calculateurs le 2 février (au CNC) et le 9 (à l'ECA), et procéder à l'arrêt quasi simultané des 2 anciennes machines (les anciens calculateurs Beaufix et Prolix ont été définitivement stoppés le 11 et le 15 février).

Le facteur de gain réellement mesuré sur les applications de prévision numérique du temps durant la phase de « miroir » (période précédant la bascule opérationnelle, durant laquelle on fait tourner sur l'ancien et le nouveau calculateur des versions d'Arome et Arpege scientifiquement identiques) est très légèrement supérieur à 5. Ainsi, l'objectif qui était affiché à cet égard par Météo-France dans le cadre de la COP 2017-2021 a pu être tenu, malgré un contexte technologique peu favorable : la « *Loi de Moore* » (qui permettait jadis d'avoir, peu ou prou, un doublement de la puissance de calcul à budget constant tous les 18 mois) a désormais vécu, les progrès en matières de miniaturisation des composants étant maintenant très lents. Pour gagner en performances, il faut donc désormais travailler sur le parallélisme massif, sur la localisation des données et la minimisation des transferts sur les réseaux internes ... sans négliger de renforcer le budget dévolu au calcul intensif.

Cet accroissement de puissance de calcul a permis de repositionner Météo-France au niveau international. Les nouvelles machines sont actuellement classées aux alentours de la 50<sup>e</sup> place du classement inter-

national TOP500 des 500 calculateurs les plus puissants au monde. Ce gain de performance va notamment permettre à Météo-France, au printemps 2022, de porter la résolution des modèles ensemblistes opérationnels de prévision numérique du temps à celle des modèles déterministes actuels (à savoir 1,3 km), et de faire également bénéficier nos départements et territoires d'Outre-mer de prévision à une taille de maille équivalente à celle de la métropole. À plus longue échéance, en 2023/2024, la mise en œuvre de systèmes de prévision à une résolution de l'ordre de 500 m est envisagée sur des zones à forts enjeux.

Mais Météo-France doit, d'ores et déjà, se préparer à l'étape suivante : à échéance 2024/25, un saut technologique majeur sera probablement indispensable pour que la progression de la puissance de calcul puisse se poursuivre. Des actions de grande envergure sont en cours de lancement pour initier le portage de nos principaux codes de calcul sur technologie GPU (processeurs graphiques), voie qui semble parmi les plus prometteuses pour le futur. Mais il s'agit là d'un défi de grande ampleur : le travail à réaliser est considérable, et ne pourra être mené à terme qu'en coopération avec nos partenaires nationaux et internationaux. 🌈

**ALAIN BEURAUD,**  
**DIRECTION DE LA STRATÉGIE, MÉTÉO-FRANCE**