

HPC : LE CALCUL HAUTE PERFORMANCE DE LA R&D D'EDF AU SERVICE DU GROUPE

Dès le début des années 1980, la R&D d'EDF se lançait dans la recherche *in silico* avec un célèbre Cray, la marque des ordinateurs les plus puissants du monde à l'époque. Un quart de siècle plus tard, de nombreux chercheurs d'EDF utilisent l'un des supercalculateurs internes, dont le dernier né, baptisé Gaïa, pointait dans le Top 100 mondial en puissance lors de sa mise en service début 2019. [Page 02](#)

CAHIER TECHNIQUE

UN NOUVEAU CALCULATEUR POUR EDF LABS

Avec Gaïa, EDF s'offre un supercalculateur HPC non seulement puissant mais aussi adaptable grâce à son architecture. [Page 05](#)

HPC ET ÉTUDES DES MATÉRIAUX : 20 ANS D'HISTOIRE COMMUNE

Dès les années 90, les ingénieurs du département Matériaux et Mécanique des Composants de la R&D d'EDF se sont initiés au calcul parallèle sur les ordinateurs de l'époque, déjà partagés avec l'ONERA, le CEA, et la DGA. [Page 10](#)

CODE_ASTER S'OFFRE UNE MEILLEURE SYNERGIE AVEC LE HPC POUR SES 30 ANS

Initié en 1989, Code_Aster (pour Analyse des Structures Thermo-mécaniques pour des Études et des Recherches) est à la fois un moyen de recherche pour la R&D et un outil d'exploitation quotidien pour l'ingénierie d'EDF. [Page 12](#)



3 QUESTIONS À AUDE PELLETIER

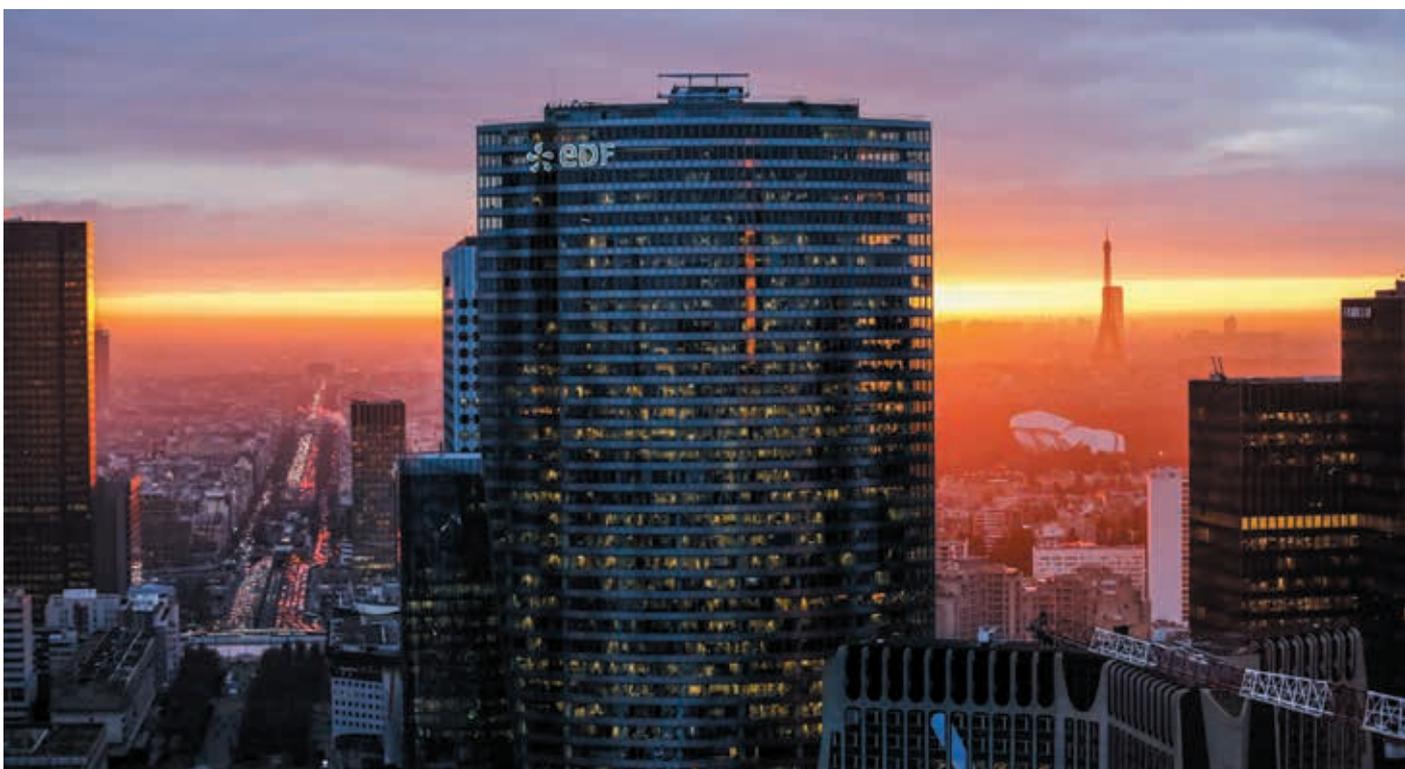
Chef de département délégué OSIRIS - Optimisation, Simulation, Risques et Statistiques pour les marchés de l'énergie. [Page 14](#)



HPC : LE CALCUL HAUTE PERFORMANCE DE LA R&D D'EDF AU SERVICE DU GROUPE

C'est en 1993 que la R&D d'EDF se lançait dans la recherche *in silico*¹ avec un célèbre Cray, la marque des ordinateurs les plus puissants du monde à l'époque. Un quart de siècle plus tard, de nombreux chercheurs d'EDF utilisent l'un des supercalculateurs internes, dont le dernier né, baptisé Gaïa, pointait dans le Top 100 mondial en puissance lors de sa mise en service début 2019. Le Calcul Haute Performance (High Performance Computing ou HPC) est en effet l'outil roi pour améliorer la précision des simulations, booster la rapidité des calculs de risques, ou encore l'efficacité des algorithmes d'Intelligence Artificielle.

QUALITÉ, PRÉCISION, SÛRETÉ : LES GRANDS OBJECTIFS DU HPC



Une stratégie initiée et guidée par les grands codes de calcul

Dès la fin des années 80, EDF avait compris l'importance d'investir dans le développement de logiciels de calcul scientifique à la fois performants et ouverts (open source), afin de répondre de manière efficace et transparente aux demandes d'informations des pouvoirs publics (en particulier l'Autorité

de Sécurité Nucléaire). Plusieurs codes de calcul ont ainsi été développés, et le sont encore aujourd'hui. « Cette politique de grands codes a rapidement nécessité des moyens de calcul significatifs en puissance crête², et a conduit au regroupement des clusters existants, à travers **une large démarche de mutualisation** : d'abord au sein de la R&D, puis avec le Groupe EDF, en particulier l'ingénierie nucléaire et Frama-

tome ; et probablement avec des partenaires extérieurs à l'avenir » explique Stéphane Tanguy, Directeur des Technologies et Systèmes de l'information à la R&D d'EDF. Cette mutualisation présente également des avantages économiques (investissement partagé), et organisationnels : les calculateurs HPC d'EDF affichent ainsi **un taux d'occupation de l'ordre de 80 %**.

1- Expériences virtuelles, réalisées via des moyens de calcul informatisés.

2- Puissance maximale, délivrée dans des conditions optimales et standardisées.

HPC : OBJECTIF QUALITÉ

L'apport du HPC pour les projets est de plusieurs natures, et notamment lié aux caractéristiques de ces machines :



Mémoire disponible par cœur



Processeur de type Gold ou Platinum



Stockage parallèle pour une écriture et une lecture très rapide des fichiers



Nombre de cœurs mobilisables par un utilisateur (> 40 000 sur Gaïa)



Réseau à faible latence pour le parallélisme de code



Forte disponibilité du système liée à son architecture

Ceci se traduit par des gains notables sur :

Les temps de réalisation des études car :

- La durée d'étude est réduite à périmètre égal.
- Le périmètre des études peut être étendu sans augmenter les temps de calcul : prise en compte d'un plus grand nombre de paramètres / contraintes / variables dans les modèles.
- Un grand nombre d'études peuvent tourner en même temps (par exemple lors des Grands Challenges avec plus d'1 million de calculs en 24 h)
- Les ressources attribuées à chaque calcul sont plus importantes, avec 25 000 cœurs de calcul reliés entre eux de manière performante
- Le post traitement est réalisé directement sur les moyens de calcul, sans transfert de données, et pour de grands volumes de données post-traitables.

La qualité des études (précision...) car :

- Des détails géométriques d'ordres de grandeur plus petits que la géométrie globale peuvent être pris en compte (ex : des détails géométriques de l'ordre d'un centimètre dans une géométrie de l'ordre du mètre ou de la dizaine de mètre).
- La précision des calculs est accrue. Elle permet par exemple de passer d'un pas de temps de 30 min à un pas de temps de 10 min, ou de simuler une durée supérieure (passer de 1 an de simulation à 3 ans).
- Plus de paramètres peuvent être inclus dans les études de sensibilité.

Calculer plus vite pour calculer mieux

Sans surprises, Gaïa est beaucoup plus puissant que son aîné : 3 pétaFlops contre 1 pour EOLE, lancé en 2017. Merci à la loi de Moore³ ! « C'est tout de même un peu plus compliqué⁴, tempère Stéphane Dupré la Tour, Directeur Innovation Avancée et Technologie Numérique à la R&D d'EDF. Le super-calculateur doit surtout répondre aux besoins des différents départements de la R&D et in fine à ceux des métiers.

La puissance de calcul n'est pas un objectif en elle-même, ce sont les applications qui comptent : certains utilisateurs souhaitent réduire la maille d'une simulation, d'autres augmenter le nombre de variables, d'autres encore tester des hypothèses plus rapidement pour confirmer ou fermer des pistes de travail... Nous avons aussi pris en compte la nature technologique des codes de calcul utilisés par les chercheurs : langage, architecture, organisation des données... Sans oublier les besoins futurs, comme l'IA, en fort dévelop-

pement, et bien sûr les performances économiques et environnementales. » L'une des réponses tient en trois lettres : GPU, pour Graphics Processing Unit, un type de microprocesseur inventé pour l'affichage graphique (en clair : les jeux vidéo !) et capable de réaliser très rapidement de très nombreux calculs, grâce à une structure hautement parallèle. Gaïa en compte 64. « C'est encore peu, mais ces processeurs demandent une adaptation particulière des codes de calcul, déjà faite pour quelques-uns. La proportion de GPU est à l'évidence

3- Ce fondateur d'Intel énonça dès 1975 que la puissance des microprocesseurs doublerait tous les 2 ans.

4- Le triplement de puissance de Gaïa versus Eole reste une décision stratégique de la R&D d'accueillir Framatome sur ses clusters pendant 3 ans et de répondre à la croissance de ses propres besoins (volume des études, besoin de précision supplémentaire).

LA R&D D'EDF ACCUEILLE LE GRATIN DE L'INFORMATIQUE SCIENTIFIQUE DE DEMAIN

En décembre 2018, le CEA et la R&D d'EDF ont organisé conjointement le colloque « Réinventer l'informatique : Au-delà des architectures Von Neumann », à EDF Lab Paris-Saclay.

Architecture physique des composants, algorithmes neuro-morphiques, calculateurs quantiques, consommation d'énergie des supercalculateurs, portes logiques optiques... de nombreuses pistes ont été discutées.

Seul consensus pour ces experts : **la loi de Moore ne sera bientôt plus d'actualité** car on arrive aux limites physiques du modèle « transistor » (chaleur, vitesse de l'information, voltage...). Le moment est donc venu d'inventer l'informatique du 21^e siècle.

Côté EDF R&D, 2019 verra le lancement d'une **évaluation de l'ordinateur quantique** pour des cas d'usages précis, comme certains algorithmes d'optimisation (recuit simulé quantique), les études probabilistes de sûreté en dynamique, ou encore les applications liées à la cybersécurité.

destinée à augmenter, d'autant plus qu'ils sont devenus particulièrement économes en énergie de par leur utilisation dans les véhicules autonomes » explique Stéphane Dupré la Tour.

IA et HPC : deux faces d'une même pièce

Le lien entre l'architecture matérielle et les logiciels est encore plus fort en matière d'intelligence artificielle. « Les concepts de l'IA sont en réalité assez anciens sur l'échelle du temps de l'informatique. Son renouveau s'appuie en grande partie sur de nouvelles synergies avec le hardware. Là encore, les

GPU (et bientôt les TPU de Google) sont une solution intéressante, notamment pour la simulation basée sur des réseaux de neurones, qui demandent une forte puissance en calculs parallèles et la manipulation de données massives. On observe d'ailleurs une effervescence autour de ce rapprochement hard/soft, par exemple avec les puces neuro-morphiques en développement, dont l'architecture interne cherche à reproduire les synapses du cerveau, ou encore les calculateurs quantiques (probabilistes), dont les données de bases ne sont pas binaires (0 ou 1) mais une distribution statistique de ces deux états ! » précise Stéphane Dupré la Tour. À la R&D d'EDF, l'IA est déjà utilisée pour des simulations et la résolution de problèmes complexes : plan de rechargement des assemblages de combustible nucléaire, équilibre des réseaux électriques...

Accompagner les chercheurs pour optimiser les usages HPC

En attendant l'ordinateur quantique, les chercheurs doivent s'approprier leur nouvelle Formule 1 du calcul scientifique. « Au-delà de la vectorisation des

codes, nous les adaptions et les enrichissons pour y intégrer les approches mathématiques les plus pertinentes pour tirer parti de la puissance de calcul. C'est aussi vrai à l'échelle de la définition des calculs lancés par les chercheurs : ils doivent trouver **la bonne stratégie pour exploiter au mieux le HPC**. Nous les y aidons notamment via le DIL (Data Innovation Lab) dont les experts peuvent proposer des choix d'algorithmes et de méthodes de calcul en fonction des problématiques » précise Stéphane Tanguy, aussitôt rejoint par Jean-Paul Chabard, Directeur Scientifique d'EDF qui conclut : « **La puissance de calcul ne fait pas tout !** Il est indispensable de ne pas oublier le sens physique, dans le choix des modèles, de la granularité, de la précision attendue, en gardant en tête l'exploitation concrète des résultats pour EDF. Et surtout, **soyons ambitieux et imaginatifs** pour mettre notre puissance de calcul au service de la résolution de problèmes concrets et de l'avancée des connaissances ». ●

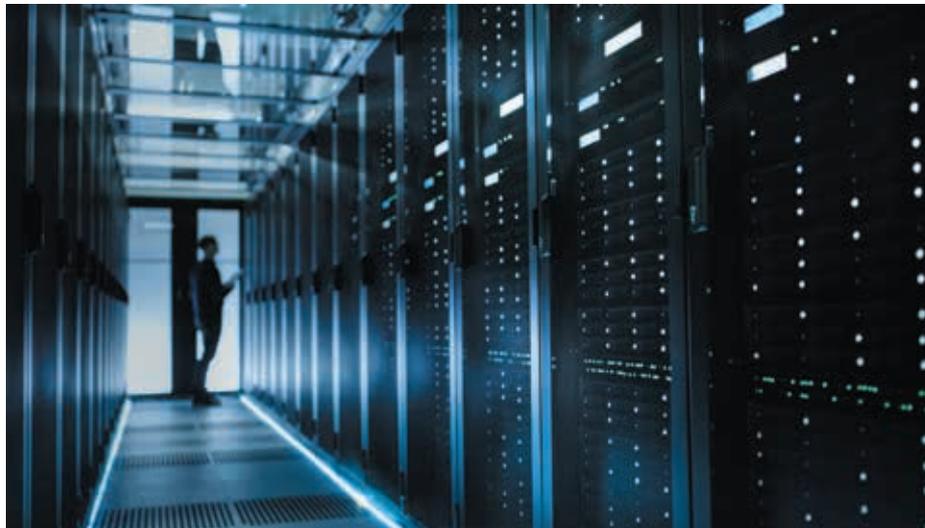


UN NOUVEAU CALCULATEUR POUR LA R&D D'EDF

Avec Gaïa, EDF s'offre un supercalculateur HPC non seulement puissant mais aussi adaptable grâce à son architecture. Pour l'éprouver avant son ouverture opérationnelle, la Direction des Technologies et Systèmes d'Information (DTSI) a choisi de lancer « *Les Grands Challenges du HPC* », l'occasion de faire le point sur les capacités et les possibles utilisations de cet « ordinateur » hors-norme.

Mais pourquoi est-il si puissant ?

Une étude menée par la DTSI, en collaboration étroite avec l'ensemble des départements Métiers et des directions Programmes associées à chaque domaine d'activité, a permis d'établir une roadmap des besoins en moyens de calcul HPC pour la période 2019-2022. Pour le nucléaire, l'intensité croissante de la réglementation, et la complexité des technologies laissent ainsi prévoir un **besoin croissant en calcul scientifique de l'ordre de 30 % par an !** « *L'arrivée de Framatome sur nos infrastructures a donc entraîné le choix d'une machine permettant de répondre aux besoins de l'ensemble des métiers des deux structures. Le tout s'inscrit dans le projet EDF de Mutualisation de l'Informatique Scientifique (MISC), qui prévoit le partage de ressources, à la fois en temps de calcul et en infrastructure de stockage entre les co-investisseurs du cluster* » rappelle Alain Martin, responsable SI Scientifique d'EDF R&D. En 2018, EDF disposait de trois clusters HPC, pour une puissance de calcul de près de 2 200 TFlops. Deux d'entre eux s'arrêtent au premier semestre 2019 et



ne restera qu'EOLE (1 200 TFlops), acquis tant par la R&D d'EDF (800 TFlops) que par la DIPNN¹ (400 TFlops). Avec l'arrivée de Gaïa et ses 3 000 TFlops, la R&D d'EDF dispose aujourd'hui de deux machines HPC pour une puissance totale de **4 PFlops** (4 millions de milliards d'opérations/seconde) soit quelques

530 millions d'heures de calcul par an consommables. Gaïa embarque notamment 64 cartes graphiques Nvidia « GPU Volta », le « top du top » à l'heure actuelle. Les machines sont hébergées dans le Data Center NOE situé à Val-de-Reuil (Eure).



GPU : PUISSANT MAIS AVEC LE BON CODE

Classiquement, les processeurs d'un ordinateur sont des **CPU** (Central Processing Unit) : ils contiennent aujourd'hui plusieurs cœurs (18 par processeur pour Gaïa), et sont capables de réaliser des opérations très **diverses**, de manière **séquentielle comme en parallèle**.

À l'inverse, les **GPU** (Graphics Processing Unit), sont très **spécialisés** : leurs multiples cœurs (43 000 Tensor Core/GPU pour les Nvidia Tesla V100 de Gaïa !) peuvent être utilisés **simultanément pour des calculs ou de l'Intelligence Artificielle**.

En théorie, les GPU de Gaïa peuvent selon certaines configurations être 50 fois plus performants qu'un CPU classique. « *En pratique cela dépend beaucoup de la façon dont les codes de calcul (ndlr : les logiciels) tirent profit de l'architecture GPU. Il faut donc réaliser un travail de « vectorisation » de ces codes, pour espérer des gains de performances variables, compris entre 2 et 50* » explique Cyril Baudry, ingénieur chercheur et architecte HPC à la DTSI.

À ce jour, les deux codes les plus consommateurs du HPC à la R&D ont été adaptés, ou en cours d'adaptation, pour tirer parti de cette puissance : Code_Saturne (mécanique des fluides) et VASP (études de matériaux à l'échelle atomique).

1- Division Ingénierie et Projets Nouveau Nucléaire.

LES GRANDS CHALLENGES : OPEN BAR NUMÉRIQUE !

Comme cela avait déjà été fait lors du lancement d'EOLE, cette période privilégiée permet aux chercheurs de réaliser des calculs dans des configurations dites « extrêmes » difficilement accessibles en production. « Il s'agit de pousser la machine dans ses retranchements en réservant une partie voire l'intégralité de la puissance pour une seule simulation. C'est aussi l'occasion de réaliser des études exploratoires difficiles à caser au sein des usages plus standards, car il est rarissime qu'un cluster HPC soit totalement libre pour des dizaines heures ou jours » décrit Alain Martin. Ce sont ainsi près de 50 millions d'heures-cœur sur les nœuds standards et 75 000 heures sur les nœuds GPU qui ont été offertes. « Cela nous a permis d'évaluer les nouveaux GPU Nvidia Volta pour le calcul, le Deep Learning, l'Intelligence Artificielle et, de tester le portage des codes de calcul » ajoute Cyril Baudry, proche des testeurs au quotidien pour définir et lancer leurs calculs de manière optimale sur ce nouveau cluster.

Les Grands Challenges ont pris fin le 11 mars 2019 avec une demi-journée de bilan qui a rassemblé plus d'une centaine de chercheurs avec Bernard Salha (Directeur) et Stéphane Tanguy (Directeur des Systèmes et Technologies de l'Information) pour l'introduction et Jean-Paul Chabard (Directeur scientifique) et Stéphane Dupré la Tour (Directeur Innovation Avancée et Technologies Numériques) en conclusion.

Gaïa est désormais « bon pour le service en exploitation » ! ●



LISTE DES GRANDS CHALLENGES :

- Simulation massivement parallèle de l'hydrodynamique dans un lit fluidisé de particules de gaz réactif polydispersé à l'échelle industrielle
- Modélisation du primaire du générateur de vapeur en situation diphasique accidentelle avec SALOME_CFD
- Simulation SPH de vagues impactant une centrale nucléaire
- Convergence infrastructures DIL / HPC pour la Data Science : architecture et use cases deep learning.
- DEEPLAN : un réseau de neurones au service de l'optimisation du planning nucléaire
- Apprentissage par renforcement pour le rechargement du cœur d'un réacteur nucléaire
- Étude thermique de la structure des internes de cuve du palier N4
- Étude de l'écoulement dans une portion de guide de grappe de type 1300 Mwe
- Calcul en sortie de boîte à eau GV pour la mesure du débit primaire
- Construction de l'étude de sûreté long terme du projet CIGEO
- Fragilisation non durcissante de l'acier de cuve : premières prévisions atomistiques sur GAIA
- Sensivity analysis for wind turbine structure damage
- GIROSCOP : Générateur Industriel et Opérationnel de Scénarios de Consommation et de Production pour Enedis

ORGANISATION : LE SLURM À L'AIDE

Historiquement le HPC de la R&D d'EDF tourne en permanence à plein régime, soit 80 % de la charge maximale. Et l'arrivée de la puissance de Gaïa ne devrait rien y changer : « *D'une part, les besoins croissent, et d'autre part, on assiste généralement à un effet rebond lorsque l'on augmente la puissance machine : des calculs impossibles et/ou auxquels on aurait renoncé sont lancés* » analyse Cyril Baudry, qui est d'ailleurs en charge du paramétrage des ressources de calculs, une démarche dénommée Quality of Service (QoS). Il est assisté dans sa tâche par SLURM, un logiciel open source d'ordonnement. « *SLURM permet simplement d'appliquer les règles de gouvernance du cluster, qui concernent principalement les conditions et la priorité des calculs : ressources allouées sur un seul type de nœud à la fois, calculs limités à 3 jours, priorité aux calculs courts et à l'antériorité, possibilité d'inscrire des opérations non critiques dans une liste de calculs « bouche trous »... Bref, nous essayons tous les jours de satisfaire tout le monde en nous basant sur des règles connues de tous* » dévoile Cyril Baudry. ●

PUISSANT MAIS ÉCONOME EN ÉNERGIE

Les moyens de calcul HPC, premier centre de consommation électrique du datacenter, participent à la certification ISO 50001 du centre par la mise en place de mécanismes d'économie d'énergie sur les nœuds de calcul (extinction des nœuds non sollicités durant 60 minutes).

LES PRINCIPALES APPLICATIONS DU HPC À LA R&D D'EDF



Simulation mécanique des fluides

(Code_Saturne – EDF) : 50 % du temps de calcul.



Simulation et études sur des matériaux

(Code VASP – Université de Vienne) : 20 % du temps de calcul.



Performance et sûreté des centrales nucléaires :

études sismiques, résistance des matériaux à l'échelle atomique, simulation neutronique, planification et optimisation de la recharge en combustible

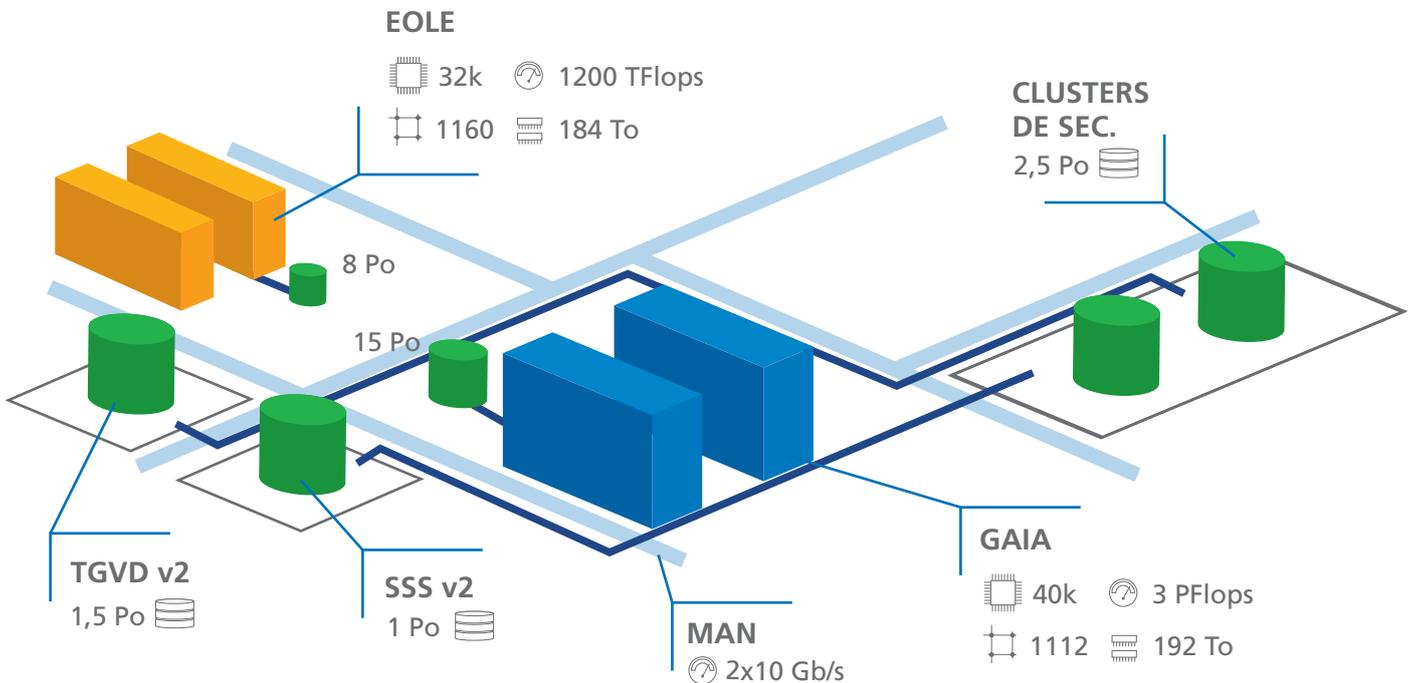


Optimisation de l'équilibre offre/demande

UN PEU DE VOCABULAIRE...

- **Nœud** (ou lame) : un ordinateur, sous la forme d'une carte électronique insérée dans une armoire.
- **Cluster** : grappe de plusieurs nœuds, partageant des données et parfois de la mémoire vive.
- **PétaFlops** : 1 million de milliards d'instructions par seconde. Le milliard d'instructions par seconde (gigaFlops) a été dépassé en 1985, les 1 000 milliards (tétraFlops) en 1997, et le pétaFlop en 2008. Prochaine étape : l'exaFlops, avec un milliard de milliards de calculs par seconde !
- **Cœur** (ou unité de calcul) : une partie du processeur, capable d'exécuter un programme en autonomie.
- **CPU** (Central Processing Unit) : processeur classique, semblable à celui de votre PC.
- **GPU** (Graphics Processing Unit) : processeur graphique, qui renferme plusieurs milliers de cœurs, et donc capable de traiter un grand nombre de calculs simultanément : le calcul parallèle.
- **TPU** (Tensor Processing Unit) : processeur lancé par Google en 2016, et adapté à une bibliothèque logicielle mathématique utilisée pour l'apprentissage automatique (machine learning).

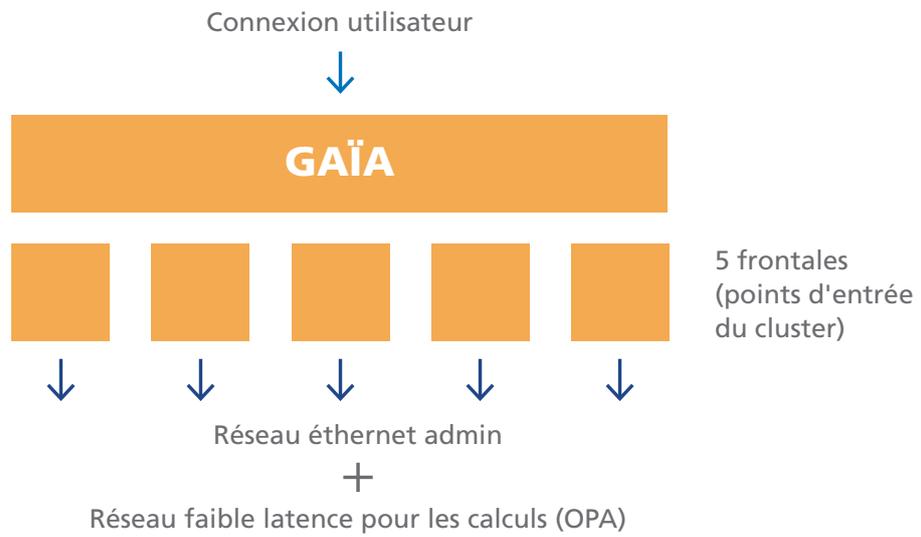
ÉCOSYSTÈME HPC À LA R&D



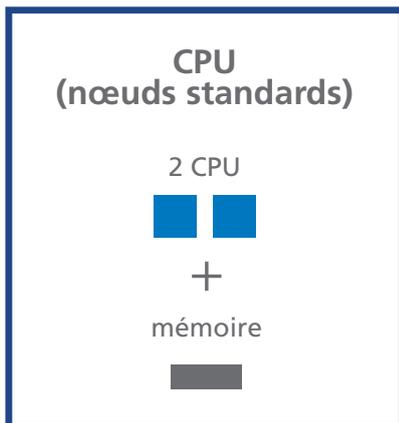
GAIA CARACTÉRISTIQUES TECHNIQUES

Fabricant	ATOS/BULL
Puissance	3 PFlops (97 ^e cluster le plus puissant au monde fin 2018)
Nœuds	<ul style="list-style-type: none"> ■ Standards : 1 112 nœuds (40 032 cœurs) ■ BigMem : 80 nœuds (1 984 cœurs) ■ GPU : 32 nœuds (1 376 000 cœurs Tensors) et 64 cartes GPU Nvidia V100
Mémoire vive totale	244 To (soit environ l'équivalent de 15 000 d'ordinateurs portables)
Stockage	15 Po (15 000 disques dur d'1 To)
Réseau	Intel Omni-Path
Système de gestion de fichier	IBM Spectrum Scale (GPFS)
Job Manager	Slurm
Hébergement	Datacenter NOE
Empreinte au sol	<p>37 armoires 42 U :</p> <ul style="list-style-type: none"> ■ 18 armoires équipées de portes à eau ■ 19 refroidies par air (couloir froid) <p>Soit à peine 40 m² au sol !</p>

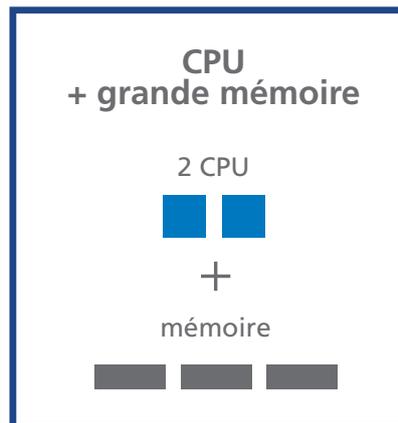
ARCHITECTURE DU CLUSTER GAÏA



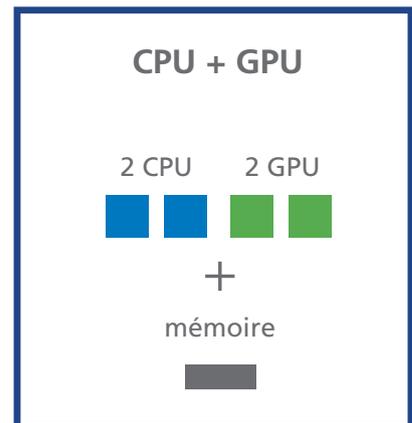
TYPES DE NŒUDS



1 112 nœuds
+ stockage local 2x1 To

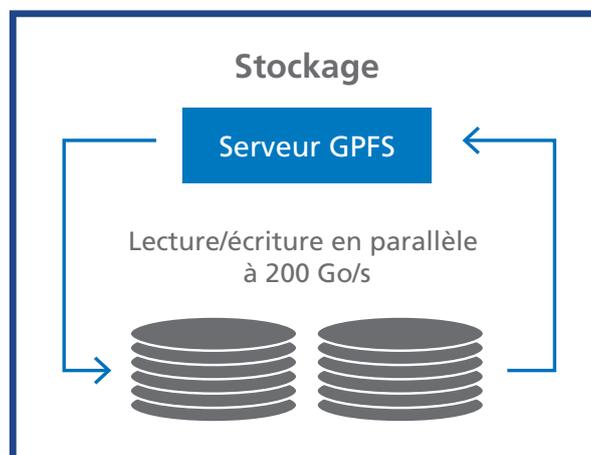


96 nœuds
+ stockage local 2x1 To



32 nœuds
+ stockage local 2x1 To

+



2 000 disques durs de 8 To

HPC ET ÉTUDES DES MATÉRIAUX : 20 ANS D'HISTOIRE COMMUNE

Dès les années 90, les ingénieurs du département Matériaux et Mécanique des Composants de la R&D d'EDF se sont initiés au calcul parallèle sur les ordinateurs de l'époque, déjà partagés avec l'ONERA, le CEA, et la DGA. Leur travail est aujourd'hui indissociable du HPC, qui intervient à toutes les étapes des simulations, en particulier pour prévoir le vieillissement des matériaux soumis aux radiations. Le challenge : prévoir l'évolution sur 500 000 heures de composants de plusieurs mètres-tonnes à partir de mécanismes élémentaires à l'échelle de la picoseconde¹ et de l'angström².



Le nucléaire : un client important et exigeant

Inutile de rappeler l'importance de la sûreté des centrales nucléaires pour l'ASN, ainsi que leur durée de vie pour l'exploitant. La cuve du réacteur, qui n'est ni réparable, ni remplaçable, en est un bon exemple : il est fondamental d'en anticiper le vieillissement. Une étude complexe mêlant des données d'exploitation (dans le cadre du pro-

gramme de surveillance, des échantillons témoins de l'acier de la cuve sont placés dans le cœur et leur vieillissement accéléré analysé périodiquement), des expériences en réacteur expérimental (permettant des études paramétriques sur les conditions d'irradiation telles que température, flux de neutrons, composition chimique des matériaux, avec cependant des limitations dès que l'on irradie avec des neutrons), et des simulations numériques. « Et ce n'est

pas tout ! intervient Christophe Domain, ingénieur de recherche. *Chaque cuve est en réalité unique, notamment parce que la composition chimique présente des légères variations générées lors de la fabrication / l'élaboration de la cuve. L'échelle est d'ailleurs notre enjeu numéro un car nous devons prévoir un phénomène mécanique final tangible, comme le durcissement dû à des amas de taille nanométrique formés sous irradiation, dont l'origine se joue dans les interactions entre atomes voisins. Or un seul centimètre cube de matière contient un nombre d'atomes exprimé à la puissance 22 ! Déjà trop pour entrer dans la mémoire vive du plus gros ordinateur du monde. Nous avons donc recours à la modélisation multi-échelle qui consiste à simuler les différents mécanismes de vieillissement depuis l'échelle atomique et les mécanismes élémentaires aux phénomènes macroscopiques. »* Le HPC reste indispensable, notamment pour les modélisations à l'échelle atomique basées sur la mécanique quantique, particulièrement bien adaptée au calcul parallèle (sur des CPU et des premiers calculs sur des GPU) : une bonne pratique consiste d'ailleurs à répartir les calculs en distribuant quelques atomes par cœur de processeur.

1- Un millionième de millionième de seconde

2- Un dixième de milliardième de mètre

L'étude des matériaux : un bon client du HPC

La durée typique d'un calcul HPC lancé par le département Matériaux et Mécanique des Composants est de l'ordre de la journée avec 1 000 cœurs. Un usage stable malgré la progression de la puissance disponible : « Il y a 15 ans, nous simulions des défauts sur un seul type d'élément à la fois (le cuivre de l'acier de cuve par exemple). Nous pouvons aujourd'hui travailler à la fois de manière plus large, en analysant les synergies entre défauts, et plusieurs éléments simultanément (Cu, Mn, Ni, Si...); et avec une bien meilleure finesse, par exemple en étudiant en détail des défauts à l'échelle atomique. Songez que la simulation de l'impact d'un neutron dure aujourd'hui quelques heures, contre une semaine en 97 ! » se rappelle Christophe Domain. Sans surprise, le département Matériaux et Mécanique des Composants est donc l'un des principaux utilisateurs du HPC de la R&D, avec plusieurs millions

d'heures CPU par an, et devrait le rester car d'autres usages frappent à la porte : « L'IA se développe fortement (voir notre encadré), et nous sommes amenés à réaliser de plus en plus de calculs pour des matériaux fonctionnels : pour les batteries et piles à combustible, et également pour l'étude des défauts dans les semi-conducteurs photovoltaïques, qui conditionnent leurs performances et leur durée de vie » précise Christophe Domain. En effet, la filière solaire cherche sans cesse de nouveaux matériaux, plus performants et/ou moins chers, ou mieux recyclables. La recherche *in silico*, via des simulations numériques, permet ainsi d'envisager (relativement) rapidement des centaines de combinaisons de matériaux et évaluer leurs propriétés photovoltaïques théoriques. Et que fait-on de toutes ces données ? On les garde bien sûr. Le big-data est donc un compagnon naturel du HPC, d'autant plus que l'IA nécessite aussi de grandes quantités de données. ●

MACHINE LEARNING ET HPC, QUELLES APPLICATIONS POUR LES MATÉRIAUX ?

Les calculs d'irradiation concernent des échelles allant du milliardième de mètre au mètre, imposant des calculs intermédiaires (simulation multi-échelle), et utilisant des « physiques » différentes (diffusion, transport, conductivité...). Dans le cadre du vieillissement des matériaux de structure, la microstructure d'échantillons issus des différents aciers de cuves de réacteurs du programme de surveillance a été analysée finement pour quantifier des amas en partie responsables du durcissement de l'acier. Leur formation a également été modélisée avec au final plusieurs milliers de calculs prenant en compte différentes configurations (composition chimique et arrangement local des éléments chimiques). Ces données calculées ont ainsi pu nourrir des **réseaux de neurones pour prédire directement les énergies de migration des atomes**, et donc éviter de nombreux calculs dans le cadre des modélisations multi-échelles et multi-physiques. Par ailleurs, l'IA/machine learning commence à être utilisé pour rechercher de nouveaux matériaux, par exemple des structures métalliques (alliages à haute entropie) ou des matériaux fonctionnels pour les futures batteries et cellules photovoltaïques.



CODE_ASTER S'OFFRE UNE MEILLEURE SYNERGIE AVEC LE HPC POUR SES 30 ANS

Initié en 1989, *Code_Aster* (pour Analyse des Structures Thermo-mécaniques pour des Études et des Recherches) est à la fois un moyen de recherche pour la R&D et un outil d'exploitation quotidien pour l'ingénierie d'EDF. Mis en Open Source, ses performances et sa forte intégration ont déjà séduit plus d'un millier d'utilisateurs réguliers. Conçu à une époque où le HPC n'existait pas encore, il a dû être adapté progressivement pour profiter pleinement des puissances de calcul HPC.

30 ans de progression continue

Une des forces de *Code_Aster* tient dans l'intégration dans un seul et même « solveur » de nombreux phénomènes physiques (thermique, acoustique, hydratation, séchage...), éléments finis et lois de comportement (acier, béton, géomatériaux...) en lien avec la mécanique. Il embarque plus de 200 modélisations, qui répondent à des besoins très variés, à des échelles allant du micromètre au kilomètre. « Jusqu'à présent, les évolutions étaient principalement scientifiques. La progression très rapide des puissances de calcul n'était pas totalement exploitée. Avec le soutien du « programme TI » et de partenariats fructueux, l'équipe *Code_Aster* a donc réalisé un important travail pour adapter une partie du code afin de lui apporter des gains de performance significatifs en HPC. **La nouvelle version sera disponible pour de premières utilisations en 2019, pour les 30 ans du code !** » annonce Dominique Geoffroy, chef de projet PSM (Plateforme de Simulation en Mécanique).

Des verrous technologiques supprimés

Au-delà des progrès en vitesse de calcul, l'implémentation du HPC dans *Code_Aster* ouvre véritablement de nouveaux horizons. Certains calculs sont tout simplement impossibles à réaliser sans HPC, faute de mémoire vive suffisante :

par exemple la modélisation sismique à l'échelle du kilomètre autour d'un barrage ou d'une centrale nucléaire. « **Le HPC permet d'intégrer des échelles très différentes dans un même calcul, alors qu'il fallait auparavant coupler des modèles différents pour chaque échelle, au détriment de la précision finale. Nous pouvons ainsi désormais modéliser la propagation d'une fissure de quelques dizaines de microns sur un tuyau de plusieurs mètres, ou encore mélanger plusieurs modélisations dans un même calcul, par exemple pour les études de CIGEO¹, qui associe des calculs d'excavation et de comportement de roche naturelle** » se réjouit Dominique Geoffroy. Autre nouveauté, la possibilité de réaliser des **études paramétriques**, qui demandent une multitude de calculs identiques, mais avec des données d'entrées différentes car incertaines. Ces simulations permettent notamment de coller au plus près de la réalité, à travers des scénarios variés.

Apprivoiser la machine

« *Sans maîtrise, la puissance n'est rien* » disait la pub. Pour profiter pleinement d'EOLE et Gaïa, les chercheurs doivent préparer encore plus soigneusement leurs calculs en amont : « À chaque physique correspond un ou plusieurs algorithmes ou modèles. Il faut donc choisir le plus adapté. Le HPC y ajoute la nécessité de réfléchir à des tactiques pour bien « découper » ses calculs et les distribuer sur les différentes machines et/ou configurations du calculateur. En clair, **la contrepartie de la puissance consiste à élaborer une véritable stratégie de résolution du problème, à la fois physique et informatique** » prévient Dominique Geoffroy. Et ce n'est pas fini : l'équipe *Code_Aster* planche déjà sur des évolutions dans l'architecture du code, qui s'appuiera encore plus profondément sur le HPC. ●



1- Futur site d'enfouissement de déchets radioactifs pour l'ANDRA.



EDF DataCenter NOÉ : centre de stockage de données, infrastructure éco efficace, situé dans l'Eure, près de Rouen

LE CALCUL HAUTE PERFORMANCE DES ÉCOULEMENTS INDUSTRIELS COMPLEXES AVEC CODE_SATURNE

Code_Saturne est le logiciel EDF de simulation numérique d'écoulements à l'échelle locale (simulation « CFD » pour Computational Fluid Dynamics) utilisé dans de multiples domaines (thermo-hydraulique des circuits des réacteurs nucléaires, hydraulique, simulation d'incendies, écoulements atmosphériques, souterrains...). Outre son périmètre d'utilisation étendu, l'un des atouts majeurs de Code_Saturne est sa capacité à tirer le meilleur parti de la puissance de calcul des supercalculateurs d'EDF afin de réaliser des études de plus en plus précises sur des géométries complexes, tout en visualisant les résultats des calculs sur les nœuds graphiques des supercalcula-

teurs grâce au module paraviz de la plateforme Salomé. À titre d'exemple, lors des Grands Challenges Gaïa, des simulations Code_Saturne ont notamment permis d'accéder aux caractéristiques locales des écoulements hautement turbulents au niveau des guides de grappes de combustible, ou (via le solveur NEPTUNE_CFD dédié aux écoulements multiphasiques, qui s'appuie sur les capacités HPC de Code_Saturne) de simuler le comportement du circuit primaire d'un réacteur nucléaire au niveau d'un générateur de vapeur sur un maillage de plus d'un milliard d'éléments.

3 QUESTIONS À AUDE PELLETIER

Chef de département délégué OSIRIS - Optimisation, Simulation, Risques et Statistiques pour les marchés de l'énergie



Plus de 130 chercheurs œuvrent au sein du département OSIRIS (Optimisation, Simulation, Risques et Statistiques pour les marchés de l'énergie) sur la thématique du management d'énergie - trait d'union entre la production à l'amont et la consommation à l'aval. Leurs travaux portent sur l'optimisation de l'équilibre offre demande et la gestion des risques de marché, et sont fortement liés aux mutations du système énergétique. Ils réalisent les études et développent et adaptent les méthodes et outils qui permettent une gestion optimisée du portefeuille d'EDF. Des travaux basés sur une très forte expertise en mathématiques appliquées, informatique scientifique et sciences de la donnée, gros consommateurs de calculs scientifiques et donc de HPC. Explications autour d'une application essentielle pour le Groupe EDF : l'équilibre offre demande.

Quels sont les enjeux de l'équilibre offre demande ?

Assurer cet équilibre, c'est prévoir au mieux la consommation de l'ensemble des clients, puis gérer et optimiser le portefeuille de production, les achats, et les ventes, en adéquation avec cette prévision. Avec pour objectif de maximiser

les gains pour l'entreprise. La gestion de portefeuille se prépare très longtemps à l'avance et se pilote jusqu'au dernier moment. Trois horizons de temps sont ainsi considérés : la structuration de portefeuille à long terme (vision prospective à plusieurs décennies), la gestion de portefeuille à moyen terme (une ou quelques années), et enfin la gestion à court terme (hebdomadaire, journalier, infra-journalier). Au département OSIRIS, nous travaillons sur ces problématiques au service des optimiseurs du groupe EDF, en particulier la Direction Optimisation Amont/aval & Trading. Pour le court terme, la R&D d'EDF a mis au point pour la Direction Optimisation Amont/aval & Trading un outil industriel appelé Apogène, capable de proposer chaque jour pour le lendemain un programme de production optimal des centrales EDF. Il s'agit d'un problème mathématique très complexe : une équation à 250 000 inconnues, à résoudre en temps opérationnel contraint, une dizaine de minutes ! L'enjeu est de taille : les coûts variables annuels mis en

jeu dans l'optimisation de production sont de l'ordre de 4 milliards d'euros, un gain de 1 point dans les résultats d'optimisation peut ainsi faire gagner 40 millions d'euros.

Côté gestion moyen terme, c'est-à-dire un horizon annuel à pluriannuel, les ingénieurs d'OSIRIS apportent des méthodes et outils d'aide à la décision pour la gestion des vallées et stocks hydrauliques, ou encore pour la gestion du planning d'arrêt des centrales nucléaires. La génération d'un planning nucléaire repose sur des méthodes itératives, qui nécessitent de tester des centaines de milliers de solutions : autant dire que le cluster est un prérequis pour s'attaquer à ce type de problème. Les enjeux financiers sont là aussi significatifs : un planning très bien optimisé peut permettre de réaliser plusieurs dizaines de millions d'euros de gains annuels. Sur l'horizon long-terme, notre outil « Continental Model » est utilisé par les équipes de la Direction de la Stratégie, pour l'exercice stratégique du CHypSE (Corps d'Hypothèses Stratégique Européen)



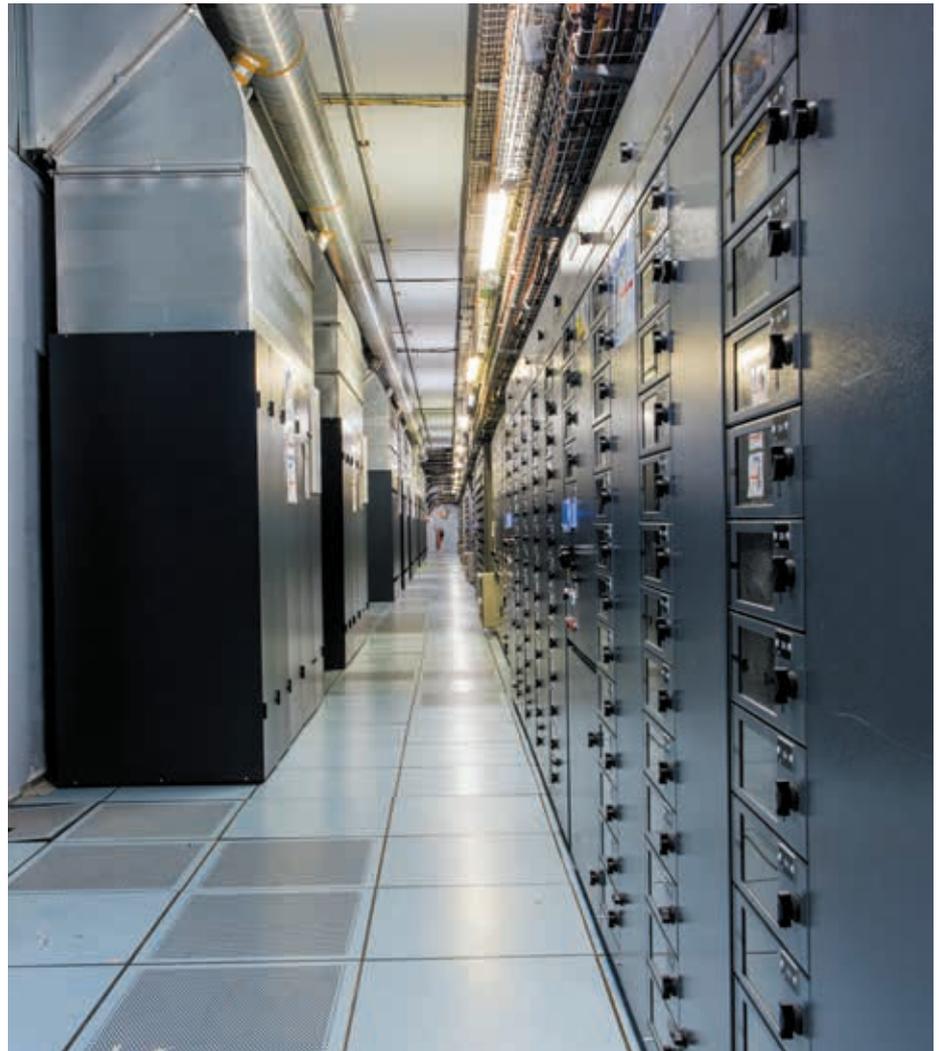
Quel lien entretenez-vous avec le HPC ?

Les clusters HPC sont un outil du quotidien pour les chercheurs d'OSIRIS. La complexité et la taille des problèmes mathématiques auxquels nous nous attaquons rendent souvent son utilisation incontournable pour la mise au point des outils livrés à nos clients, ainsi que pour certaines études que nous menons pour leur compte en interne.

Le calculateur HPC est ainsi absolument nécessaire pour la mise au point et le test d'Apogène et de ses évolutions. En effet, Apogène est un outil industriel dont la fiabilité et la performance doivent être assurées. Pour qualifier chaque nouvelle version, les chercheurs d'OSIRIS effectuent plus de 5 000 « tirs » qui doivent tous être réussis. Avec des machines scientifiques classiques, il faudrait 7 journées complètes pour avoir un résultat qui prend moins de 8 heures sur le cluster. Sans HPC, la qualification de l'outil serait longue et difficile, le panel des tests moins complet, et l'outil de moindre qualité, sujet à problèmes opérationnels. Grâce aux clusters HPC, la R&D d'EDF peut en garantir la qualité mais aussi proposer de nouvelles versions plus souvent et répondre très rapidement aux demandes de maintenance en conditions opérationnelles.

Les technologies HPC permettent aussi des innovations : pour les plannings des arrêts nucléaires, nous avons testé en 2019 de nouvelles approches à base de réseaux de neurones sur Gaïa dans le cadre des Grands Challenges. Ces méthodes devraient nous permettre à terme d'améliorer encore les gains.

Dernier exemple : notre outil GIROSCOP, aussi présenté dans le cadre des Grands Challenges, qui permet de générer des scénarios conjoints de production-consommation, utiles au dimensionnement du réseau de distribution en pre-



nant en compte la production diffuse des énergies renouvelables, la météo, l'évolution du parc, et la consommation : les calculs ont pris respectivement 15 h sur Eole, et 5 h sur Gaïa, contre... 3 000 h avec un calculateur classique, soit tout de même plus de 4 mois ! Vous comprenez donc pourquoi nous sommes attachés au HPC.

Comment voyez-vous évoluer votre usage du HPC ?

OSIRIS s'est développé historiquement autour des activités de management d'énergie du système centralisé. Avec les mutations actuelles du système énergétique, de nouveaux problèmes d'optimisation apparaissent : intégration d'énergies renouvelables décen-

tralisées et faiblement pilotables, évolution des modes de consommation vers plus de local, développement de la gestion des mobilités électriques, évolutions réglementaires associées, etc. Tandis que de nouvelles techniques se développent dans les sciences de la donnée et l'intelligence artificielle : nous les testons pour la gestion des risques financiers sur les marchés, pour la gestion de portefeuille, et pour la prévision de consommation. Ces nouveaux besoins et opportunités plaident pour un usage croissant du HPC, pour adapter les outils actuels et développer ceux de demain. Nos attentes sont donc très fortes vis-à-vis de la performance actuelle et future de ce formidable outil qu'est le HPC. ●



EDF
22-30 avenue de Wagram
75382 Paris cedex 08 - France
Capital de 1 054 568 341,50 euros
552 081 317 R.C.S. Paris

www.edf.fr

Publication EDF R&D
EDF Lab Paris-Saclay
7, boulevard Gaspard Monge
91120 Palaiseau France
Directeur de la publication : Bernard Salha
Directeur de la rédaction : François Molho

Comité de rédaction : Équipe de la communication EDF R&D
Contact : sandrine.dyevre@edf.fr

Crédits photos : EDF - Allard Denis - Caraveo Marc - Eranian Philippe I Getty
Images I Shutterstock - Sdecoret
Conception et réalisation : Okédito