

# Comment la recherche pose un nouveau regard sur le monde.

Avec les serveurs Lenovo ThinkSystem et leurs processeurs AMD EPYC™, l'Institut de Physique du Globe de Paris bénéficie de performances 5 fois plus élevées par rapport à l'infrastructure précédente pour faire progresser les recherches en sciences de la Terre et de l'Univers.

Solutions d'Infrastructure Lenovo  
pour les Organisations Centrées sur les Données

1

## **Le contexte**

Etablissement d'enseignement supérieur et de recherche de renommée mondiale, l'Institut de physique du globe de Paris (IPGP) regroupe plus de 500 personnes – post-doctorants, doctorants, chercheurs, ingénieurs, techniciens et agents administratifs – qui partagent la même passion pour les sciences de la Terre, de l'Univers et de l'Environnement.

L'IPGP étudie la Terre et les planètes, de leur noyau à leur atmosphère. Cette mission est assurée par 17 équipes dont les sujets de recherche vont de la mesure des variations du champ magnétique terrestre à la surveillance des volcans actifs sur le territoire français et l'étude de l'activité sismique.

2

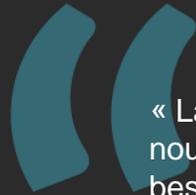
## L'enjeu

La recherche à l'IPGP évolue constamment. Ses équipes utilisent des outils d'imagerie, de modélisation et d'apprentissage machine de plus en plus sophistiqués pour étudier des systèmes physiques très complexes. Cette activité repose sur – et produit – d'énormes volumes de données dont l'analyse est de la responsabilité du Service de calcul parallèle et d'analyse de données (S-CAPAD) de l'IPGP.

Avec la convergence du calcul intensif (HPC) et de la science des données, l'IPGP fait face à un besoin croissant de ressources de traitement et de stockage. Alexandre Fournier, responsable scientifique de S-CAPAD, commente : « Les nouvelles méthodologies comme les systèmes de détection acoustique distribuée (Distributed Acoustic Sensing) et l'interférométrie radar satellitaire capturent ou créent d'énormes volumes de données. Les nouvelles possibilités offertes par l'apprentissage machine permettent aux chercheurs d'exploiter ces données au maximum. Or l'infrastructure HPC précédente du S-CAPAD s'est révélée insuffisante tant en termes de capacité que de performance pour répondre à nos besoins en calcul intensif et analyse de données massives. »

Pour continuer à progresser dans des domaines de recherche essentiels, l'IPGP s'est associé au Laboratoire Astroparticules et Cosmologie (APC) appartenant à Université de Paris, dans la création du projet HPC appelé DANTE (pour multi Data ANalysis and compuTing Environment for science).

L'objectif principal de DANTE est de mutualiser, au sein de l'IPGP, les ressources matérielles et logicielles du S-CAPAD et du centre François Arago (FACe) de l'APC pour satisfaire les besoins croissants en moyens HPC pour la recherche et la formation en sciences de la Terre, de l'Univers et de l'Environnement, et autres utilisations. Pour atteindre ce but, l'IPGP devait se doter d'une plateforme de calcul et de stockage optimisée pour les jeux de données massifs.



« La recherche devenant de plus en plus complexe et gourmande en données, nous devons faire évoluer notre environnement HPC pour répondre aux besoins des chercheurs. »

**Alexandre Fournier**

responsable scientifique de la plateforme DANTE, IPGP

## **Pourquoi Lenovo et AMD ? Plus de performance et plus d'efficacité grâce aux processeurs de pointe AMD.**

A la suite du lancement d'un appel d'offres et de l'évaluation de plusieurs propositions, l'IPGP a retenu Lenovo comme fournisseur technologique majeur pour DANTE. Au cœur de la nouvelle solution HPC se trouvent les serveurs Lenovo ThinkSystem SR645, équipés de deux processeurs AMD EPYC™ 7002 Series de deuxième génération. Possédant jusqu'à 128 cœurs physiques par serveur, le SR645 fournit des performances de haut niveau associées à une conception qui en fait un système au rapport coût/efficacité optimal et économe en énergie – un avantage déterminant pour l'IPGP.

Geneviève Moguilny, responsable opérationnelle du S-CAPAD de l'IPGP, explique : « Pour la plateforme DANTE, nous nous étions fixé des limites strictes tant pour le budget que pour la consommation énergétique. Il est très important que Lenovo ait pu nous fournir un nombre de cœurs de calcul aussi élevé tout en respectant ces contraintes. »



« La proposition de Lenovo équipée des processeurs AMD EPYC™ est celle qui répondait le mieux à nos exigences en termes de prix, qualité et nombre de cœurs. »

**Geneviève Moguilny**

responsable opérationnelle du S-CAPAD, IPGP

## Doper la puissance de calcul et la capacité de stockage.

Pour sa mise en œuvre, l'IPGP a travaillé étroitement avec Lenovo Services et Neoteckno, un partenaire de Lenovo spécialisé dans le déploiement et la gestion des architectures HPC. Collaborant étroitement, l'équipe a installé un cluster composé des serveurs Lenovo ThinkSystem SR645 et SR670, incluant plus de 7 000 cœurs de calcul et 12 cartes graphiques de dernière génération pour les traitements massivement parallèles (GPGPU). DANTE affiche une puissance de calcul maximale (théorique) de plus de 298 téraflops pour ses seuls processeurs, complétée par un surplus significatif de performance grâce aux GPGPU. La plateforme fonctionne sous Red Hat Enterprise Linux et utilise SLURM Workload Manager pour l'ordonnancement des tâches.

L'IPGP recourt aussi aux technologies Lenovo pour stocker et gérer de gros volumes de données, qu'elles soient issues de l'observation ou générées par les applications tournant sur la plateforme DANTE : un ThinkSystem DM7100H Hybrid Flash Array pour le stockage à long terme des données d'observation et des résultats de calculs, et un ThinkSystem DE6000H Hybrid Storage Array, directement connecté au cluster HPC, pour que les travaux de simulation accèdent rapidement aux jeux massifs de données.

Outre la mise en œuvre de la plateforme DANTE, Lenovo Services et Neoteckno devraient accompagner l'IPGP dans la formation des utilisateurs. La formation aux bonnes pratiques est prévue pour durer tout au long des 7 années de vie du projet.

Alexandre Fournier remarque : « L'équipe en charge de DANTE est assez réduite, il est donc important que les étudiants et les chercheurs puissent utiliser la plateforme en toute autonomie. Lenovo et Neoteckno ont un rôle déterminant à jouer dans l'élaboration du plan de formation. Les nouveaux utilisateurs devront se familiariser rapidement avec la plateforme pour en tirer le meilleur parti. »



« Non seulement le cluster HPC de Lenovo est extrêmement performant mais ses capacités de stockage sont impressionnantes. Il s'agit d'une véritable originalité de notre plateforme : les utilisateurs peuvent faire tourner leurs applications à proximité immédiate de gros volumes de données, ce qui rend la recherche plus efficace. »

**Alexandre Fournier**

responsable scientifique de la plateforme DANTE, IPGP

3

## Les résultats

Inauguré en octobre 2021, DANTE est utilisé à ce jour par 12 équipes de recherche au sein de l'IPGP – un nombre qui va augmenter. Une étudiante s'en sert pour modéliser, en profitant de l'apprentissage machine, la relation entre la composition d'un magma et sa viscosité : cette propriété contrôle le degré d'explosivité volcanique, et donc la dangerosité. L'importante capacité de stockage offerte par les GPGPU lui permet d'exécuter son modèle en une seule fois, ce qui nécessitait auparavant un traitement par lots.

De plus, les serveurs Lenovo, équipés des processeurs AMD EPYC™, fournissent une puissance de calcul 5 fois supérieure à l'infrastructure HPC précédente. Grâce à ce gain de performance, les chercheurs vont plus loin dans l'exploration de leurs données et sont capables d'émettre de nouvelles hypothèses, le tout plus vite et plus facilement. DANTE les aide également à tester les performances et l'évolutivité de leurs algorithmes, les résultats devenant un avantage compétitif pour accéder aux ressources nationales de calcul intensif.

« Pour pouvoir utiliser les moyens nationaux de calcul haute performance, les chercheurs doivent en effet démontrer à la fois l'importance scientifique de leur projet et la capacité de leur application à évoluer et à faire un usage efficace de ces ressources », souligne Alexandre Fournier. « Une des principales utilisations de DANTE consiste à affiner et perfectionner les simulations numériques pour qu'elles puissent être exécutées sur les équipements nationaux. Il est aussi beaucoup plus facile de travailler avec de très gros volumes de données. Comme il n'est toujours pas pratique de transférer de telles quantités de données à travers les réseaux, DANTE a l'avantage de les stocker juste à côté des ressources HPC. »



Geneviève Moguilny ajoute : « Le gain de vitesse des serveurs Lenovo est une vraie satisfaction pour les utilisateurs. Pour les applications qui font appel au multithreading, disposer de deux processeurs 64 cœurs avec multithreading automatique au sein d'un même serveur permet de multiplier la vitesse de calcul par 100 sans pour autant modifier le code source. Accéder à de nombreux cœurs sur un seul serveur physique élimine toute la complexité liée à la parallélisation du code et à la distribution de la mémoire à travers plusieurs serveurs physiques. »

Avec DANTE, l'IPGP continue à mettre ses ressources HPC à la disposition de la communauté des chercheurs et a vocation à s'ouvrir à d'autres disciplines comme les mathématiques, la médecine ou la linguistique.

Alexandre Fournier conclut : « En nous permettant de faire un bond significatif en matière de ressources HPC et d'analyse de données massives, le cluster DANTE de Lenovo nous aide à continuer à repousser les limites de la connaissance scientifique. »

- ✓ **Performance 5 fois plus élevée par rapport à l'infrastructure précédente**
- ✓ **Plus de 7 000 cœurs de calcul**
- ✓ **Puissance maximale de 298 téraflops pour ses seuls processeurs**
- ✓ **Consommation électrique maximale de 54kW pendant benchmark HPLinpack**



« Notre partenariat avec Lenovo ouvre de nouvelles perspectives pour la recherche. Il contribue à la réalisation de projets innovants qui approfondissent notre compréhension de notre planète et de l'Univers. »

**Geneviève Moguilny**

responsable opérationnelle du S-CAPAD, IPGP

## Que ferez-vous avec les solutions Lenovo HPC?

Les organisations centrées sur les données accélèrent leurs recherches vitales grâce à des solutions d'infrastructure intelligentes de Lenovo, alimentées par AMD.

[Etudiez les Solutions HPC de Lenovo](#)



Lenovo et le logo Lenovo sont des marques ou des marques déposées de Lenovo.

AMD, le logo AMD, EPYC et leurs combinaisons sont des marques commerciales d'Advanced Micro Devices, Inc.

Les autres noms de sociétés, de produits et de services peuvent être des marques commerciales ou des marques de service de tiers.

© Lenovo 2022. Tous les droits sont réservés.