



Le Big Data stocke l'espace : le CNRS projette d'héberger le plus grand catalogue d'objets célestes

Paris, le 8 juin 2015

Depuis 2012, un partenariat lie le CNRS et la société Dell dans le but de développer des solutions novatrices autour du traitement massif de données pour les besoins de la recherche. Cette année, le partenariat s'est focalisé sur le projet LSST (*Large Synoptic Survey Telescope*) qui promet des avancées majeures dans les domaines de l'astrophysique et de la cosmologie.

Implanté au Chili, le *Large Synoptic Survey Telescope* (LSST) est un télescope de 8,4 mètres de diamètre équipé d'une camera comptant 3,2 milliard de pixels. A partir de 2022, le télescope photographiera le ciel de manière systématique durant dix ans : chaque partie du ciel sera revisitée toutes les quatre nuits, créant ainsi un film en trois dimensions de la portion d'Univers visible. L'un des objectifs de ces observations est de mesurer avec une précision inégalée les quantités physiques associées à l'énergie noire qui tend à accélérer l'expansion de l'Univers.

Le traitement des données issues du LSST constitue un vrai défi puisque, chaque nuit, environ 2000 images seront prises, soit plus de quinze Téraoctets de données produits, transportés, stockés et exploités. Ces données subiront des traitements spécifiques visant à l'identification, à la classification et au catalogage automatique d'objets célestes afin de produire un jeu de données structuré, exploitable par les chercheurs. Ces traitements seront effectués dans deux centres informatiques, le Centre de calcul de l'Institut national de physique nucléaire et de physique des particules du CNRS (CC-IN2P3) en France et le *National Center for Supercomputing Applications* (NCSA) aux Etats-Unis, et reposeront sur un système de bases de données massivement distribuées.

À terme, ce sont 37 milliards d'objets astronomiques (galaxies, étoiles, astéroïdes...) qui seront catalogués, soit un volume total de 500 Pétaoctets de données¹ (dont 50 Pétaoctets indexés dans la base de données), ce qui en fait l'un des plus grands projets

¹ Soit 500 millions de Gigaoctets, l'équivalent de 100 millions de DVDs.

Big Data référencé. Avec le concours de Dell, la technologie associée à ces datacenters est à la pointe de l'innovation et a nécessité un travail conséquent se situant à la frontière entre la R&D technique et la recherche en informatique.

Pour mener à bien ce projet, le CNRS devait disposer d'un prototype de ce que sera, à terme, le catalogue le plus complet d'objets célestes jamais réalisé. Grâce au partenariat signé entre les deux entités, c'est désormais chose faite puisque Dell vient de fournir une plateforme expérimentale équipée de 50 serveurs (soit 500 Téraoctets de stockage) destinée à héberger un système de base de données nommé « Qserv », développé spécifiquement pour les besoins du LSST par le *SLAC National Accelerator Laboratory* aux États-Unis. Ce matériel permettra aux experts du CC-IN2P3 (CNRS) et du SLAC de vérifier la stabilité du système, de préparer la future plateforme d'indexation de données spécifique au LSST et de tester sa capacité à traiter l'ensemble des données.

Contacts presse Dell			
Odile Polge	Dell France	Odile_Polge@Dell.com	06 34 53 35 48
King Houndekpinkou	Cohn&Wolfe	king.houndekpinkou@cohnwolfe.com	01 49 70 43 32
Contacts presse CNRS			
Contact presse	CNRS	presse@cnsr.fr	01 44 96 51 51