

Lyon, le 2 mai 2005

L'expérience D0 entame une nouvelle reconstruction de données via la grille

Plusieurs centres de calcul¹, dont celui de l'IN2P3/CNRS, situés dans six pays, ont mis en commun leur puissance informatique afin de « reconstruire » les innombrables données expérimentales issues de l'expérience D0 dédiée à l'étude des interactions protons/antiprotons à de hautes énergies. Grâce aux technologies de grille, ces ordinateurs vont traiter en six mois plus de 250 téraoctets² (TO) de données, soit assez pour réaliser une pile de CD's aussi haute que la Tour Eiffel. Le CC-IN2P3 a été l'un des premiers centres à démarrer cette reconstruction des données, il y a un mois.

3000 processeurs gigahertz Pentium III : c'est l'équivalent de la puissance informatique récemment atteinte pour retraiter les données issues des interactions protons/antiprotons à de hautes énergies étudiées dans l'expérience D0³, installée auprès du collisionneur Tevatron (FNAL⁴, États-Unis) et à laquelle collaborent près de 650 scientifiques originaires d'une vingtaine de pays.

Les données collectées par le détecteur D0 sont enregistrées puis doivent être « reconstruites » (ou retraitées) sous une forme utilisable par les physiciens. Une nouvelle reconstruction devient nécessaire lorsque les scientifiques impliqués dans l'expérience ont accompli des avancées significatives, par exemple dans l'étalonnage du détecteur. Ces nouveaux éléments sont alors intégrés au logiciel utilisé pour un traitement plus efficace et une étude plus rigoureuse des données de physique.

Lors d'un retraitement en 2003, le CC-IN2P3 avait déjà brillamment reconstruit 36% des données brutes envoyées par le FNAL vers les autres sites. Une réelle performance puisque les serveurs du centre représentaient seulement 8% des ressources globales du projet.

Avec les avancées technologiques et logicielles réalisées depuis, un nouveau retraitement s'avère nécessaire pour l'ensemble des données collectées jusqu'à ce jour. La mise en commun des ressources des différents sites est effectuée grâce à la grille de calcul Samgrid, qui permet de traiter en six mois plus de 250 téraoctets (TO) de données, ce que le FNAL mettrait trois ans à reconstruire tout seul.

Le CC-IN2P3 et Westgrid au Canada ont été les premiers centres à démarrer ce retraitement, il y a un mois. A ce jour, près de 23 téraoctets de données (issues de près de 120 millions de collisions de particules) ont déjà été transférés du FNAL vers les serveurs HPSS⁵ du CC-IN2P3. Au terme de l'opération, 20 % des données envoyées par le FNAL vers les différents sites seront traitées au CC-IN2P3, à une moyenne de 2 millions de collisions par jour. Cette

¹ CC-IN2P3 à Lyon, Westgrid au Canada, GridKa en Allemagne, FZU en République Tchèque, GridPP au Royaume-Uni et les Universités du Wisconsin, de l'Oklahoma et du Texas aux États-Unis. Des centres situés au Brésil, en Inde, en Corée et en Chine pourraient venir grossir les rangs des sites participant à cette opération.

² 1 téraoctet = 1000 gigaoctets (10¹² octets)

³ www-d0.fnal.gov

⁴ Fermi National Accelerator Laboratory

⁵ HPSS (High-Performance Storage System) est un système qui permet le stockage hiérarchique de gros volumes de données, avec un accès aux fichiers par nom.

efficacité est rendue possible grâce à un effort particulier d'adaptation de la grille aux ressources du centre.

Patrice Lebrun (IPNL) et Tibor Kurca (détaché au CC-IN2P3) sont responsables de ce programme de reconstruction au sein de D0 France. Laurent Duflot (LAL), également impliqué dans cette opération, a notamment en charge le transfert des données au CC-IN2P3.

Contact chercheur

Patrice LEBRUN (IPN de Lyon)

Tél. : 04 72 44 84 43

lebrun@in2p3.fr

Contact communication

Gaëlle SHIFRIN (CC-IN2P3)

Tél. : 04 78 93 08 80

gshifrin@in2p3.fr