

Vue du tunnel du LHC en août 2007.

Le calcul scientifique des expériences LHC

Une grille de production mondiale

Fairouz Malek (fairouz.malek@in2p3.fr)

Responsable scientifique du projet LCG-France, membre de l'expérience ATLAS du LHC
LPSG, CNRS-IN2P3 et Université Joseph Fourier, 53 Av. des Martyrs, 38 000 Grenoble

Le LHC, c'est avant tout l'accélérateur de particules le plus puissant au monde ; il a nécessité des travaux d'infrastructure titanesques. C'est également une aventure humaine et une entreprise intellectuelle toutes deux uniques dans l'histoire des sciences.

L'analyse de l'énorme quantité de données enregistrées par ses expériences exploitera une infrastructure mondiale inédite, la « Grille de Calcul », comparable en puissance de calcul à l'infrastructure de « Google », mais autrement plus performante pour le traitement, le transfert et le stockage des données.

Cette infrastructure de calcul a donné naissance à une grille de production mondiale, qui a attiré d'autres disciplines scientifiques.

Le LHC, acronyme anglais pour Grand Collisionneur de Hadrons (Large Hadron Collider), est le plus puissant accélérateur de particules subatomiques au monde. Les particules accélérées à une vitesse proche de celle de la lumière, y circulent dans un anneau de 27 km de circonférence, enterré à 100 m sous terre et situé à cheval sur la frontière franco-suisse près de Genève.

Mis en œuvre par le laboratoire européen du CERN (Organisation européenne pour la recherche nucléaire), le LHC et les équipements expérimentaux associés ont été conçus et construits par une collaboration internationale. Il a démarré en septembre 2008 mais, le 19 du même mois, de graves incidents techniques ont nécessité un arrêt qui a duré plus d'un an. Il fonctionne de nouveau depuis le 20 novembre 2009 et a livré plus de 100 millions de collisions à des énergies atteignant 7 TeV (tétra⁽⁴⁾ eV) dans le centre de masse, la moitié de l'énergie nominale, attendue dans deux ans.

Quatre gigantesques expériences, ALICE (A Large Ion Collider Experiment), ATLAS (A Toroidal LHC ApparatuS), CMS (Compact Muon Solenoid) et LHCb (où le « b » signale que l'expérience est dédiée à la physique du quark *b*), observeront les collisions entre les particules accélérées pour tenter de percer quelques-uns des mystères fondamentaux de l'origine de

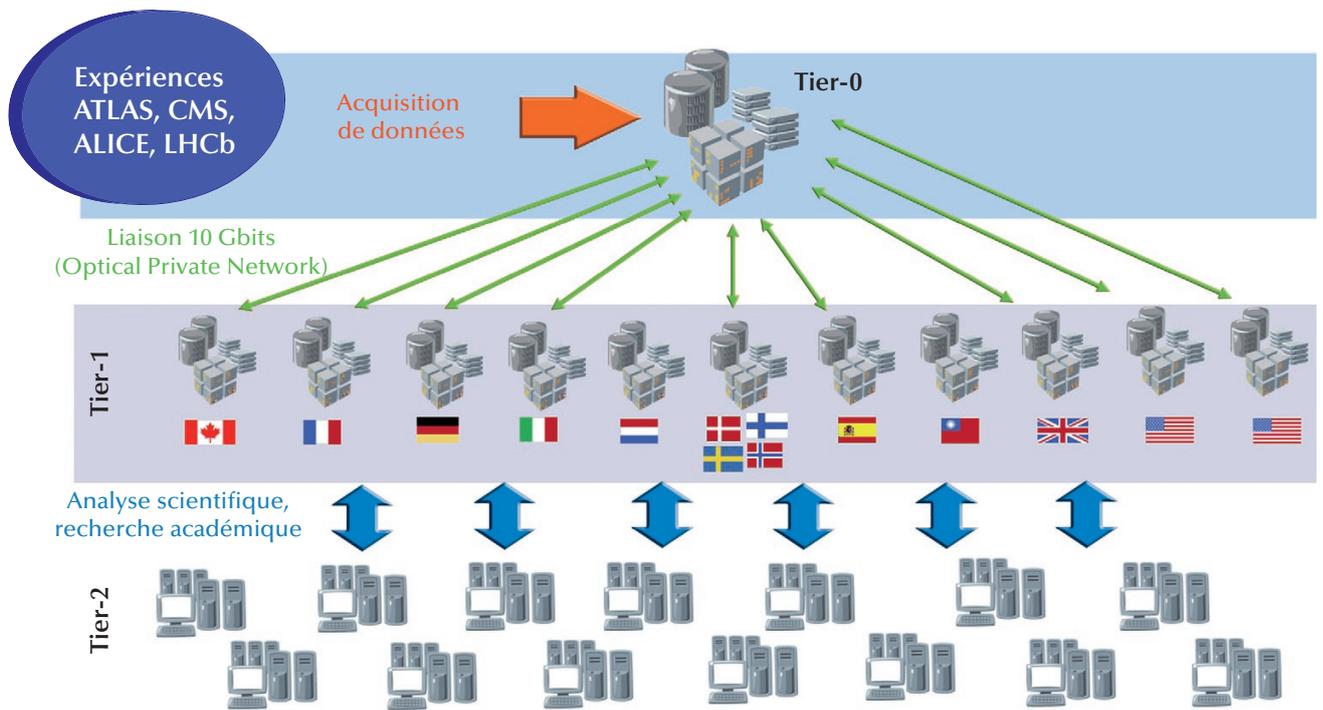
l'Univers. Plus de 8000 scientifiques du monde entier sont associés à cette aventure humaine hors du commun.

L'article de B. Mansoulié dans *Reflets* n°14, pp. 5-9, décrit en détail l'accélérateur, les expériences, la démarche et les objectifs scientifiques attendus. Cet article-ci vous fera découvrir les moyens et les innovations technologiques mis en œuvre pour le calcul scientifique, qui permettront aux physiciens des quatre expériences d'analyser les données du LHC.

Le traitement des données

Lorsque le LHC fonctionnera en régime de croisière, 40 millions de particules accélérées entreront en collision chaque seconde, et des milliers de particules secondaires seront générées et traverseront les détecteurs des expériences. *In fine*, le passage d'une particule dans un détecteur se traduit par des impulsions électroniques à partir desquelles les physiciens peuvent déterminer sa trajectoire et son identité. Les millions d'impulsions créées par chaque collision sont alors numérisées par des circuits électroniques intégrés aux détecteurs avant d'être transmises par fibres optiques aux ordinateurs d'acquisition de données. Les données numériques y sont regroupées sous formes de fichiers binaires de





© CERN

1. Architecture du calcul des expériences LHC.

Définitions

- (a) **Giga** = 10^9 ; **Téra** = 10^{12} ; **Péta** = 10^{15} .
- (b) **Octet** : une unité de mesure de quantité de données. Un octet est lui-même composé de 8 bits, soit 8 chiffres binaires. 1 Moctet = 1 million d'octets.
- (c) **Ferme de calcul** : une grappe d'ordinateurs fournissant de la puissance de calcul pour traiter l'information.
- (d) **Tier** : mot anglais voulant dire niveau ou rang. L'image la plus proche de cette définition est l'oignon que l'on épluche rang par rang.
- (e) **MoU** : Memorandum of Understanding = protocole d'accord de collaboration, dont la référence est : CERN-C-RRB-2005-01.
- (f) **1 CPU** = 1 processeur (1 cœur) des ordinateurs les plus performants de nos jours.
- (g) **IN2P3** : Institut National de Physique Nucléaire et de Physique des Particules, un des instituts nationaux du CNRS.
- (h) **IRFU** : Institut de Recherche sur les lois Fondamentales de l'Univers de la DSM (Direction des Sciences de la Matière) du CEA.

>>>

« données brutes ». Chaque collision constitue, au niveau de chacune des quatre expériences, un événement dont la taille en données brutes est de l'ordre du million d'octets^(b) (1 Mo). Afin de réduire la quantité de données, un tri intelligent est réalisé en un temps ultra-court (moins d'une seconde) pour ne garder que les événements potentiellement intéressants. Ce tri réduit le taux d'acquisition dans une première étape (quelques microsecondes) à 100 000 événements par seconde, grâce au circuit électronique de déclenchement, puis, dans une deuxième étape à environ 100 ou 200 événements par seconde, cette fois-ci en traitant les informations à l'aide de ressources (fermes de calcul^(c)) et d'algorithmes dédiés. Les données brutes, après tri, sont finalement enregistrées à un taux équivalent à 1,5 CD par seconde pour les quatre expériences, sachant que la capacité d'enregistrement d'un CD est de 700 Mo.

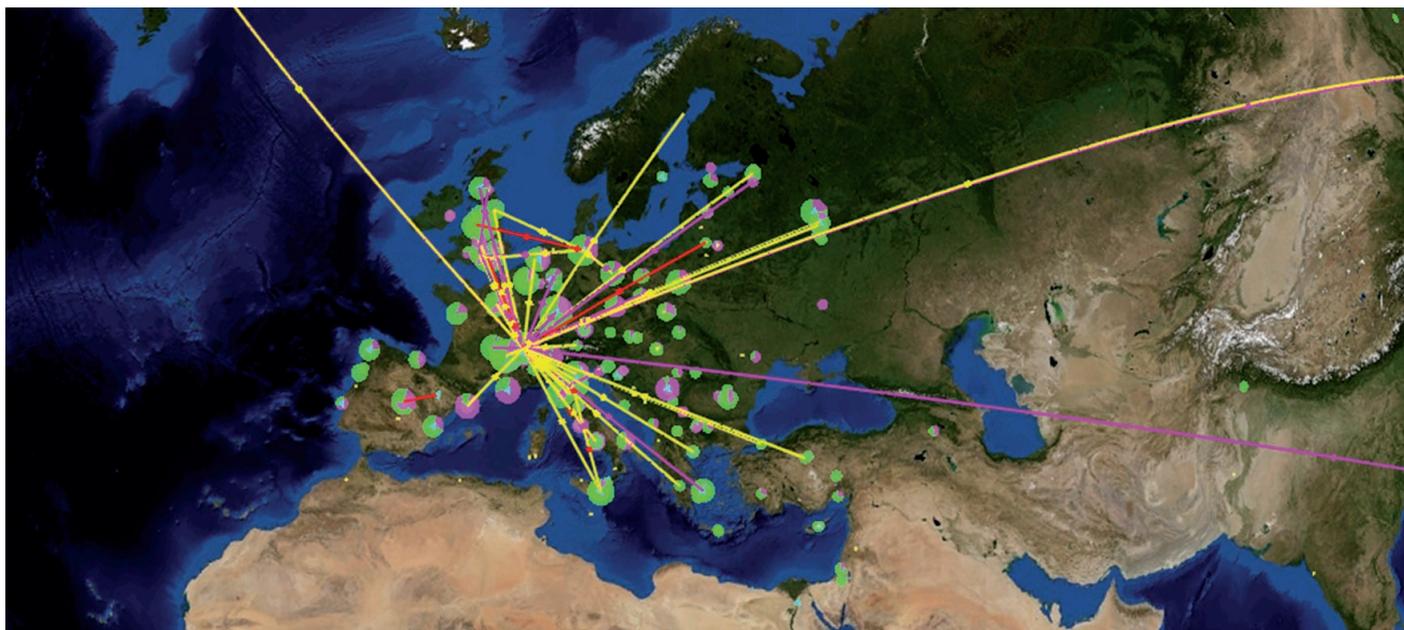
Le processus de traitement des données brutes peut alors commencer. Il s'effectue en deux étapes. La première transforme les millions d'informations enregistrées en des informations dites de « physique », contenant les caractéristiques de chaque trajectoire, l'impulsion et l'identité de chacune des particules détectées. Dans la deuxième étape, ces informations sont analysées et les résultats sont regroupés de façon à pouvoir identifier et interpréter les phénomènes physiques. Ces traitements réduisent la taille de

l'information d'un facteur 2 à 10. Le processus est en général répété plusieurs fois par an sur l'ensemble des données acquises depuis le démarrage, lorsqu'elles sont produites dans des conditions similaires. Ce «(re)processing» est effectué pour tenir compte du développement continu des algorithmes de traitement et d'étalonnage du détecteur.

La grille de calcul : l'état de l'art

Dans les conditions nominales de fonctionnement du LHC, les expériences enregistreront chaque année 10^{10} collisions ; ce dernier chiffre est comparable au nombre d'étoiles dans la Voie Lactée. Traiter les données ainsi accumulées représente un véritable défi informatique, autant pour les flux (de l'ordre du gigaoctet^(a) par seconde) que pour leur volume (plusieurs dizaines de pétaoctets^(a) chaque année). À tout instant, plusieurs milliers de chercheurs du monde entier solliciteront des ressources de calcul et de stockage pour analyser ces données. Pour faire face à ce défi, la solution du calcul distribué, ou « grille de calcul », s'est imposée. Elle a été mise en œuvre par la collaboration W-LCG (World wide LHC Computing Grid) [1].

Les supercalculateurs – dont le Blue Gene/L d'IBM – offrent des capacités de calcul de l'ordre de milliers de milliards d'opérations à la seconde, en utilisant des architectures de type « massivement paral-



2. Image instantanée de la grille EGEE. Les cercles représentent les machines qui effectuent les tâches à l'instant de la prise d'écran (en vert : tâches en exécution ; en rose : tâches en attente d'exécution ; la taille du cercle représente la quantité de tâches en exécution par rapport au total mondial). Les lignes représentent les tâches transférées entre sites (en jaune : tâches faites et réussies ; en rouge : tâches avortées ; en rose : tâches en attente d'exécution).

lèles ». Ces ordinateurs utilisent les meilleures technologies disponibles au moment de leur conception. Chaque processeur permet le traitement d'une portion du problème à traiter, et l'échange des données entre les processeurs se fait à un très haut débit. L'inversion d'une matrice est le calcul type pour lequel ce genre d'ordinateur est efficace.

Les grilles informatiques, quant à elles, mettent en commun des ressources de calcul et de stockage de tous les objets reliés à la grille. Elles associent plusieurs générations de processeurs usuels diffusés à grande échelle dans le commerce comme, par exemple, les Intel Xeon et AMD opteron. Sur la grille, chaque processeur traite un seul événement. Les événements étant indépendants les uns des autres, il y a donc peu de données à transférer entre les processeurs. Le principal intérêt de ce système, fondé sur la mise en commun des ressources, réside dans ses capacités quasi infinies et dans sa capacité d'évolution en continu, à la différence des machines massivement parallèles. Il suffit, en effet, d'ajouter de nouvelles entités pour augmenter d'autant les capacités de calcul de la grille. Celles-ci peuvent égaler les supercalculateurs en puissance de calcul, bien qu'elles soient limitées par les capacités de transfert entre processeurs ; mais leur principal atout est qu'elles peuvent traiter des quantités énormes (mesurées en pétaoctets) de données réparties sur la grille.

Si nous évaluons les capacités des supercalculateurs en Tflops/sec, ce sont, pour les grilles informatiques, les pétaoctets de données stockées, échangées et distribuées qui sont la bonne mesure. Les grilles sont centrées sur l'analyse d'énormes quantités de données en s'appuyant sur de très nombreux ordinateurs, mondialement distribués. Ceci permet, dans une certaine mesure, de partager le coût proportionnellement aux ressources mises à disposition par chacun des pays contributeurs.

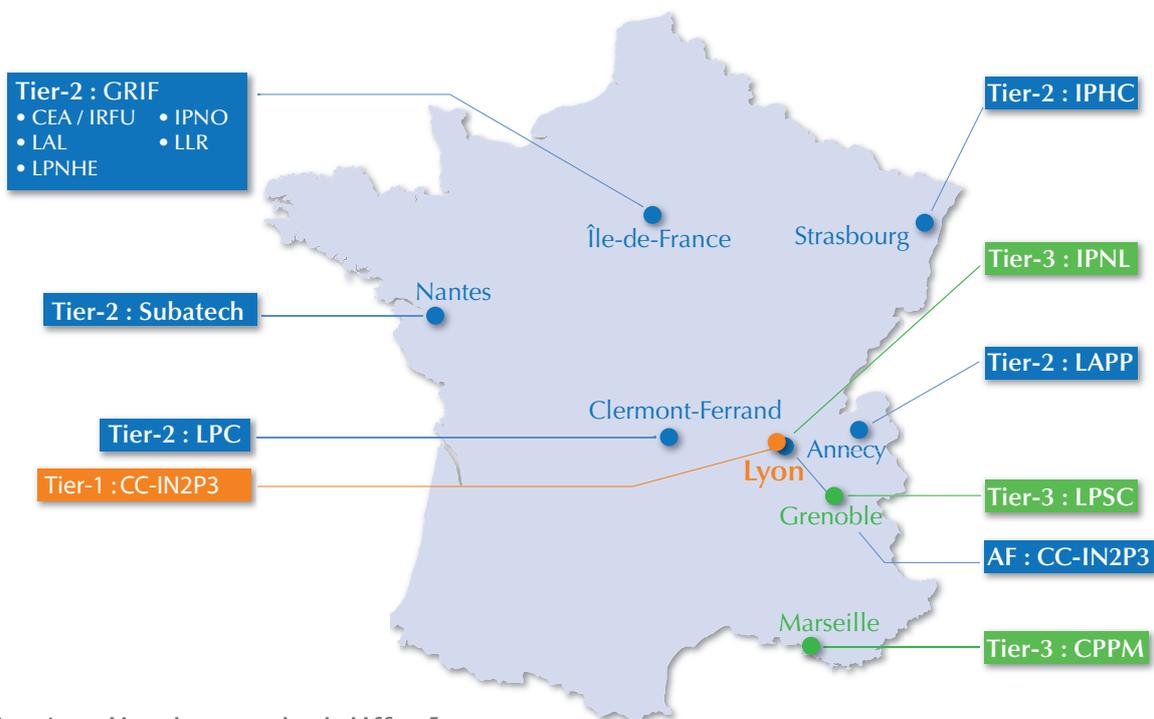
Le principe de la grille W-LCG s'appuie sur des centres de calcul répartis dans le monde entier et qui sont organisés en réseau. Elle s'articule autour de deux grilles principales, technologiquement différentes mais interopérables, l'une en Europe, EGEE [1] (Enabling Grids for E-sciencE), et l'autre aux États-Unis, OSG [1] (Open Science Grid), auxquelles sont associées des grilles régionales ou nationales. Lorsqu'une requête est formulée par un utilisateur à partir de l'un des nombreux points d'entrée du système, la grille de calcul établit l'identité de l'utilisateur, vérifie ses autorisations, recherche les sites disponibles et susceptibles de fournir les ressources requises, en termes de stockage, de capacité de calcul, de disponibilité des logiciels d'analyse... pour finalement satisfaire la requête de l'utilisateur, sans que ce dernier ait à se soucier du détail du processus.

La grille W-LCG est hiérarchisée en quatre niveaux, ou rangs, appelés « Tier »⁽⁴⁾. L'appartenance à un rang donné fixe la nature et la qualité des services qu'un centre de calcul doit fournir. La quantité de ressources à fournir est à la discrétion des agences de financement, selon un protocole d'accord rediscuté chaque année (MoU)⁽⁶⁾.

Les fonctions exigées pour chacun des rangs (fig. 1) sont les suivantes :

- Tier 0 : il est unique, c'est le centre de calcul du CERN. Il est responsable de la collecte et du stockage des données brutes enregistrées par les expériences et de leur distribution vers les onze centres de rang inférieur, les Tier 1.
- Tier 1 : ce sont de grands centres nationaux ; ils sont chargés de stocker une copie des données brutes, de les traiter pour en extraire des données de physique et de distribuer ces dernières vers les centres de rang inférieur, les Tier 2.
- Tier 2 : au nombre d'environ 150, ce sont les centres des instituts ; ils sont en charge de l'analyse finale des données par les utilisateurs et de la production des simulations.
- Tier 3 : ce sont également des centres dans les instituts ; ils fournissent des ressources complémentaires aux centres Tier 2 pour un usage local, sans engagement vis-à-vis de W-LCG.

>>>



3. Répartition géographique des centres de calcul LCG en France.



Des liaisons haut débit, de 10 Gocets/s ou plus, spécialement conçues pour le LHC (LHC Optical Private Network), connectent le centre Tier 0 aux centres Tier 1. Elles assurent le transit des données venant des expériences et des données de physique, traitées à un taux de 800 Mocets/s. Les centres Tier 2 sont connectés aux Tier 1 par des liaisons de l'ordre du Gocet/s.

Pour traiter l'ensemble des données enregistrées par les expériences du LHC, les besoins en capacité de calcul sont de l'ordre de 100 000 CPU⁽⁶⁾. Pour stocker la totalité de ces données il faudra une capacité de l'ordre de 15 pétaoctets par an, soit l'équivalent de 15 millions de CD qui, empilés, atteindraient une hauteur de 20 km.

Actuellement, la grille W-LCG compte près de 200 sites répartis sur trois continents, en Europe, Amérique et Asie. Avant le démarrage du LHC, elle était déjà opérationnelle, avec plus de 50 000 tâches traitées quotidiennement, pour la production de données simulées et le traitement de données de tests. À partir du mois de mars 2010, W-LCG a vu passer 650 000 tâches par jour ; le million sera bientôt atteint au mois de juin 2010, ce qui tend vers les 100 000 CPU par jour attendus. La figure 2 est une image instantanée du site de contrôle en ligne de la grille EGEE. On peut, entre autres, observer les échanges de données entre sites et la capacité instantanée de calcul fournie par chacun des sites.

La contribution de la France

La France héberge au centre de calcul de l'IN2P3 (CC-IN2P3), situé à Villeurbanne dans la région lyonnaise, l'un des onze centres Tier 1 de W-LCG. Elle dispose également de plusieurs sites Tier 2 et Tier 3 (fig. 3). La gestion de ces ressources, le fonctionnement, le financement et les relations avec W-LCG sont assurés par le projet LCG-France, créé en 2004 [2].

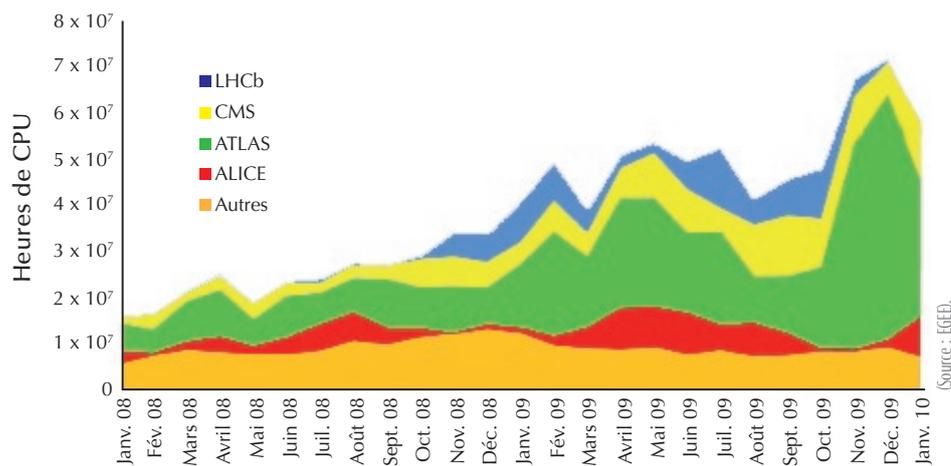
Le CC-IN2P3, financé conjointement par l'IN2P3⁽⁶⁾ du CNRS et l'IRFU^(h) du CEA, offre ses services aux quatre expériences du LHC, à hauteur de 15% de l'effort mondial des Tier 1. Le CC-IN2P3 héberge également un Tier 2 et un Tier 3, regroupés sous l'appellation de « centre d'analyse » (AF). La contribution relative de la France à l'effort mondial des Tier 2 est voisine de 10%. Il est important de noter que les ressources françaises ne sont pas réservées exclusivement aux chercheurs français, mais qu'elles sont accessibles à l'ensemble des chercheurs participant au projet LHC selon l'accord de collaboration (MoU)⁽⁶⁾. Réciproquement, les chercheurs français ont accès à l'ensemble des ressources mondiales.

Vers quel futur ?

La grille de calcul exploitée par la communauté de la physique des particules, historiquement pionnière dans cette application, a vu avec le démarrage du LHC et les études de faisabilité du successeur du LHC, le nombre de ses utilisateurs croître

Pour en savoir plus

- 1 • Les sites de W-LCG, EGEE et OSG se consultent, respectivement, aux adresses suivantes : <http://lcg.web.cern.ch/lcg/>, www.eu-egee.org/, www.opensciencegrid.org/
- 2 • Le site Web du projet LCG-France : <http://lcg.in2p3.fr/>
- 3 • EGI : <http://web.eu-egi.eu/>
- 4 • Institut des Grilles du CNRS : www.idgrilles.fr/
Un GIS est un groupement d'intérêt scientifique ; on peut citer le grand accélérateur GANIL, dont le statut est similaire.
- 5 • Livre blanc sur les grilles de production : oai:hal.in2p3.fr:in2p3-00408379_v1.
- 6 • Article de presse CNRS du 31 janvier 2007 : « Malaria : plus de 4,3 millions de médicaments testés grâce aux grilles de calcul ». www.in2p3.fr/presse/communiques/archives/2007/01_egee_wisdom.htm



4. Évolution mensuelle, en heures de CPU, de l'utilisation de la grille de calcul EGEE par les quatre expériences du LHC, de janvier 2008 à janvier 2010. En orange, utilisation par des expériences extérieures au LHC. (Source : EGEE).

continuellement. De plus, le pouvoir attractif d'une grille de calcul bien rodée est indéniable. Les chercheurs d'autres disciplines connexes, telles que la physique nucléaire et celle des astroparticules, ne manqueront pas de vouloir utiliser ce fantastique outil. Ainsi, dès 2012, près de 85% des ressources de calcul utilisées par les disciplines regroupées au sein de l'IN2P3 le seront au travers de la grille de calcul. Celle-ci sera progressivement utilisée dans tous les domaines de la science et trouvera naturellement des retombées bien au-delà du monde de la recherche.

L'avènement de la grille de calcul comme outil de référence pour de très grandes infrastructures de recherche comme le LHC, nécessite une pérennisation tant à l'échelle mondiale ou européenne que nationale. L'effort entrepris jusque-là pour mettre sur pied une grille de calcul en France doit être vigoureusement poursuivi, avec une attention particulière pour les ressources dédiées au LHC. De cette façon, nos chercheurs pourraient relever le défi de l'analyse de ces données dans un contexte de très forte compétition internationale.

Le projet EGI (European Grid Initiative) [3] est en train de prendre le relais du projet EGEE financé par l'Europe. Il repose sur des initiatives nationales (NGI), dont le « GIS » France Grilles est le représentant français et l'Institut des Grilles du CNRS le mandataire [4]. Le besoin total pour la France est estimé à environ 25 000 cœurs et 75 pétaoctets de disque d'ici 2012 [5].

Les grilles de calcul sont devenues en quelques années un outil indispensable à la communauté nationale dans plusieurs domaines scientifiques de premier plan, tels que la physique subatomique, les sciences du vivant ou les sciences de la planète. Elles apparaissent très clairement comme la ressource informatique complémentaire des grands supercalculateurs utilisés par des disciplines comme la génomique, la météorologie, la dynamique des fluides, la chromodynamique quantique sur réseau, les simulations d'accidents, etc. La figure 4 illustre la croissance des cas d'utilisation de la grille de calcul, de janvier 2008 à janvier 2010, toutes disciplines confondues. Environ 30% de cette utilisation concerne des projets autres que ceux du LHC.

Citons un cas exemplaire d'utilisation de la grille de calcul en sciences du vivant : l'expérience Wisdom (*World-wide in silico docking on malaria*) [6]. La stratégie de Wisdom repose sur un criblage virtuel de molécules aux performances thérapeutiques encourageantes : elle permet aux chercheurs de calculer la probabilité qu'une molécule active se fixe sur une protéine cible, altérant ainsi son activité biologique, et, dans le cas de la malaria, la prolifération du parasite. Wisdom s'est alors associée à plusieurs grilles de calcul internationales : la grille régionale Auvergrid en Auvergne, les grilles EELA en Amérique du Sud, EuChinagrid en Chine, EuMedgrid autour de la Méditerranée et ASGC TWGrid ont ainsi mis à disposition leurs ressources de calcul, complétant celles

de la grille EGEE. Au cours des 10 semaines d'expériences, le projet Wisdom a permis d'analyser près de 80 000 médicaments potentiels par heure ; par conséquent, 4,3 millions de molécules actives (médicaments potentiels) ont été testées, et plus de 140 millions de liaisons entre ces molécules et les protéines cibles pour combattre la malaria ont été calculées. Ce programme a mobilisé simultanément plus de 5000 ordinateurs dans 27 pays, produisant plus de 2000 gigaoctets de données, ce qui représente l'équivalent de plus de 400 années de calcul sur un seul ordinateur.

En France, plusieurs laboratoires du CNRS ont apporté leur contribution de façon significative : le Centre de calcul de l'IN2P3 a géré les ressources mises à disposition pour cette initiative et le LPC-Clermont a supervisé l'utilisation des ressources pour les différents calculs scientifiques. La contribution de l'ensemble des laboratoires français sur la grille EGEE dans le cadre de Wisdom représente environ 15% des 400 années de calculs. Grâce à la grille de calcul, Wisdom a ouvert la voie à de nouvelles perspectives thérapeutiques pour le traitement du paludisme, mais également pour combattre d'autres maladies tropicales. Ce travail a débouché sur le dépôt de brevets pour des molécules actives contre le parasite de la malaria. Le même type de calcul sur d'autres cibles biologiques, pour combattre la grippe aviaire et le diabète, a été effectué avec comme résultat le dépôt de brevets pour des nouvelles molécules prometteuses. ■