

**Simon Penel** (simon.penel@univ-lyon1.fr), **Christophe Blanchet** (christophe.blanchet@ibcp.fr), **Pascal Calvat** (pcalvat@cc.in2p3.fr), **Yonny Cardenas** (cardenas@cc.in2p3.fr), **Clément Gauthey** (clement.gauthey@ibcp.fr), **Jean-Yves Nief** (nief@cc.in2p3.fr), **Bruno Spataro** (bruno.spataro@univ-lyon1.fr)

### **Overview:**

Large scale phylogenomics and comparative genomics require complex computational methods (parsimony, maximum likelihood, bayesian methods, MCMC, *etc.*) associated with massively distributed calculation. In this respect, grid computing plays a crucial role. Here we present how we processed exhaustive similarity searches on several millions of sequences with BLAST, using two different grids (TIDRA and GRISBI).

### **Enjeux scientifiques, besoins de la grille :**

L'étude de l'évolution du monde vivant au niveau moléculaire (l'évolution et la dynamique des génomes, la construction de l'arbre du vivant, l'histoire des espèces, les transferts horizontaux de gènes, *etc.*) est un domaine de recherche qui nécessite d'associer l'algorithmie, le développement de méthodes de calcul et, de plus en plus, le calcul intensif.

Par exemple, les études de génomique comparative (comparaison systématique des génomes de différents organismes) nécessitent de rechercher des similarités de séquence dans un grand nombre de génomes. Typiquement, la base de familles de gènes HOGENOM [1], qui fournit les arbres phylogénétiques de tous les gènes de plusieurs centaines de génomes, a nécessité de comparer deux à deux plus de 3 millions de séquences.

En outre, la phylogénie, qui vise à reconstruire l'histoire d'un gène chez un ensemble d'espèces, implique des calculs dont la durée augmente avec le nombre d'espèces et avec la complexité des méthodes utilisées (parsimonie, maximum de vraisemblance, méthodes bayésiennes, chaînes de Markov). Dans le cas de HOGENOM, il s'agissait de calculer la phylogénie de plusieurs centaines de milliers de familles de gènes.

Dans un contexte où le nombre de génomes séquencés explose, l'utilisation des grilles de calcul et/ou du *cloud computing* est une condition nécessaire à la maintenance de la base HOGENOM.

### **Développements, déploiement sur la grille :**

Nous avons utilisé deux grilles de calcul, la grille TIDRA (<http://www.tidra.org/>) et la grille GRISBI (<http://www.grisbio.fr/>).

TIDRA est destinée aux scientifiques d'organismes publics ou privés de la région Rhône-Alpes, toutes disciplines confondues. Elle inclut une grille de calcul régionale, basée sur l'intergiciel gLite, constituée des 5 noeuds suivants : le LAPP (Annecy), le LPSC (Grenoble), l'IPNL (Lyon), l'IBCP (Lyon) et le CC-IN2P3 (Lyon). La grille met potentiellement à disposition des chercheurs plus de 7000 cœurs (CPUs). Le CC-IN2P3 propose dans le cadre de cette grille une expertise du stockage massif de données avec iRODS (<https://www.irods.org/>). Nous avons mis en place une application pilote permettant l'utilisation massivement distribuée du logiciel de

recherche de similarités BLAST [2]. Les calculs de comparaison de séquences en très grand nombre ont permis de tester les performances de cette grille et du système iRODS et finalement de définir une stratégie de production. Ce travail a débouché sur le développement de DTM (J. Cardenas, <http://cc.in2p3.fr/docenligne/947>), un gestionnaire de tâches distribuées. Nous avons utilisé JJS (P. Calvat, <http://cc.in2p3.fr/docenligne/269>) un outil de soumission de jobs sur la grille qui comprend des commandes « grille » simplifiées (soumission de jobs, statut des jobs , gestion des proxies, *etc.*).

GRISBI est une grille de calcul nationale dédiée à la bioinformatique. GRISBI, basée sur l'intergiciel gLite, est constituée des 7 plateformes bioinformatiques suivantes : IBCP (Lyon), LBBE (Lyon) , ABiMS (Roscoff), GenOuest (Rennes), MIGALE (Jouy-en-Josas), CBiB ( Bordeaux), GenoTool (Toulouse), et de la collaboration avec 3 mésocentres pluri-disciplinaires à Bordeaux, Lille et Strasbourg. Elle propose l'accès à 856 coeurs, sur lesquels sont installés et maintenus tous les logiciels de bioinformatique et de biostatistiques standards ainsi que les bases de données biologiques internationales. L'utilisation du système de fichier XtremFS (<http://www.xtremfs.org/>) est possible sur certains noeuds de la grille et en cours d'installation sur les autres. En outre le projet GRISBI a adapté un ensemble de commandes simplifiées pour les opérations les plus communes sur grille (soumission de jobs, statuts, gestion des proxies, *etc.*). Ici aussi les calculs de comparaison de séquences en très grand nombre ont permis de tester les performances de la grille et du système XtremFS.

### **Outils, difficultés rencontrées :**

Dans les deux cas de figure, les problématiques rencontrées sont identiques:

1-l'optimisation du temps passé en calcul et la gestion des tâches

Lorsqu'il existe un grand nombre de tâches à effectuer et que ces tâches ont une durée courte et/ou variable il peut être intéressant de créer des jobs "pilote" ou "agent" qui exécutent plusieurs tâches chacun : cela 1) maximise le temps de calcul d'un job, 2) minimise le temps passé à une action commune à toute les tâches (par exemple le téléchargement d'un fichier utilisé par l'ensemble des tâches) et 3) simplifie la gestion des tâches .

Lors de notre première production sur la grille TIDRA (2009-2010) nous avons choisi une approche qui maximise le temps de calcul effectué par un job et qui maximise le nombre de machines compatibles avec le job à exécuter. Concernant le premier point, il s'agissait d'éviter qu'un job s'arrête avant d'avoir atteint son temps d'exécution maximum. Concernant le deuxième point, la limitation venait des machines qui présentaient la mémoire la plus faible : 2Go. Nous avons donc défini une stratégie reposant sur des tâches unitaires demandant au maximum 2 Go de mémoire, chaque job de la production devant effectuer le plus grand nombre de tâches possible. La répartition optimale des données pour la deuxième production (2011) - 5 millions de séquences à traiter- consistait en environ 170 000 fichiers de 30 séquences. Cette production a représenté au total plusieurs milliers de jobs paramétriques.

Lors de notre première production sur la grille GRISBI (2011) nous avons testé une approche similaire, inspirée de notre expérience précédente, dans laquelle les 170 000 tâches ont été réparties sur quelques milliers de jobs paramétriques, ainsi qu'une approche plus classique dans laquelle chaque job exécutait une seule tâche.

## 2-la gestion des données et des résultats.

Une production type consiste dans le traitement de 5 millions de séquences. La répartition optimale de ces données consiste en environ 170 000 fichiers de 30 séquences ce qui implique d'accéder à ces 170 000 fichiers en entrée puis de stocker les 170 000 fichiers de résultats.

Si chaque job exécute une tâche unique, le fichier d'entrée peut être fourni via la *SandBox* (dans la limite de la taille maximum de la *SandBox*, typiquement moins de 10 Mo). La sortie du logiciel BLAST peut aussi être récupérée via la *SandBox*, à condition que celle-ci ne soit pas trop grosse. Il faut donc trouver un compromis entre une taille de fichier de sortie minimum et un temps de calcul maximum, mais il est possible que la limitation sur la taille du fichier de sortie limite directement le temps passé à calculer et ainsi pénalise l'efficacité du job.

Dans le cas d'une approche via une liste de tâches, chaque job exécutant plusieurs fichiers, il devient plus simple d'utiliser une alternative à la *SandBox* et de déposer tous les fichiers sur la grille. Cependant il n'existe pas de commande « grille » permettant de gérer un grand nombre de fichiers par lot, et il est nécessaire d'utiliser des outils dédiés à cela.

Sur la grille TIDRA nous avons utilisé l'outil iRODS pour mettre en place les données sur la grille et pour leur traitement par les jobs. iRODS permet, en utilisant un mot de passe, la gestion et l'accès aux données stockées sur la grille via des commandes en ligne (CLI) linux, via un navigateur ou encore via une API C. Le client iRODS peut être installé facilement et permet un accès de n'importe quelle machine. En outre iRODS propose différents services (duplication des données, micro-services, métadonnées, *etc.*). Il s'agissait de la première intégration en production de l'outil iRODS avec une grille de calcul.

Sur la grille GRISBI nous avons utilisé l'outil XtremFS. Ce système de fichiers virtualisé (« cloud file system ») permet de créer et d'avoir accès à un répertoire partagé en réseau par tous les jobs de grille. XtremFS permet ainsi de traiter et de répliquer facilement les données de grilles de manière totalement transparente.

## 3-la gestion des proxies.

L'expiration des proxies au cours d'une production a posé des problèmes qui ont été résolus par l'utilisation de proxies de longue durée et le renouvellement automatique des proxies avec l'outil myProxy.

## Résultats scientifiques :

Nous maintenons une base de *hits* BLAST chez l'ensemble des séquences de protéine disponibles (séquences Uniprot, CDS traduits des génomes de Ensembl et d'autres consortiums). La version actuelle contient environ 12 millions de séquences. La mise à jour de la base a demandé les calculs BLAST de 5 millions de séquences contre trois banques de respectivement 5 millions (TIDRA et partiellement GRISBI), 7 millions (cluster local) et 150 000 séquences (GRISBI). Les résultats consistent en plusieurs centaines de milliers de fichiers totalisant au total plus d'un téraoctets (To) de données et décrivant des milliards de hits blasts.

### Temps de calculs :

Sur TIDRA, la première production (22 248 jobs) a consisté en 937 588 heures *HS06* soit 115 751 heures (13 ans) sur 1 processeur Intel Xeon. La deuxième production (15 797 jobs, répartis sur une dizaine de soumissions) a consisté en 513 812 heures *HS06* soit 63 433 heures (7 ans) sur 1 processeur Intel Xeon. Nous avons atteint des pics de 650 jobs en parallèle soit un temps effectif idéal de moins d'une semaine. Mais en tenant compte des ajustements et du temps passé entre les soumissions le temps effectif était de 2 mois.

Sur GRISBI, la production (47 713 job) a consisté en 71 362 heures *HS06* soit 8,810 heures (1 an) sur 1 processeur Intel Xeon. Nous avons atteint des pics de 830 jobs en parallèle, soit un temps effectif théorique de 10 heures, en réalité de plusieurs jours.

### Perspectives et conclusion :

Un outil de gestion automatique des tâches pour la grille GRISBI est en cours de développement au CCIN2P3. La technologie iRODS semble bien supporter un accès intensif aux données. En outre iRODS permet l'utilisation de métadonnées et de micro-services que pourrions exploiter dans l'avenir.

La technologie XtreamFS semble très prometteuse. Nous avons observé un certain nombre d'échecs lors d'un accès intensif aux données, mais ces problèmes ont été résolus en augmentant le délai d'attente du montage.

Même si la stratégie choisie est identique, l'utilisation des deux grilles est différente. Choisir quelle grille utiliser dépend du contexte. GRISBI offre aux bioinformaticiens un panel de logiciels et de bases de données parfaitement adapté à leur thématique particulière. Elle est adaptée à des analyses bioinformatiques classiques.

TIDRA étant généraliste ne propose pas à priori d'outils liés à une thématique scientifique (il est donc nécessaire d'installer les logiciels et données nécessaires au début de chaque job), mais propose une prise en charge de la gestion des job adaptée au traitement d'un grand nombre de tâches. Elle est adaptée à des analyses à très grande échelle et/ou peu classiques.

### Références :

- [1] Penel S, Arigon AM, Dufayard JF, Sertier AS, Daubin V, Duret L, Gouy M and Perrière G (2009) "Databases of homologous gene families for comparative genomics" *BMC Bioinformatics*, **10** (Suppl 6):S3
- [2] Altschul, S.F. & Gish, W. (1996) "Local alignment statistics." *Meth. Enzymol.* **266**:460-480