

Charte utilisateur des ressources de calcul et de stockage du LBBE/PRABI

1 – Introduction

Le **PRABI-AMSB** (plateforme UCBL, FR BioEnviS), correspond à la composante « Analyses et Modélisation des Systèmes Biologiques » du PRABI (Pôle Rhône-Alpes de Bioinformatique).

Le Laboratoire de Biométrie et de Biologie Évolutive (LBBE) et le PRABI-AMSB disposent de ressources de calcul (cluster de 1200 cœurs de calcul) ainsi que de capacités de stockage (~250 To hautes performances, ~300 To de stockage) adaptées notamment au traitement des données NGS.

L'installation, l'administration et la maintenance des ressources informatiques sont sous la responsabilité du pôle informatique du LBBE. Le PRABI-AMSB prend quant à lui en charge l'accompagnement de projets bioinformatiques (e.g. RNA-seq, analyse de données métagénomiques, construction de banques de familles de gènes) et la formation des utilisateurs à l'utilisation de la ligne de commande Linux et des outils bioinformatiques. L'infrastructure de calcul et de stockage est cofinancée depuis sa mise en place par le LBBE et le PRABI. Elle est hébergée pour une partie dans la salle serveur de l'étage PRABI (bâtiment Mendel) et pour l'autre dans une salle dédiée au sous-sol du bâtiment Omega. Une partie du matériel informatique et de sa maintenance est financée par la FR BioEnviS (Fédération de Recherche Bio-Environnement et Santé). Suite à plusieurs demandes remontées au conseil des utilisateurs du PRABI-AMSB, le LBBE et ce dernier ont décidé de faire bénéficier (en mode best effort) à l'ensemble du personnel membres de la FR BioEnviS (permanents et non-permanents) d'un accès aux ressources de calcul et de stockage et à l'expérience acquise par le personnel dans ce domaine.

L'objet de cette charte est d'informer les utilisateurs de la FR BioEnviS quant aux bonnes pratiques relative à l'usages des ressources informatiques mutualisées LBBE/PRABI.

L'utilisation des ressources informatiques susmentionnées entraînera une validation tacite de cette charte.

Pour toutes informations complémentaires, contacter le PRABI-AMSB par mel à l'adresse suivante : contact@prabi.fr.

2 - Périmètre fonctionnel

Les ressources de calcul et de stockage mises à disposition sont spécifiquement adaptées à des projets conséquents d'analyse de données (génomique, transcriptomique, protéomique) et de manière plus générale à tous projets de

bioinformatique que ce soit en terme de capacité de calculs (1200 cœurs) de RAM (~9 To cumulés) et de stockage (~ 550 To cumulés).

Ces ressources informatiques sont tournées vers un public d'utilisateurs expérimentés ayant un minimum de bases en informatique et plus particulièrement en lignes de commandes Linux (voir prérequis). Les utilisateurs débutants pourront bénéficier des conseils du PRABI-AMSB, du pôle informatique du LBBE et des autres utilisateurs sur des aspects plus techniques concernant par exemple l'installation et la configuration des outils bioinformatiques, l'écriture de scripts et de programme ou la parallélisation de calculs. Les utilisateurs sont ainsi conviés à suivre les réunions d'information organisées régulièrement par le pôle informatique au sujet des ressources de calcul et de stockage. Par ailleurs, le service **d'atelier du Jeudi** du PRABI-AMSB est ouvert à toutes questions et demandes de conseil relatives au traitement bioinformatique et biostatistique des données (rendez vous par mail uniquement via contact@prabi.fr).

Le cluster de calcul est complémentaire au service galaxy.prabi.fr qui réclame moins d'investissement de la part de l'utilisateur mais offre moins de flexibilité et de ressources. Par ailleurs, la mise à disposition des ressources se fait en continuité avec les missions d'accompagnement de projets bioinformatique et de formations dispensées par le PRABI-AMSB.

3 - Prérequis et comptes utilisateurs

Pré-requis

1. Avoir des bases concernant l'environnement Linux et les lignes de commandes associées. Pour les utilisateurs inexpérimentés le PRABI-AMSB pourra ainsi proposer **des formations ou des ateliers de mise à niveau** sur les aspects techniques liés à l'environnement, aux lignes de commandes Linux et à l'installation de programmes.
2. Suivre la formation à l'utilisation d'un cluster de calcul, formation **obligatoire** pour les nouveaux utilisateurs. Cette formation dispensée deux fois par an par le pôle informatique du LBBE. Elle est par ailleurs fortement conseillée à tous les autres utilisateurs afin que ceux-ci soient tenus informés de l'évolution du système.
3. **S'inscrire** à la liste de diffusion calcul du LBBE calcul.lbbe@liste.univ-lyon1.fr. Vous serez ainsi informé des dernières nouveautés et de l'état de fonctionnement des ressources. Vous pourrez utiliser cette liste pour partager votre expérience avec d'autres utilisateurs.

Compte utilisateur

Les informations suivantes devront être envoyées à contact@prabi.fr pour la création d'un compte :

- **Prénom, Nom**
- **Adresse mel institutionnelle**
- **Laboratoire d'affiliation**

Pour les non-titulaires, il faudra obligatoirement renseigner :

- **Prénom, Nom, Adresse mel institutionnelle du titulaire responsable**

L'utilisateur s'engage à ne pas partager son login/mot de passe avec d'autres utilisateurs.

Stockage des données

Les utilisateurs sont entièrement responsables de leurs données. Le PRABI-AMSB et le LBBE ne pourront en aucun cas être tenus responsables de la perte éventuelle de données (par exemple lors d'une mauvaise utilisation du service par l'utilisateur, de l'arrêt des ressources informatiques, de la mise à jour du système etc.).

Les ressources étant limitées, pour que l'ensemble des utilisateurs puisse en bénéficier équitablement et dans les meilleures conditions, chaque utilisateur devra sauvegarder ses données brutes (*par exemple* : fichiers de séquences *.fastq*) et devra récupérer/archiver les données générées par ses analyses (*par exemple* : fichiers d'alignement *.bam*) le plus rapidement et régulièrement possible. Par ailleurs, afin d'appréhender au mieux les projets à fortes volumétrie (>500 Go de données brutes), il est demandé de contacter le LBBE/PRABI-AMSB en amont (contact@prabi.fr) afin de discuter de la faisabilité du projet et des solutions disponibles.

La définition des quotas disques reste à l'entière discrétion des administrateurs et peut être modifiée à tout moment afin d'assurer la pérennité des ressources. Les quotas par défaut seront présentés lors des réunions de formations et d'information.

Les utilisateurs seront informés *via* la liste de diffusion [<calcul.lbbe@liste.univ-lyon1.fr>](mailto:calcul.lbbe@liste.univ-lyon1.fr) de toute opération de maintenance et de mise à jour.

Stockage et Archivage des données. Le pôle informatique du LBBE met à disposition un espace de stockage de données « tièdes » (c.a.d. des données non utilisées pour le calcul) sous iRODS. L'utilisation d'iRODS sera présentée lors des formations dispensées par le pôle informatique.

Cycle de vie des données. Pour les étudiants (stagiaires, doctorants, etc.), CDD et tout personnel en fin de contrat (Ingénieur, Post-doctorant, etc.) leur responsable s'engage à rapatrier l'ensemble des données avant la fin de la convention ou du contrat. Les données seront effacées deux mois au plus tard après la fin du contrat ce qui correspond à la durée après laquelle les comptes/badges à l'UCBL sont coupés. Par ailleurs les utilisateurs s'engagent à archiver les données publiées sur des archives appropriées (ENA, SRA, etc.).

4 – Calculs et logiciels

Les utilisateurs s'engagent à suivre les recommandations liées au bon usage des ressources (mémoire et CPU notamment) présentées lors des formations utilisateurs ou communiquées par la liste de diffusion.

Les administrateurs pourront être amené à mettre en place des outils ou des éléments de configuration (files d'attente, priorisation des jobs...) spécifiques pour les utilisateurs de la FRBioenvi.

Les utilisateurs peuvent installer des programmes, logiciels ou modules sur leurs espaces de stockage. Dans ce cas ils sont responsables des aspects liés à la sécurité et à la mise à jour. S'ils considèrent que l'outil est d'intérêt général, ils peuvent demander aux administrateurs à ce qu'il soit installé à un niveau plus global.

Banques de données. Les utilisateurs devront se renseigner de la disponibilité au préalable des bases de données mises à disposition par les administrateurs.

5 - Remerciements associés à l'utilisation du service galaxy.prabi.fr

Pour conclure, les utilisateurs s'engagent à utiliser la phrase de remerciement suivante : ***"This work was performed using the computing facilities of the LBBE/PRABI."***, si des résultats obtenus grâce aux ressources de calculs et de stockages sont utilisés dans le cadre de publications ou tout autres formes de communication (poster, présentation à des séminaires, à des conférences, etc.).

Les personnels impliqués apprécieront de voir apparaître leur nom dans les remerciements pour tout aide ponctuelle apportée ou dans la liste des coauteurs pour une implication significative dans l'accompagnement de projets d'analyse ou de développement bioinformatique.

Rédigé le 04/02/2018,

Vincent Navratil, Responsable Technique - PRABI-AMSB

Guy Perrière, Responsable Scientifique - PRABI-AMSB

Christine Oger, PRABI-AMSB,

Dominique Guyot, PRABI-AMSB,

Bruno Spataro, LBBE

Stéphane Delmotte, LBBE