

Cette lettre d'information est destinée aux membres des équipes de recherche utilisant la plate-forme bio-informatique GenoToul. Elle a pour but de vous informer sur les évolutions de l'équipe, les nouveaux outils, services, projets et formations mis en place.

1/ Évolution de l'équipe :



Dans le cadre du partenariat avec la plate-forme PlaGe (<http://genomique.genotoul.fr/>), Nicolas Allias vient de rejoindre l'équipe, le 1^{er} octobre, pour traiter les données issues du 454 et développer des modules pour les chaînes de traitements génériques élaborées par la plate-forme.

Son CDD a une durée de 18 mois.

2/ journée régional GENOTOUL bio-info le jeudi 3 décembre 2009 :

Cette journée, soutenue par le GIS GenoToul, se veut être l'occasion d'initier une série de rencontres et d'échanges entre les équipes de bio-informatique régionales, aussi bien dans le cadre de leur activité de recherche méthodologique que, sur le plan opérationnel, de l'exploitation des données par la mise oeuvre ou le développement d'outils appropriés.

Nous invitons l'ensemble des équipes régionales travaillant dans le domaine de la bio-informatique (organismes publics et entreprises privées) à proposer une présentation en deux parties dont la première, brève (5mn), décrira les thématiques traitées dans leur équipe, et la seconde, plus longue (environ 20mn), développera un sujet particulier.

Nous limitons volontairement le nombre de présentations par équipe à une au plus, afin que chaque équipe ait l'occasion de s'exprimer.

Ciblée sur les développements en bio-informatique, cette journée offrira aussi l'occasion de présenter le nouveau master de Bio-informatique et Biologie des Systèmes de l'Université Paul Sabatier.

La journée est ouverte à tous.

Lieu: Salle de conférence de l'INRA, Auzeville

- Date limite de soumission de résumé : 15 novembre 2009
- Date limite d'inscription : 20 novembre 2009

Pour soumettre un résumé et/ou vous inscrire : <http://bioinfo.genotoul.fr/main/index.php?id=99>

3/ Évolution de la configuration matérielle : nouveaux serveurs, nouveaux équipements réseaux :

Afin de répondre au besoin de continuité de service et d'évolution de nos serveurs, une solution de virtualisation des systèmes vient d'être mise en place :

Cette solution s'appuie sur différents matériels et logiciels notamment:

- La baie de disques CLARION (SAN) de marque DELL
- Le logiciel VSPHERE Entreprise 4 de l'éditeur VMWARE
- 2 switch Fiber Channel 4Gbs de marque BROCADE
- 2 switch Gigabit Ethernet de marque HP
- 3 nouveaux serveurs de marque BULL

Chacun des serveurs possède quatre processeurs INTEL (hexacore xeon) et 64 Go de RAM (extensible à 256 Go).

Nous disposons à ce jour de 22 machines virtuelles pour une estimation à 50 d'ici 2/3 ans.

Cette nouvelle infrastructure évolutive est tolérante aux pannes : basculement automatique des services sur les machines disponibles.

Elle remplacera les serveurs qui ne sont plus sous contrat de maintenance. A ce jour, tous les serveurs accessibles par les utilisateurs de la plate-forme ont été migrés sur cette nouvelle architecture hormis TATA, POLYA et INTRON.

Les projets hébergés sur le serveur TATA vont progressivement être transférés sur de nouvelles machines virtuelles (regroupées par fonctionnalité ou par organisation).

Les bases de données seront hébergées sur un serveur dédié : INTRON.
Le serveur POLYA servira de serveur de fichiers et de serveur de bases de données de secours.

Perspectives :

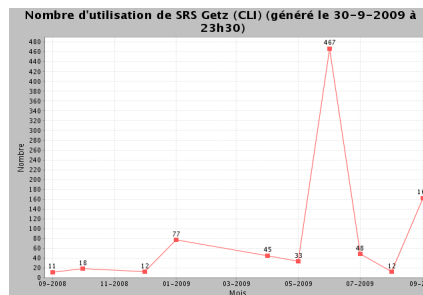
Étant donnée la forte augmentation de la volumétrie des données et l'arrivée en fin de vie de la solution actuelle de sauvegarde, une nouvelle infrastructure de stockage / sauvegarde doit être mise en place. Ce projet donnera lieu à des consultations et à des acquisitions courant 2010/2011.

4/ Evolution de l'accès aux banques de données publiques sur la plate-forme :

La plate-forme bio-informatique Genotoul tient à jour et met à disposition des utilisateurs des banques publiques pouvant être interrogées soit par recherche de similarités (NCBI blast) soit par recherche textuelle à l'aide de l'outil SRS.

Nous envisageons dans les prochaines semaines d'arrêter le service SRS pour le remplacer, pour certaines banques, par un service de recherche textuelle basée sur l'outil lucegene, faisant partie de la suite GMOD (<http://gmod.org/wiki/Lucegene/>).

Comme le montre le graphique de droite l'interrogation des banques en mode texte est très peu utilisée en ligne de commande, mais aussi depuis le serveur web. Le niveau d'utilisation ne justifie pas le coût annuel de la licence du logiciel.



Les banques listées ci-dessous sont actuellement toutes accessibles via SRS. Seules celles figurant dans la première partie du tableau seront transférées sous lucegene.

genbank uniprot dbest geneontology taxonomy enzyme	Transférées sous lucegene
interpro prosite rfam omim prosite pfam	Non transférées (fichiers bruts accessibles ou interrogeable par blast)

Cette opération se fera durant les mois de novembre et décembre 2009.

5/ Labellisation IBISA et animations RENABI:

La plate-forme a obtenu la labellisation IBISA en 2009. Le projet financé permettra :

1. d'appuyer l'équipe dans sa mission de mise à disposition des banques de données ;
2. de poursuivre les développements de la plate-forme d'annotation des ARN non codant, RNAspace (<http://www.rnaspaces.org/>), en collaboration avec une équipe de l'IGM (Orsay), l'équipe SEQUOIA du LIFL/INRIA (Lille) mais aussi avec les équipes toulousaines INRA SIGENAE et SaaB,
3. de remplacer, en 2010, le robot de sauvegarde ;
4. de renforcer les capacités de calcul en vue des besoins générés par l'exploitation des données issues du séquençage haut débit.

6/ Quelques animations à venir du Réseau National de Bioinformatique (ReNaBi):

1. Journées BIOGRALE (BIOinformatique à GRAnde échelle) les 24 et 25 novembre 2009 à Rennes
2. Journée "Evolution et biodiversité bactérienne : impact du séquençage de nouvelle génération" le 07 décembre 2009 à l'Institut Pasteur
3. Journées « Bioinformatique des ARN non codant » les 24, 25 et 26 février 2010 à Bousens (31)

7/ Formation unix et utilisation du cluster de calcul:

Une nouvelle session de formation pour utiliser l'infrastructure de la plate-forme et les outils de bio-informatique mis à disposition est prévue les 8 et 9 mars 2010.

Pour toute demande d'information ou de travaux, veuillez envoyer un mail à support.genopole@toulouse.inra.fr en précisant vos nom et coordonnées.