



Le réseau régional d'ingénieurs en bioinformatique de Lille, le PPF bioinformatique et le PPF calcul scientifique intensif vous convient à une conférence jeudi 23 Mai 2013, à 15h30, Amphithéâtre de l'Institut de Biologie de Lille, 1, rue du Pr Calmette, LILLE. Ce séminaire sera donné en français et suivi d'un tutoriel le 24 mai de 9h00 à 12h00, salle de formation du CRI, Bat M4, Université Lille1. Inscription obligatoire pour le TP (limitée à 20 participants). Contact : sophie.gallina@univ-lille1.fr

IDB-cloud : Providing Bioinformatics Services on Cloud

Christophe Blanchet

Infrastructure Distribuée pour la Biologie – CNRS-IBCP FR3302 – Lyon

Recent improvements in biological experimental technologies force scientists to face a deluge of data that require relevant tools and sufficient computer resources for their analysis. Then scientists need large scale infrastructure, but they also often combine many bioinformatics tools from the arsenal of thousands available from the international community. Common interfaces are standard portals and web services for the ease of use, but also the command line to run scripts or their own software. All these analyses rely most of times on reference public databases that need to be easily available (NFS) and updated. Such multiple requirements can be very difficult to satisfied in public infrastructures (clusters, portals) available to a whole scientific community such as in Biology. To solve these difficulties, we have deployed a cloud infrastructure, devoted to Biology, that ease the on-demand deployment of bioinformatics infrastructures. We developed several bioinformatics appliances ready to run on a standard cloud infrastructure. For example one of these appliances provides scientists with their own dedicated Galaxy portal with common NGS tools. Several other virtual appliances are also available and provide scientists with a large range of biological tools. For example the Proteomics appliance with OMSSA and X!Tandem tools, or the BioMapReduce appliance with the standard Hadoop MapReduce framework coupled to common bioinformatics tools. All these virtual machines are connected to updated public databases such as Uniprot, EMBL, Prosite ; or to different genomes and indexes such as HG19. The IDB-cloud is available to the academic community to run their own analysis, on a registration basis and standard use. The use for different purposes will help to identify the different requirements from the Bioinformatics community, and also to evaluate such cloud infrastructure in the context of the Institut Français de Bioinformatique that is planning to deploy a cloud core facility for the French community.

