



**Master Informatique, Mathématiques, Multimédia & Télécommunications**  
**Spécialité « informatique »**  
**Spécialité « Réseaux de Télécommunications, Multimédia et Automatique »**

## **Proposition de sujet de Stage Recherche 2016-2017**

**Titre: *Sélection économique des structures d'optimisation des bases de données sur cloud***

**Laboratoire d'Informatique et d'Automatique pour les Systèmes (LIAS/ISAE-ENSMA)**

**Encadrant(s) : Ladjel BELLATRECHE (bellatreche@ensma.fr, 05 49 49 80 77)**

**Mots clés : cloud, optimisation de requêtes, modèles économiques, vues matérialisées**

### **Description du sujet :**

La technologie de bases de données a subi des évolutions spectaculaires durant une période assez courte: de bases de données centralisées au cloud. Le cloud se propose comme solution afin de stocker des masses de données à peu de frais et de manière efficace. Pour optimiser cette masse de données, des structures de données doivent être définies comme les vues matérialisées, les index, etc. Si nous nous penchons sur la littérature, ses structures sont souvent sélectionnées en utilisant des modèles de coût qui prennent en compte que les aspects liés à l'optimisation (le temps de réponse de requêtes), mais négligent l'énergie consommés par les data centers. Pour offrir une optimisation "verte" (*green computing*), la sélection des structures d'optimisation doit être guidée par les aspects énergétiques. Si un propriétaire des données exige un temps de réponse rapide, il doit payer plus cher. Dans ce cas, un modèle économique prenant en compte les contraintes de temps de réponse et les aspects énergétiques doivent être pris en compte dans l'esprit pollueur payeur. Ce sujet rentre dans le cadre de thèse de Mr. Ahcène BOUKORCA qui étudie le développement de structures de données efficaces pour le cloud.

Le travail demandé s'organisera en plusieurs étapes :

1. l'étude bibliographique autour de la technologie cloud
2. identification l'ensemble des paramètres permettant le développement d'un modèle de coût (des paramètres liés à la base de données, les structures de données, le cloud, les aspects économiques, le profil de propriétaire des données, etc.);
3. la proposition un algorithme de sélection dirigée par un modèle de coût prenant en compte les aspects énergétiques et temps de réponse;
4. le prototypage de la solution proposée et expérimentation sur une plateforme cloud.