

*BD**2 : des BD à BD*

(Des Bases de Données à Big Data)

Professeur Serge Miranda, (Directeur MBDS, Université de Nice, France)

Mail: miranda.serge@gmail.com

GUEST SPEAKER/ CONFÉRENCIER INVITÉ



Serge Miranda est Professeur des Universités en informatique à l'Université de Nice Sophia-Antipolis (UNS) depuis octobre 1983. Il a soutenu une thèse d'État à l'Université Paul Sabatier de Toulouse en 1980 et une thèse de Master à UCLA (Univ de Californie Campus de Los Angeles) avec une bourse de postdoc de l'IRIA en 1976. Il est devenu professeur des universités de « classe exceptionnelle » en 2011.

De 1982 à 1998, il a créé et dirigé l'équipe recherche en base de données de l'UNS qui a « produit » plus de 16 enseignants chercheurs dont six Professeurs des Universités.

Depuis 1998, il dirige le Master « MBDS » de l'UNS (*Mobilité, Bases de Données et intégration de Systèmes*) qu'il a fondé comme D.U. en 1992-1993 suite aux donations logicielles et matérielles de partenaires industriels. MBDS a été délocalisé au Maroc, Tunisie, Madagascar, Haïti et Russie. MBDS est original par son centre de ressources et ses partenariats industriels qui en font un véritable « laboratoire d'innovation sur les services d'information du futur » avec un modèle de valorisation original. MBDS est devenu un véritable d'innovation sur les services d'information du futur dans les domaines de la maison, de la voiture, du tourisme, du paiement et de la ville intelligente.

Serge Miranda a été, de décembre 2009 à Juin 2011, le premier président de la fondation partenariale universitaire DreamIT (de l' UNS) issue de la loi sur l'autonomie des universités pour conduire de manière pluridisciplinaire de l'innovation et de la recherche sur les services d'information du futur à la convergence du téléphone mobile devenu ordinateur et d'Internet (*mobilité*). La fondation DreamIT, qu'il a créée sur le réseau des partenaires du MBDS, a conduit des projets de prototypage de services mobilitaires innovants au Maroc, en Inde, en Russie et en Haïti. En Avril 2011, a été inauguré le premier bâtiment reconstruit de la faculté de sciences de l'Université d'Etat d'Haiti qui héberge le master MBDS délocalisé à la faculté des Sciences sur un modèle du futur: antisismique, à énergie renouvelable (éolien et solaire) et piloté complètement par un téléphone mobile NFC (le premier bâtiment de ce type des Amériques).

ABSTRACT / RÉSUMÉ

Nous sommes rentrés dans l'ère des Tsunami de données temps réel venant du Web, des réseaux sociaux (twitter, Face Book, Instagram,...), des objets tagués (NFC, QR Code, OCR), des capteurs (smartphones), ... représentant plus d'un Exaotets(10^{18}) chaque jour fin 2012 avec des projets visant le Yotta octets de données (10^{24}) en 2020.

BIG DATA intègre aussi bien les données publiques de type OPEN DATA (CSV, PDF, RDF, ..), les LINKED DATA (Web Sémantique et ses standards RDF, OWL et SAPRQL) et les MOBIQUITOUS DATA (venant des objets tagués). Tous les secteurs scientifiques vont être touchés et tous les grands secteurs de l'économie mobiquitaire (tourisme, commerce, MNO, m-paiement) et de la gestion des risques (avec des cartographies temps réel).

M.Stonebraker a donné les « 4 V » caractérisant le BIG DATA (Volume, Velocity, Variety, Variability) et posé les bases d'une discussion que nous étudions dans cette conférence sur les passerelles entre le monde des Bases de Données SQL et le monde BIG DATA.

Dans cette conférence, après avoir défini le concept de Big Data, nous comparons les 3 approches essentielles pour appréhender cet univers de données :

- SQL3/ODMG basée sur une approche top down, avec un schéma fixe, les propriétés TIPS et ACID des transactions
- NO-SQL (Not Only SQL) basée sur une approche bottom up, temps reel avec un schema variable, les propriétés BASE et le théorème CAP ; les 4 types de base NO SQL seront abordées sur un exemple
 - o Clé-valeur (Ex Hadoop et MAP Reduce)
 - o Orientée Colonne (Ex Big Table)
 - o Orientée Document (Ex Mongo DB)
 - o Orientée Graphe (Ex Neo4J) pour les réseaux sociaux
- NEW SQL de systèmes hybrides entre SQL et NO SQL avec les 2 types de base
 - o Les SGBD Spécifiques comme VoltDB (de Stonebraker), MySQL ou NImbusDB
 - o Les passerelles Hadoop vers le monde SQL (Oracle, Ibm, Microsoft et Teradata)

KEYWORDS / MOTS-CLÉS

(EN) BIG DATA, OPEN DATA, MOBIQUITOUS DATA, SQL3/ODMG, NO-SQL, NEW SQL, système hybride.