Soutenance de thèse de Baraa MOHAMAD

Nom : MOHAMAD, **Prénom** : Baraa

Laboratoire : LIMOS

Directeur de thèse : Pr.Farouk Toumani Université Blaise Pascal, France

Pr. Le Grenwald Université d'Oklahoma, United-State

Codirecteur de thèse : Dr.Laurent D'orazio Université Blaise Pascal, France

Date de soutenance : 23 juin 2015

Membre de jury :

Rapporteurs:

Pr.Fréderique Laforest Télécom Saint-Etienne, Université Jean Monnet, Franc

Dr. Genoveva Vargas-Solar Laboratoire d'Informatique de Grenoble, France

***Examinateur*:**

M.Serge Torti Chef d'entreprise YANSYS Medical, France

Directeurs:

Pr.Farouk Toumani University of Blaise Pascal, France

Pr.Le Grenwald University of Oklahoma, United-State

Dr. Laurent d'Orazio University of Blaise Pascal, France

Titre de thèse : Medical Data Management on the cloud

Résumé :

Medical data management has become a challenge due to the emergence of new imaging technologies providing high image resolutions.

This thesis focuses in particular on the management of DICOM (i.e. Digital Imaging and Communications in Medicine) files. DICOM is one of the most important medical standards. DICOM files have special data format where one file may contain regular data, multimedia data and services. These files are extremely heterogeneous (the schema of a file cannot be predictable) and have large data sizes. The characteristics of DICOM files added to the requirements of medical data management in general – in term of availability and accessibility- have led us to construct our research question as follows:

Is it possible to build a system that: **(1) is highly available, (2) supports any medical images (Different specialties, modalities and physicians' practices), (3) enables to store extremely huge/ever increasing data, (4) provides expressive accesses and (5) is cost-effective .**

In order to answer this question we have built **a hybrid (row-column) cloud-enabled storage system**. The idea of this solution is to disperse DICOM attributes thoughtfully, depending on their characteristics, over both data layouts. All with exploiting the interesting features of the cloud that enables us to ensure the availability and portability of medical data. Storing data on such hybrid data layout opens the door for a second research question, **how to process queries efficiently over this hybrid data storage with enabling new and more efficient query plans**

The experimental prototypes implemented in this thesis show interesting results and open the door for multiple optimizations and research questions.