MEDINFO 2015: eHealth-enabled Health
I.N. Sarkar et al. (Eds.)

doi:10.3233/978-1-61499-564-7-929

A Medical Image Backup Architecture Based on a NoSQL Database and Cloud Computing Services

Luan Henrique Santos Simões de Almeida[a], Marcelo Costa Oliveira[a]

[a] *Laboratório de Telemedicina e Informática Médica, Instituto de Computação, Universidade Federal de Alagoas, Maceió, AL, BR*

Abstract

The use of digital systems for storing medical images generates a huge volume of data. Digital images are commonly stored and managed on a Picture Archiving and Communication System (PACS), under the DICOM standard. However, PACS is limited because it is strongly dependent on the server's physical space. Alternatively, Cloud Computing arises as an extensive, low cost, and reconfigurable resource. However, medical images contain patient information that can not be made available in a public cloud. Therefore, a mechanism to anonymize these images is needed. This poster presents a solution for this issue by taking digital images from PACS, converting the information contained in each image file to a NoSQL database, and using cloud computing to store digital images.

Keywords:
Cloud Computing; NoSQL Database; PACS.

Introduction

The Picture Archiving and Communication System (PACS) is an approach widely used to store digital images. In 2010, Teng et al. [1] stated that in 2014 the United States produced approximately 100 Petabytes of medical images. Nevertheless, PACS is limited in current scenarios, with high investment in data centers and limited physical storage. Therefore, adopting a scalable environment that can persist data as long as needed and withstand varying loads is essential.

Cloud Computing is able to provide an elastic and appropriate environment for a set of applications besides providing hardware and software resources as a service, providing a low-cost storage with high reliability and security [2]. Thus, cloud computing becomes the ideal setting for several health applications, providing a great resource for storage and digital image processing. However, DICOM images contain private patient information; therefore, a mechanism to anonymize these images is necessary. The objective of this poster is to build an automated backup environment of anonymized medical images derived from PACS into a public cloud, powered by a NoSQL database.

Materials and Methods

The experiments were developed using the PACS solution from dcm4che project, the Amazon S3 public cloud and 1761 computed tomography (CT) images in the DICOM standard. The patient information contained in the DICOM file was stored in a MongoDB NoSQL database since these DICOM files are formed by tags that are similar to the key-value documents in MongoDB. As shown in Figure 1, the images are generated by acquisition modality equipments, sent to the PACS, read by the backup manager, and uploaded to the cloud after removing personal information.

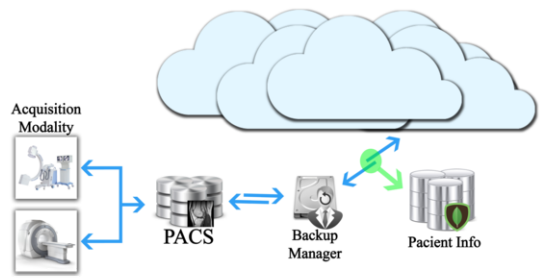

Figure 1 - Architecture overview

Results and Conclusion

The environment was built to be a fail-safe system by using MongoDB documents as log reports. Through these log reports, the environment will be able to recognize in which task the backup process stopped and continue from that point. The reverse way is also possible and is in development. In other words, the system will be able to identify the patient ID from the patient name, recover the file from the cloud, then place the patient information inside this file, and identify them. The query for patient information was optimized due to the use of MongoDB documents instead of regular SQL databases since it is possible to parallelize queries and these documents consisting of key-value pairs. Using a NoSQL database allows integration between other health bases within the big data concept in a near future.

References

[1] Teng CC, Mitchell J, Walker C, et al. A medical image archive solution in the cloud. International Conference on Software Engineering and Service Science (ICSESS), IEEE, 2010; 431-434.

[2] Zeng W, Zhao Y, Ou K, Song W. Research on cloud storage architecture and key technologies. Proceedings of the 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human, 2009. p. 1044–1048, New York: ACM.

Address for correspondence
luanhssa@ic.ufal.br