*LIFE: International Journal of Health and Life-Sciences*
*ISSN 2454-5872*

*Badarneh et al., 2016*

*Volume 2 Issue 3, pp. 18-25*

*Date of Publication: 19th November, 2016*

*DOI-https://dx.doi.org/10.20319/lijhls.2016.23.1825*

*This paper can be cited as: Al-Badarneh, A., Husari, G., & Najadat, H. (2016). A Model for Managing Medical Image Data on the Cloud. LIFE: International Journal of Health and Life-Sciences, 2(3), 18-25.*



# A MODEL FOR MANAGING MEDICAL IMAGE DATA ON THE CLOUD

**Amer Al-Badarneh**
*Associate Professor, Jordan University of Science and Technology, Irbid, Jordan*
*amerb@just.edu.jo*

**Ghaith Husari**
*Research Assistant, University of North Carolina at Charlotte, North Carolina, USA*
*ghusari@uncc.edu*

**Hassan Najadat**
*Associate Professor, Jordan University of Science and Technology, Irbid, Jordan*
*najadat@just.edu.jo*

## Abstract

*Due to the increasing demand on Hospital Information System (HIS) resources and the high cost of constructing independent information platform especially for medium to small hospitals, large scale clusters which are based on cloud computing technologies have been very popular to use in order to have a quality information platform that best suits the hospital budget. In this paper, we propose an enhanced model that provides a solution for hospital information systems. This model provides new features that will enhance the efficiency of data retrieval of medical images and the costs of processing these images in the cloud.*

**Keywords**

Cloud Computing, Healthcare Cloud, DICOM Protocol Server, Hospital Information System, Cloud of HIS Architecture

## 1. Introduction

Hospitals have massive amounts of high resolution medical images which require massive storage data centers. In healthcare systems, massive storage leads to the need of acquiring efficient techniques for storing, retrieving, and processing all of these high resolution medical images in timely manner as it involves uploading all of these images to the cloud. There are several solutions to mitigate the overhead of uploading and downloading these images, which sometimes happens very frequently and that leads to extra costs of retrieving and processing the medical images. In this paper we propose a new technique and an enhanced model to make these operations more efficient in order to have effective use of cloud computing services which leads to saving some of the costs of cloud computing in terms of computation power, reserved space and the bandwidth required. Also, apart from the costs the proposed model achieves faster performance of the whole system.

## 2. Previous Work

(Iron Mountain, 2015) presents the use of hybrid cloud and multi-tier storage. This model has onsite and offsite storage techniques that affect the efficiency of data retrieval of the entire system. The onsite part involves storing some medical images on the hospital server. This kind of data which qualifies is to be stored on the hospital server is very frequently retrieved and processed so it is a very reasonable idea to keep them on a local storage for a short period of time before uploading them to the cloud. However, the proposed solution did not specify in the proposed archiving techniques which images qualifies to be stored locally and which are not. Therefore, this work can be enhanced to obtain a better efficiency of retrieving and processing images.

(Yang, Chenb, & Yang, 2010) proposed a model for processing medical images queries using co-allocation technique with several replicas in data-grid environments. The model implements Medical Image File Accessing System (MIFAS). A data grid is a model that has many servers working together to deal with medical information in distributed environments. MIFAS main goal is to simulate Picture Archiving and Communication System (PACS) over

data grid environment using co-allocation strategy. That aims to handle and mitigate PACS challenges and limitations like load-balancing, fault-tolerance, security and stability.

### 2.1 The Multi-tier Framework for Medical Image Storage

In Figure 1, we can see a multi-tier model (Iron Mountain, 2015) which has several tiers. In this model, there are several tiers to store medical images. Each tier performs a very important functionality in terms of efficient retrieval for medical image on the cloud. The first tier is PACS storage archive which is where the PACS reside in the proposed model. The second tier is a high availability gateway to provide medical images rapidly and frequently in a very fast fashion. Then there are two additional archives to provide medical images for retrieving and processing. These images have been taken in a certain date, for example tier 1 typical capacity is for six to twelve months. These archives directly connect to rapid gateway according to their date and usage terms.

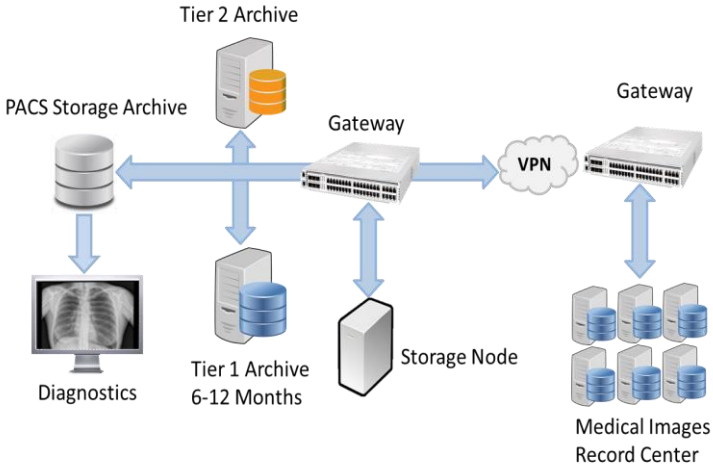


**Figure 1:** *Multi-tier Medical Storage*

### 2.2 File Accessing System in a Co-Allocation Data Grids

The proposed model in (Yang, Chenb, & Yang, 2010) is to solve some cloud availability and performance problems by presenting a MIFAS in co-allocation data grids. The framework of the proposed mode is shown in Figure 2.

When an image query is sent to the system with the desired image features, it is forwarded to the broker that has the ability to look for the corresponding replicas which the queried image reside on. When all replicas are identified, the broker chooses the best replica, which has the best availability to the physical location of the requesting client. Then the broker returns the physical

address of that replica to the client in order for it to retrieve the desired medical image. MIFAS use Anticipative Recursively-Adjusting Mechanism (ARAM) (Yang, Chen, Chou & Wang, 2010) for co-allocation because ARAM has good management of replicas and network fault-tolerance. This solution has the following advantages:

- Reduce transfer time for co-allocation of medical images.
- Has a better form of network fault tolerance.
- Has a parallel download feature by using several replicas to download medical images for the same client.
- Overcomes PACS by providing a more stable and reliable system to retrieve medical images from the replicas.

The proposed model in (Yang, Chenb, & Yang, 2010) is an efficient one for data retrieval and parallel usage of data grids which contains the hospital medical images. However, the co-allocation data grids don't archive the medical images in the best usage fashion as they don't consider frequency of retrieving. In our proposed model, this architecture and archiving model is tailored to achieve more efficient solution.

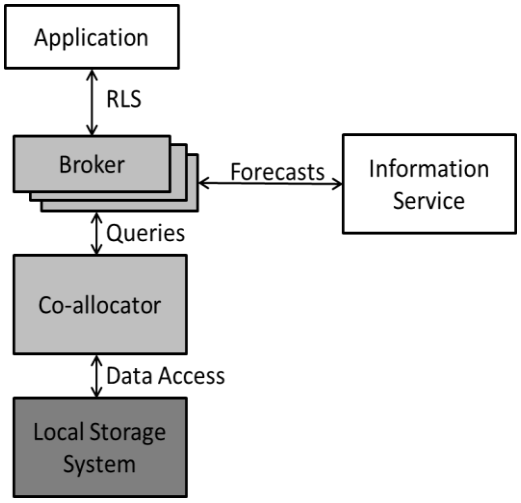


**Figure 2:** *File Accessing Workflow in Co-allocation Data Grids*

## 3. The Proposed Model

As mentioned earlier in this paper, medical images have higher resolutions so they require large storage to store, higher bandwidth to retrieve, in addition to more computation power to process. So a solution that takes into consideration the relevance of the image to the patient's

case and the frequency that the image is going to be retrieved will be more efficient to use in order to have faster and less costing model.

The probability of an image to be frequently retrieved is directly connected to the relevance of the image. In other words, even if an image is recently taken, that doesn't mean it will be frequently retrieved because the image might be not very important, such as full body scans. These images cost a lot of storage and retrieving time, but at the same time they don't often help in medical diagnoses, unless for the diseases which can be diagnosed by taking way less costing medical images such as x-rays, PET, CT scans and MRIs. So the archiving technique can be enhanced to achieve better performance.

We use DICOM standard (Moreno et al., 2002) to handle image format and communication protocol, and for data storage we tailor the multi-tier architecture to achieve better efficiency. We use DICOM format which deals with medical image format and communication amongst its formats. DICOM clients are the medical devices which contact the DICOM to store pictures such as MRI, CT-scan, ultrasound, and X-ray. DICOM standard doesn't store the image alone, but also it stores metadata along with it. Metadata are additional information about the picture that describes the picture and will be used for retrieving it after it's stored. The DICOM servers are responsible for DICOM images. Also the protocol defines the communication protocol between servers. The DICOM server extracts metadata and stores it in an SQL DB in order to retrieve and search for this information in a faster more efficient fashion. The DICOM image indexers extract header of images and its metadata to store in SQL DB. Storing images in the cloud will be based on three aspects, which are: storage, two-tier technique, and co-allocation. We discuss each of these aspects next.

- **Data Center Storage**

The most important feature of the storage system is the high speed between shared nodes. The virtualized shared disk database management system has the advantage of applying gigabit Ethernet (up to 12 GB/s) to obtain the best possible sharing speed.

- **Two-Tier Storage Technique**

In this proposed model, we present private storage and public storage. The private part has a gateway which is the access point to it. Also it has two other archives. The first one stores the medical images which have the highest probability of being accessed by the users (image that was proven relevant). The second one stores the recent medical images. So users' data in hospital

are got from the relevant archive or the recent archive, if a medical image isn't found in both archives, the public part is called. The public part stores the archives of the medical images data of all the hospitals in the cloud. To securely exchange data between both parts we use the Virtual Private Network approach (VPN).

- **Co-Allocation File Accessing**

The public part which contains all hospitals public data has the risk of attack or natural catastrophic events. So it must be replicated to avoid these risks. So it needs a protocol to locate the closest replica of the requested data to have a better performance. There is an already used technique which requires forwarding the request to a broker that has the ability to search for the available replica which holds the requested medical data and can select one of the replicas based on aspects of access time and bandwidth. There is also an already tested and used strategy that is called ARAM strategy which retrieves the physical file location of the replica.

## 4. The Model Functionality

Our proposed model main functionalities are: storing images, requesting images, and moving images between private-public parts when several conditions apply to them.

- **Storing image in DICOM server**

When a medical image is sent by DICOM client to the DICOM server, the image is sent to indexer which extracts its metadata and stores it in SQL DB (Moreno et al., 2002). Then the entire image is stored in the "most recent" archive.

- **Requesting image from DICOM server**

  When a client from the private side (inside the hospital) requests an image, a query is sent to the DICOM server and retrieved. If the image is not found on the DICOM server the query is forwarded to the SQL DB. If the image isn't found on the SQL database then it is on the public part and the broker returns the address of the best replica that has the requested image.

- **Transferring medical images between servers**

The medical images that reside on the private side must be transferred to the public side every agreed-on time to ensure its safety and integrity. The migration of the medical data must be done based on two time thresholds.

There are two local archives to be used in our model, the relevant and the recent archives. The relevant archive contains images are supposed to be relevant to the patients case. An image is considered to be relevant if the physician who is responsible for the case marked the image as relevant, otherwise it is not considered so. The other archive is the recent archive. This one stores images that were taken recently and were not marked as relevant by the physician. The images which reside on the relevant archive will be migrated after the patients case is closed, while images which reside on the recent archive will be migrated after an agreed-on time.

## 5. Summary and Conclusion

Our proposed model is intended to store and retrieve medical images more efficiently in the cloud by using DICOM server, Two-tier storage, and co-allocation. These three techniques solve more challenges in the cloud, like the required bandwidth for data retrieval of medical images and the costs of computation power for processing these images in the cloud, in addition to the required space in the cloud server which the hospital images will reside on. We hope that this enhancement may make retrieving medical images more effective and less time and monetary-cost consuming by the use of better archiving techniques for relevant and most recent images.

In the future a simulation of the proposed technique to provide results of using two-tier architecture storage and the new proposed archiving method that is meant to show how it achieves better performance in terms of data retrieval of medical images and the costs of processing these images in the cloud.

## References

Iron Mountain (2015). Retrieved from www.ironmountaineurope.com/~/media/Files/Iron Mountain/Knowledge Center/Reference Library/Brochures/Digital Record Center for Medical Images.pdf

Moreno, R. A., & Furuie, S. S. (2002, September). Integration of heterogeneous systems via DICOM and the clinical image access service. InComputers in Cardiology, 2002 (pp. 385-388). IEEE.

Yang, C. T., Chen, C. H., & Yang, M. F. (2010, October). Implementation of a medical image file accessing system in co-allocation data grids. Future Generation Computer Systems, 26(8), 1127-1140.

Yang, C. T., Chen, L. T., Chou, W. L., & Wang, K. C. (2010, December). Implementation of a medical image file accessing system on cloud computing. In Computational Science and Engineering (CSE), 2010 IEEE 13th International Conference on (pp. 321-326). IEEE.