A Flexible Storage Architecture for Large PACS

SS Furuie, N Bertozzo, M Yamaguti, MA Gutierrez

Heart Institute (InCor), University of São Paulo Medical School, Brazil

Abstract

PACS and Electronic Patient Record require large storage capacity in order to be fully exploited. A medium size Cardiovascular Hospital (500 beds) needs around 20TB for a period of 10 years.

The proposed solution combines online, nearline and offline modes, a database and a managing software in a cost-effective architecture without compromising fast recovery, scalability and safety. The data flow from one media to another, storing the oldest block of exams to less costly media. Furthermore, the online subsystem is easily scaled, since in order to expand the online visibility, it is enough to insert another storage sub-system in the network, and update the configuration file.

The described architecture has been operational at the Heart Institute since April 2000, with two RAID5 subsystem (200GB and 500GB) and a 3.5 TB DLT jukebox with 52 slots as nearline. Currently, more than 2.5TB of DICOM images have been archived.

1. Introduction

PACS and Electronic Patient Record require large storage capacity in order to be fully exploited. A medium size Cardiovascular Hospital (500 beds) needs around 20TB for a period of 10 years. Furthermore, the storage process should be flexible and cost-effective.

For instance, solutions for storing and retrieving cineangiographic images are difficult due to high volume of data (typically 300 Mbytes per exam) and dynamic requirements for visualization. DICOM [1] servers do not handle long term storage of medical images. Old images are just discarded automatically as soon as predefined threshold of free disk space is reached.

A cost-effective architecture for storage is possible without compromising fast recovery, scalability, and safety. This paper presents an affordable and scalable storage for large PACS systems.

2. Methodology

The conditions that must be met are:

a) online scalability with immediate access (< 3s) to recent data, where the fast access period (for instance, 6 months) can be easily configured;
b) long-term autonomy with operator-free access (nearline mode) to old exams (for instance, 6 - 24 months);
c) permanent storage (offline mode) for very old exams (for example, older than 24 months);
d) safe archiving;
e) cost-effective solution.

The proposed architecture combines online, nearline and offline modes in an asynchronous flow of data (figure 1). The flow is controlled by a managing software ('manager') based on information archived in a database.

The online storage is based on a RAID-5 magnetic disk system. The access time is clinically adequate (less than 3s) using fast networking (fast Ethernet and Gigabit Ethernet.)

For the nearline storage we have used a jukebox of DLT (Digital Linear Tapes) with 48 slots and 4 drives. Multiple drives are necessary to allow simultaneous reading and storage operations, improving client response time.

The offline storage is simply the tapes on the shelves.

Basically, when a defined threshold of online free-space storage is reached, the manager copies the oldest block of exams to an available tape and updates the database. If the jukebox is full, the oldest tape is moved to the shelf, and again, the database is updated[2]. However, the online capacity, that determines the temporal window of fast recovery (online visibility), changes with the inclusion of new equipments or clinical requirements. Therefore, the architecture should be flexible in order to accommodate dynamic changes.

2.1. Archiving images

The most critical storage subsystem is the online mode, since it should be configurable, reliable and fast. The nearline mode should support large volume of data and handle, automatically, individual media, which eventually might be stored on the shelves as offline mode. Therefore, each media, 70GB DLT in our case, should be identifiable and its content be known by the database manager.

Each DICOM equipment can be configured to send data to a specific online storage subsystem, thus allowing a flexible mapping of equipments, modalities and storage devices.

The proposed architecture (figure 1) consists of several (configurable) online storage subsystems in the network: ArchN, ...,Arch1 and Arch0. Each subsystem uses redundant disks, such as RAID5 devices, assuring good level of data safety. The 'Arch0' subsystem has a special function of buffering between the online and nearline modes.

Periodically, the managing software checks a configuration file (manager.ini), which contains the number of online subsystems and some defined characteristics of each one, such as network path and free-space threshold. Whenever the free-space in any ArchN...Arch1 is below a threshold (say 10GB), a block of exams (for instance, 5GB) is moved into 'Arch0' and the database is updated. If the free-space of 'Arch0' is below its threshold, a block of old data, equivalent to the capacity of the removable media, is moved into the nearline system and the database is updated. The first file in the DLT tape is a list of the content of the media. If the jukebox is full, the oldest media is moved to the shelf, new one is inserted and the database is updated again. Thus, the 'Arch0' subsystem integrates online and nearline subsystems, and the database has a complete record of the current status of the images.

The online storage system can be easily expanded simply adding another storage subsystem and updating the configuration file. The buffer 'Arch0' assures proper connection and throughput to the nearline subsystem.

The online visibility could also be extended if we copy a block of exams to the nearline subsystem, instead of moving, and delete the exams in 'Arch0' on demand.

In steady state, all archiving subsystems, except the shelves, operate with the maximum allowed load, optimizing the window of visibility for the online and nearline modes.

2.2. Managing data flow

The system should have complete knowledge of image location through image database updates for each file movement[2]. It plays a central role in the optimization for storing and retrieving medical exams (figure 2).

On the other hand, image database should be integrated with other databases, such as Hospital Information System (HIS) and Radiological Information

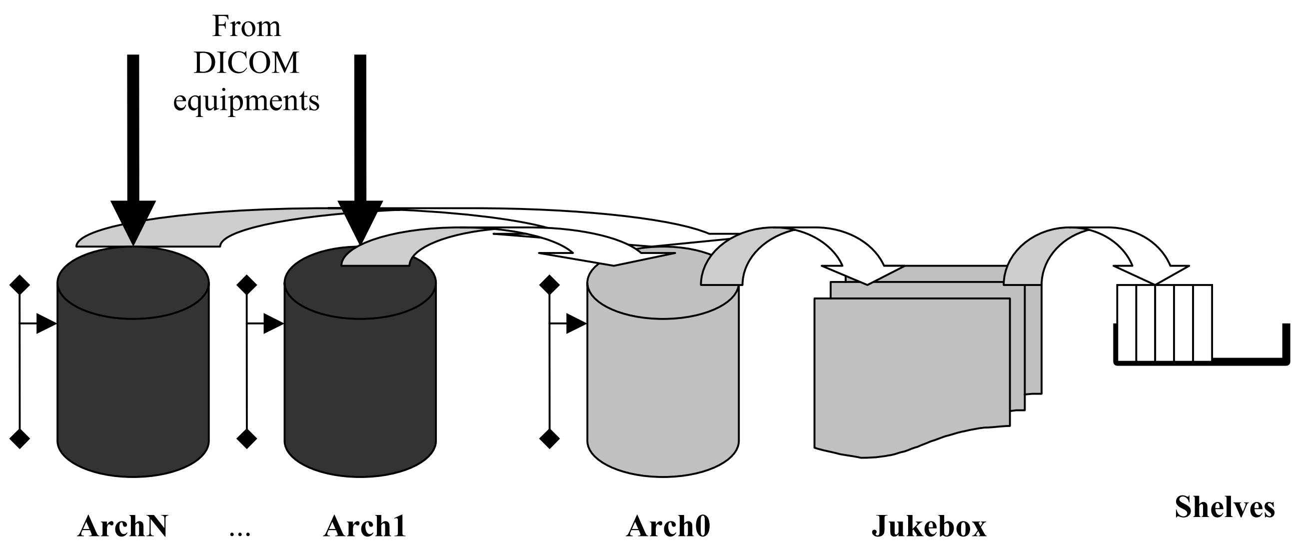

Figure 1. New images are archived in ArchN to Arch1 subsystems, depending on the DICOM server configuration. As soon as the threshold is reached, a block of old exams are moved to the Arch0 storage. If Arch0 also reaches its limits, a block is moved to the jukebox and eventually to the shelf.

System (RIS), to allow construction of robust Electronic Patient Record (EPR) [3].

Since DICOM database does not have fields for nearline and offline storage, additional fields in 'Series' table corresponding to media information are necessary [2]:

- ONLINE_STATUS (boolean): current series is (not) online;
- ONLINE_PATH (string): null or path of the directory that contains the exam;
- NEARLINE_LABEL (string): null or identification of the nearline media (tape in our case);
- OFFLINE_LABEL (string): null or identification of the offline media (tape).

Each time a DICOM equipment sends an exam to the DICOM server, all fields of the database related to DICOM data are updated (figure 3). These fields should keep a consistent track of images along time. An image series can not be nearline and offline at the same time. However, images can be online and nearline or offline, since data might have been retrieved from jukebox or shelf for visualization.

The system manages all insertions and deletions in each media. In order to maximize availability of online data, the images related to most antique patients are copied (full tape load) to a new tape in the jukebox subsystem. Each tape supports up to 70 GB (configurable). Therefore, as soon as the tape is written, disk space can be released as needed until all those images in the tape are deleted from online storage. Then another tape is used and so on. This approach assures optimal usage of online data, maximizing the number of available images for quick access.

In order to assure that online disk is almost full all the time, we are using two watermarks to manage disk space. When the watermark level 1 for disk space is reached, the images of most antique patients are deleted from online media until the disk space level reaches the watermark level 2 and this rule is applied indefinitely.

2.3. Retrieving images

Image access can be done directly querying the database or through DICOM protocol, because all functionalities of DICOM server can be kept. Since we are using the same database manager for patient and image data, an Electronic Patient Record can directly access images and take further actions, such as visualization or processing [4].

The nearline retrieval involves automatic restoring of the DICOM data to the online media, appending a short life tag (2 days). The search and load time from jukebox is around 3 minutes.

For very old exams (offline mode), the procedure for recovery involves human operator and is via scheduling, loading the proper media into the nearline storage and following the same process as for nearline data.

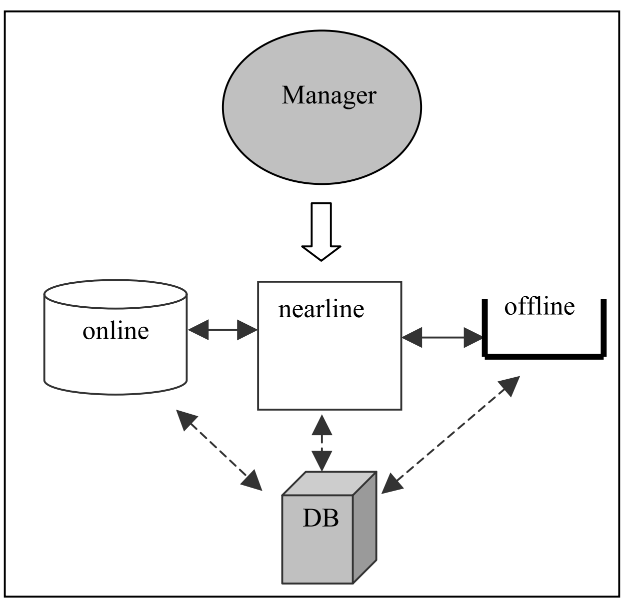

Figure 2. Manager keeps track of all file insertion, move, deletion and media conditions.

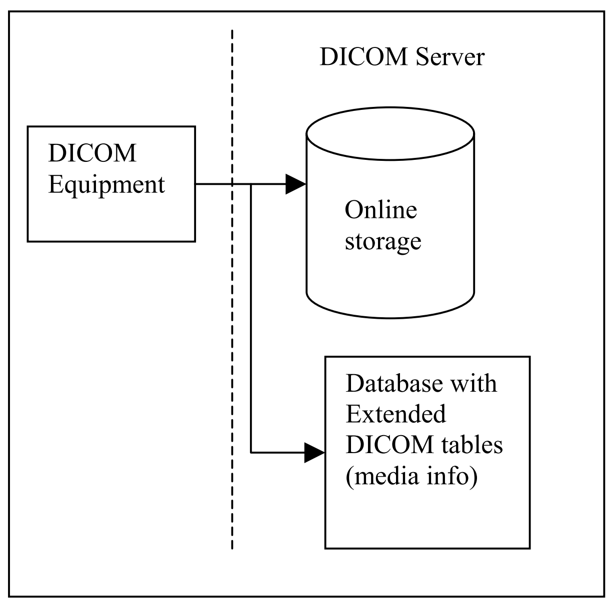

Figure 3. Archiving with media information based on DICOM fields.

3. Results

The described architecture was developed and implanted at the Heart Institute (InCor) of University of Sao Paulo in April, 2000. We used a 200GB RAID5 (Compaq, Ultrix) subsystem as Arch0 and 500GB RAID5 (Compaq, Linux) subsystem as Arch1. As nearline, we used a 3.5 TB DLT jukebox (Compaq) with 52 slots and 4 drives. Currently, more than 2.5TB of DICOM images, have been stored using this architecture.

The system is multi-task and multi-user, allowing simultaneous reading and writing, including the four drives of the nearline mode.

Table 1 shows the distribution of number of exams in August 2002, considering the modalities and the mode of storage. In average, 300 access per day are performed using our Electronic Patient Record.

Table 1. Number of exams per modality and mode of storage of the current system in August 2002.

Modality	online	nearline/ offline
Hemodynamic	3726	12916
Magnetic Resson.	1973	1938
CT	2562	2123
Nuclear Medicine	423	57

4. Conclusions and discussion

This paper proposes a flexible architecture for large PACS storage. The online capacity can be easily incremented as needed and tenths of terabyte of data can be archived in a cost-effective way, balancing online, nearline and offline modes.

An alternative approach can be taken if one wants to use the nearline media also as backup of online storage. In this case, each new exam should be stored simultaneously in the 'ArchN ... Arch1' subsystem and 'Arch0' buffer. If 'Arch0' reaches its limit, a block of exams of the size of nearline media is moved to the jukebox. If any of 'ArchN' subsystem reaches its limit, a block of exams is simply deleted, since they are already in 'Arch0' or in permanent media.

Acknowledgements

This project has been supported by FAPESP grant number 97/14206-5, FINEP 66.0138 and by Fundação E. J. Zerbini.

References

[1] DICOM Committee (2000) Digital Imaging and Communications in Medicine (DICOM), National Electrical Manufacturers Association, version 2000

[2] Furuie SS, Figueiredo JCB, Gutierrez MA, Bertozzo N, Yamaguti M. Archiving and retrieving long-term cineangiographic images in a PACS. Computers in Cardiology 1999; 26:435-8.

[3] Nardon FB, Rebelo MFS, Furuie SS, Moura LA. Modeling the Electronic Patient Record - The Heart Institute of São Paulo Case. Proceedings of Toward an Electronic Patient Record (TEPR'2000) 2000: 150-164.

[4] Tachinardi U, Gutierrez MA, Moura LA, Leão B, Meneguetti C, Melo CP. Integrating PACS and HIS in a Nuclear Cardiology Department. In: Computers in Cardiology 1993. London, UK: IEEE Computer Society Press, 1993: 819-22.

Address for correspondence.

Sérgio S Furuie.
Serviço de Informática – InCor
Av Enéas de Carvalho Aguiar, 44.
05403-000 São Paulo – SP Brazil
Sergio.furuie@incor.usp.br