

# Scalable Data Storage Analysis and Solution for Picture Archiving and Communication Systems (PACS)

Ayhan Akbal and Erhan Akbal

**Abstract**—Information technologies had become means of universal development by rapidly growing in the recent years. And the health sector is the leading one among these sectors. While medical information and laboratory test results –among the data sets generated at health institutions-- occupy a small space as data due to being in text form, digital images occupy a large space in discs due to being pictures. Thus the storage and use of digital images is becoming a significant problem for the institutions. Within the frame of this study, solutions regarding the points to be considered in Data Storage solutions -by using the data sets of Hospital of University of Firat as sample-, how the planning should be made and the optimization of data storage areas had been emphasized.

**Index Terms**—PACS, data storage, DICOM, network.

## I. INTRODUCTION

Information technologies had become means of universal development by rapidly growing in the recent years. And the health sector is the leading one among these sectors. Health sector, which had started to become digital, had generated rapidly increasing data sets along with digitalization. Digital medical images-consisting the basis of diagnosis and treatment in health are the most significant information source of these data sets. Health sector, which had started to become digital, had carried the medical images to digital environment in order to benefit from the advantages of this technology [1]. In health sector, there are many devices and sources generating medical images such as x-ray, ultrasound, MR, tomography, angiography, endoscopy, colonoscopy etc. The use of these images –generated by such sources- in the treatment of patient and the use of them as source in academic studies had gave rise to the obligation of storing them in order to follow-up the patient's illness in future periods. While the medical information and laboratory tests –among the data sets being generated at health institutions- occupy a small space due to being in text form, digital images occupy a large space in discs due to being pictures.

Thus the storage and use of digital images is becoming a significant problem for the institutions. Within the frame of this study, solutions regarding the points to be considered in Data Storage solutions -by using the data sets of Hospital of University of Firat, how the planning should be made had been emphasized.

Manuscript received January 7, 2016, revised September 13, 2016.

Ayhan Akbal is with the Department of Electrical and Electronics Engineering, Firat University Elazig, Turkey (e-mail: ayhanakbal@gmail.com).

Erhan Akbal is with the Department of Digital Forensic Engineering, Firat University Elazig, Turkey (e-mail: erhanakbal@gmail.com).

## II. PICTURE ARCHIVING AND COMMUNICATION SYSTEM

Picture Archiving and Communication Systems (PACS) is the general name assigned to software or Networks used for storage, recall, distribution and presentation of medical images. International Digital Imaging and Communication in Medicine (DICOM) protocol had been developed by European CENT TC251, Japan JIRA, American IEEE, HL7 and ANSI institutions in order to have all the different systems operate as integrated with each other by a common format in the whole world and in order to ensure unity [1], [2].

### A. DICOM Standard

Medical image, information on the relevant medical image, patient's information, hospital's information, motion videos and voice records are able to be placed within the DICOM file [3] This standard had first been constituted by the a committee consisting of American College of Radiology (ACR) and National Electrical Manufacturers Association (NEMA)[2]. The first version of this standard had been defined as ACR-NEMA 300-1985 or ACR-NEMA 1.0, and in 1988 Radiological Society of North America (RSNA) had released the second version called ACR-NEMA 2.0(ACR-NEMA 300-1988) which had been developed by the ACR-NEMA work groups. In 1993, RSNA had released a more functional form of these standards. This modernized standard had been developed as ACR-NEMA DICOM or DICOM 3.0. DICOM had been annually reviewed and updated by the work groups, and it had been named such as DICOM 3.0 2014 as per it year of update. Standard development studies are being carried out by "Medical Imaging & Technology Alliance" department which is within the body of The Association of Electrical and Medical Imaging Equipments Manufacturers located in United States of America [4]. This protocol performs data exchange in between two systems through the TCP/IP protocol. DICOM Standard consists of 18 sections in total. These are:

- PS3.1: Introduction and Overview
- PS3.2: Conformance
- PS3.3: Information Object Definitions
- PS3.4: Service Class Specifications
- PS3.5: Data Structures and Encoding
- PS3.6: Data Dictionary
- PS3.7: Message Exchange
- PS3.8: Network Communication Support for Message Exchange
- PS3.9: Removed
- PS3.10: Media Storage and File Formats for Media Interchange
- PS3.11: Media Storage Application Profiles
- PS3.12: Media Formats and Physical Media for Media

Interchange

- PS3.13: Removed
- PS3.14: Grayscale Standard Display Function
- PS3.15: Security and System Management Profiles
- PS3.16: Content Mapping Resource
- PS3.17: Explanatory Information
- PS3.18: Web Access to DICOM Persistent Objects (WADO)

Among these sections, PS3.2, PS3.3, PS3.4, PS3.5 and PS3.6 determine the network transmission and common characteristics of media sharing. PS3.7, 3.8 and 3.9 determine the basics regarding the transmission of digital image from one point to another. PS3.10, PS3.11 and PS3.12 determine the storage of media and how the file formats will be. While PS3.14 is determining the gray scale image standards, PS3.16 and PS3.17 determine source and content explanations, and PS3.18 DICOM defines the web accesses. [1].

PS3.11 section of DICOM deals with how the DICOM images will be stored and which information is required. Within this frame, details on how the DICOM images are required to be stored in physical storage mediums are being explained in detail in Fig. 1.

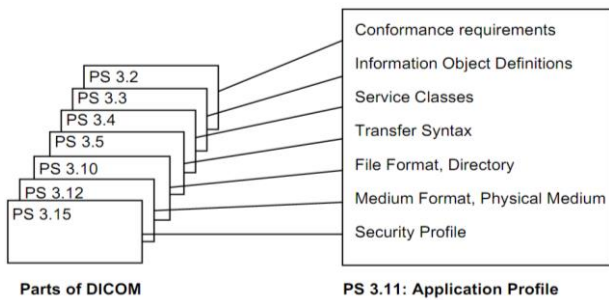


Fig. 1. Sections that DICOM PS3.11 is concerned with.

B. Data and Implementations of DICOM

The generation and storage of DICOM image data is a critical phase especially within the body of health institutions. The constructs where DICOM images are being generated at an institution had been shown DICOM images: Colonoscopy/ X-Ray (X-Ray, CR, DR)/Magnetic Resonance (MR)/ Tomography (BT, CT), Ultrasound (US), Angiography, Endoscopy

The data sources being specified generate images as per technology, and they are required to be stored in order to be communicated to relevant individuals. The user groups within the institution performing operations by using the DICOM data had been provided [5].

DICOM images -after being generated by the sources- are first being kept in local storage spaces until being examined by radiologists, or they are sometimes being kept at main storage spaces until being transmitted. But the occupying of large spaces by image sizes obtained for some sources plays a significant role in the design of physical storage mediums and in the determination of storage type [6].

C. PACS Systems

The images generated by DICOM data sources should be stored for evaluation by radiologists, for determination of surgical operation and treatment as being examined by doctors and for the prospective follow-up of health status of patients. The systems, where DICOM data are being stored

and which enable reuse of them when required, are being called PCS (Picture Archiving and Communication System). In PACS, the sources and types of transmitted images are called modality [7]. These modalities are being seen in Table I.

TABLE I: MODALITIES OF DICOM

Modality	Description
AU	Audio
CD	Colorflow Doppler
CR	Computed radiography
CT	Computed tomography
DSA	Digital Subtraction Angiography
DX	Digital Radiography
ECG	Electro cardiography
EPS	Cardiac Electrophysiology
ES	Endoscopy
GM	General Microscopy
HD	Hemodynamic Waveform
IVUS	Intravascular Ultrasound
LS	Laser surface scan
MG	Mammography
MR	Magnetic Resonance

III. STUDY ON SCALABLE DATA STORAGE MEDIUM

The shooting of storage mediums and technologies required for PACS systems is very important. Within this frame, by performing the analysis of images transmitted to PACS system from radiological images, the amount of storage space and technology to be required in the future had been determined. According to the result of this study, the issues to be considered by other institutions which will install the PACS system had been determined [5], [6].

In this study, the data to be analyzed had been received from the Hospital of University of Firat. Hospital of University of Firat is located at Elazig Turkey. It is an hospital having 1.000 beds, having daily average patient number of 1.700, where about 50 surgeries are held per day, and having employees of over 2.000. The PACS modalities –stored at the Hospital of University of Firat in between 2009-2014- and their numbers had been given in Table II, and the space which the images occupy in the disc had been given in Table III.

TABLE II: MODALITIES BEING STORED ON PACS SYSTEM AND NUMBER OF SHOOTINGS

Yıl	CR	DX	US	MR	CT
2009	1.021	-	99	10.260	340
2010	15.126	29.178	14.119	13.262	15.852
2011	52.439	28.949	14.489	14.379	15.637
2012	51.803	25.558	15.525	16.149	15.386
2013	50.230	34.552	16.412	17.694	20.518
2014	54.938	40.452	18.549	21.614	24.059

When Table I and Table II are examined, it will be observed that there is no direct relation in between number of shootings and the space occupied on disc. The reason of it is that while CR, DX have one or more image frames in partial US shootings, there are more than one frame in MR and CT. Thus, among the radiological images, CT image of a patient

occupies a large space on the disc.

TABLE III: SPACE OCCUPIED ON THE DISC BY THE MODALITIES BEING STORED ON THE PACS SYSTEM (GB)

Yıl	CR	DX	US	MR	CT
2009	12,58	-	0,67	839,00	160,80
2010	238,11	573,12	49,40	1.035,00	10.890,80
2011	803,22	662,37	38,28	1.043,00	11.801,00
2012	853,03	579,22	54,23	1.245,00	11.036,00
2013	804,99	792,57	58,40	1.296,00	13.477,00
2014	909,25	847,26	64,20	2.105,00	14.205,00

TABLE IV: MODALITY OF A PATIENT AND IMAGE SIZES

Mdltıy	Rows	Col.	Pixel Space	Bit	Size of Frame	Number of Frame	Total Size
CR,DR	2437	2557	0,14	16	11,8 MB	1	11,8 MB
BT,CT	679	888	0,54545	16	516 KB	932	470 MB
US	768	1024	-	8	770 KB	3	2,25 MB
MR	256	256	1,7969	16	142 KB	330	45,76 MB

When Table IV is examined, serial number in shooting of each patient is low despite the size of CR images is higher, but when CT image of a patient is examined, the total size of the CT image of a patient quite increases as number of frames in CT is higher despite the file size of one frame of image is small compared to CR. Within this scope, the number of patients of the hospitals, shooting numbers, modalities of shooting, number of frames shot are significant parameters in planning of storage spaces.

A. Determination of the Capacity of Data Storage Space

The hospitals should well analyze the changes in the past while prospectively planning their storage spaces. Within this frame, when the modalities, shooting numbers, data sizes at Hospital of University of Firat in between years 2009-2014 is analyzed on the basis of trimesters during the prospective storage space plan of the Hospital of University of Firat, it is being observed that CR(DR) changes as shown in Fig. 2, and CT changes as shown in Fig. 3.

These changes, annual increase rates and storage space to be required for the next 5 years with an increase rate of 20% -through the planning of systems to be procured- had been provided in Table V.

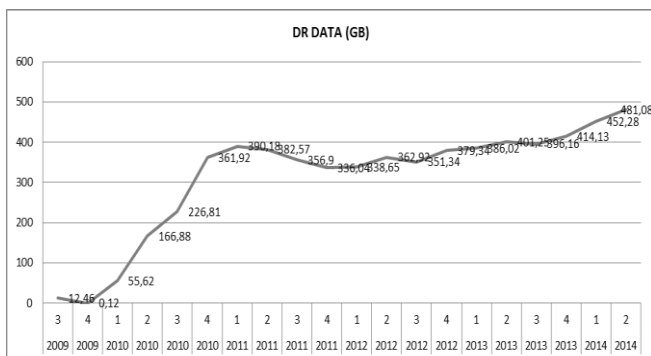


Fig. 2. Distribution in GB of DR image as per years on the basis of trimesters.

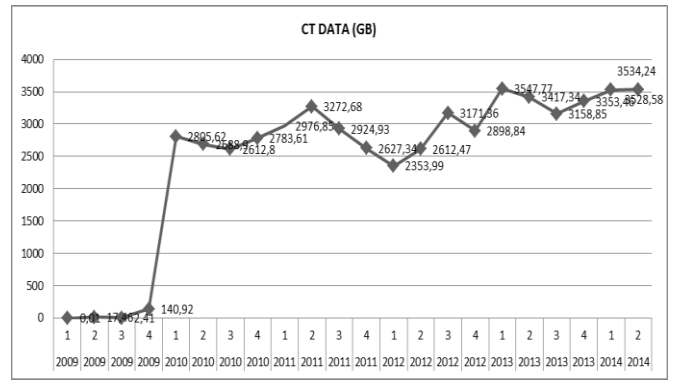


Fig. 3. Distribution in GB of CT image as per years on the basis of trimesters.

TABLE V: DISC REQUIREMENT FOR FIVE YEARS

YEAR	DISC SPACE (TeraByte)
2009	1,19
2010	14,98
2011	16,82
2012	16,18
2013	19,30
2014	57,06
2015	68,47
2016	94,49
2017	130,39
2018	179,94
2019	248,32
2020	342,69

B. Data Storage Technology and Selection of Disc

After determination of the PACS storage space requirement, it is required to determine which technology and disc speeds are required for that storage space. Because discs are significant parameters in the determination of PACS storage space costs. For this purpose, it is required to determine the number of devices generating PACS data within the institution, in other words number of users who will write data on the discs and the number of users who will read such data. Within this frame, the users at Hospital of University of Firat writing data on the disc and their number had been given in Table VI.

TABLE VI: NUMBER OF DEVICES AT THE HOSPITAL OF UNIVERSITY OF FIRAT

Modality	Number of Devices
CR	2
DX	2
US	18
MR	2
CT	2
<b>Total</b>	<b>26</b>

IV. CONCLUSION

Consequently, the selection of data storage space and

technology in a PACS system differentiates in between institutions and in between regions where service is being provided. But it is being observed that the points to be considered are common under all the conditions. Within this frame, in the determination of PACS data storage space, number of devices at the hospital generating PACS data, number of users using PACS data, types of modalities by which shooting is performed, number of frames shot by modalities performing shooting with high frame number especially such as tomography shootings are very important, and they determine the design of infrastructure.

As described in the previous sections, the capacity values required for the storage of PACS data are very high. While the costs are increasing when high speed discs (such as Flash Disc) are preferred, retrospective access to data of 15 days may be a problem with low speed discs. Thus, it is required to perform optimization. If the whole system is constituted by SSD disc, optimum solution will not be provided –as access to data older than 15 days would decrease- and installation costs will be high. Thus, while 15 days' PACS data is being kept on SSD discs, images of older than 15 days can be automatically carried to FATA disks. In this case, the costs can be decreased by keeping the number of SSD discs low, and thus high capacity data storage spaces are obtained with low costs by the FATA discs.

Moreover, the data storage spaces to be installed should be scalable. The increase in time of images transferred to scalable PACS data storage space, addition of new capacity spaces to the system prospectively, easy integration of the added systems to the current system, ability of simultaneous operation of disc technologies such as SSD (FLASH), SAS, FIBER, FATA on the same storage unit, the disc groups being added to the structure not to decrease the total speed of the system, access to discs directly over network and their easy management are required.

## REFERENCES

- [1] Digital Imaging and Communications in Medicine (DICOM) Part 11: Media Storage Application Profiles, PS 3.1. [Online]. Available: [ftp://medical.nema.org/medical/dicom/2007/07\\_11pu.pdf](ftp://medical.nema.org/medical/dicom/2007/07_11pu.pdf)
- [2] P. Mildemberger, M. Eichelberg, and E. Martin, *Introduction to the DICOM Standard*, Euro Radiol, vol. 12, pp. 920-927, 2002.
- [3] J. C. Caicedo, F. A. Gonzalez, E. Triana, and E. Roomero, "Design of a medical image database with content-based retrieval capabilities," *Advances in Image and Video Technology*, vol. 4872, pp. 919-931.
- [4] D. Caramella, "Is PACS research and development still necessary?" *International Congress Series*, vol. 1281, pp. 11-14, 2005.
- [5] J. A. Seibert, L. J. Filipow, and K. P. Andriole. (2000). Practical Digital Imaging and PACS. [Online]. 174(4). Available: <http://www.ajronline.org/doi/full/10.2214/ajr.174.4.17409322000>
- [6] R. Taberner and T. Contesti, "Digital photograph storage systems in clinical dermatology," *Actas Dermosifiliogr*, vol. 101, no. 4, pp. 307-314, 2010.
- [7] H. N. Strickland, *PACS (Picture Archiving and Communication Systems): Filmless Radiology*, Arch Dis Child 2000, vol. 83, pp. 82-86.



**Ayhan Akbal** was born in 1977 in Elazig, Turkey. He is currently working as an assistant professor at the Electrical-Electronics Engineering Department of Firat University. He received his Ph.D. degree in electrical and electronics engineering in 2008, M.S. degree in computer engineering in 2001, from Firat University, Turkey. His research interests include communication, network, wireless, FPGA security, and wireless sensor network.



**Erhan Akbal** is currently working as an assistant professor at the Digital Forensics Engineering Department of Firat University. He received his Ph.D. degree in electrical and electronics engineering in 2012, M.S. degree in computer engineering in 2007, from Firat University, Turkey and a bachelor degree in computer science from Jordan University of Science and Technology, Jordan, in 2003. His research interests include computer network security, wireless sensor network, intrusion detection and digital forensics.

# **Computer Science and Engineering Applications**

