Design of A Picture Archival and Communication System Based on Collected Data

by

Vijay D'Silva

Center for Communications and Signal Processing
Computer Studies
North Carolina State University

August 1986

CCSP-TR-86/17
# TABLE OF CONTENTS

1. INTRODUCTION ............................................................................................................. 1
   1.1 Problem statement .................................................................................................. 1
   1.2 Description of this thesis ....................................................................................... 3

2. THE PACS ENVIRONMENT: A DESCRIPTION ......................................................... 6
   2.1 Organization of the radiology department ......................................................... 6
      2.1.1 Admissions .................................................................................................... 6
      2.1.2 Diagnosis ..................................................................................................... 6
      2.1.3 The Filerooms .............................................................................................. 7
      2.1.4 The Referring services ............................................................................... 8
   2.2 Objectives of the surveys .................................................................................... 8

3. THE SURVEYS .............................................................................................................. 11
   3.1 The Fileroom survey ......................................................................................... 11
      3.1.1 Aims and objectives of the survey ............................................................. 11
      3.1.2 Results from the Fileroom survey .............................................................. 12
   3.2 The Radiological Survey ................................................................................... 19
      3.2.1 Description of the survey ......................................................................... 19
      3.2.2 Results obtained from the Radiological survey ........................................ 21
5. DESCRIPTION OF THE SIMULATION MODEL

5.1 Inputs to the model

5.1.1 Speed of the bus

5.1.2 Block sizes

5.1.3 Compression ratio

5.1.4 Influx rate

5.1.5 DBMS time

5.1.6 Diagnosis think times

5.2 Outputs from the model

5.2.1 Mean retrieval time for an image

5.2.1.1 Confidence interval calculation

5.2.1.2 The batch means method

5.2.2 Delay, mean queue length, and utilizations

5.3 The logic driving the program

5.3.1 Flow chart for the program

5.3.2 Estimation procedures

5.3.3 Programming methodology
6.1 Mean retrieval times under different conditions .................................. 70
6.1.1 Bus speed variations ............................................................................. 72
6.1.2 Compression ratio variation ................................................................. 73
6.1.3 Block size variation .............................................................................. 74
6.1.4 Influx variation ..................................................................................... 76
6.1.5 DBMS time variation ............................................................................. 77
6.1.6 Diagnosis think time variation .............................................................. 78

7. DISCUSSION AND CONCLUSION .......................................................... 87

LIST OF REFERENCES .................................................................................. 92
APPENDIX A. PROGRAM LISTING FOR THE SIMULATION ......................... 95
APPENDIX B. SAMPLE OUTPUT OF THE PROGRAM ................................... 96
## LIST OF TABLES

1. List of Current Filerooms ................................................................. 7
2. Questions on the Fileroom Survey .................................................. 12
3. Questions on the Radiological Survey ........................................... 20
4. Questions on the Admissions Survey ............................................. 29
5. Distribution of the number of procedures per patient .................. 31
6. Description of the Reading Stations ............................................. 42
7. Input Data to the Simulation Model .............................................. 45
8. Diagnosis Think Times used for the Simulation .............................. 46
9. Frequency of Access of Existing Filerooms .................................. 47
10. Average Number of Images per Diagnosis per Station ................ 49
LIST OF FIGURES

1 Current Procedure in a Radiology Department ........................................ 10
2 Variation in the number of requests for folders ..................................... 15
3 Success of recovery of Film Folders from the Fileroom ............................ 16
4 Variation of Location Times for Folders with Time of Day ....................... 17
5 Distribution of Number of Studies per Exam Category ............................. 17
6 Number of Studies vs. Time of Day for Patient Categories ....................... 18
7 Number of Old Images among Images Used for Diagnosis ........................ 25
8 Images per Study, for Storage, for Referring Services ............................ 25
9 Fraction of Images for Storage, to referring Services ............................. 25
10 Usefulness of Lab Studies ........................................................................ 26
11 Usefulness of Clinical Information .......................................................... 26
12 Usefulness of other Radiological Studies ................................................ 27
13 Fraction of Normal Studies per Exam Category ...................................... 27
14 Reprocessing of Films: Probability of Occurrence ................................ 27
15 Number of Admittances on Monday: A plot vs. exam category ............... 32
16 Distribution of Diagnostic Examinations with Time ............................... 32
17 Distribution of Nuclear Medicine Examination with Time ..................... 32
18 Distribution of CT-Body Examinations with Time ............................. 33
19 Distribution of MRI Examinations with Time .................................. 33
20 Distribution of Ultrasound Examinations with Time .......................... 33
21 Distribution of Special Procedure Examinations with Time ................. 34
22 Number of Procedures per Patient per Referring Service .................. 35
23 Diagram of PACS Architecture ..................................................... 51
24 Distribution of Number of Films per Study for Stations 0 to 3 .......... 52
25 Distribution of Number of Films per Study for Stations 4 to 7 .......... 53
26 Distribution of Number of Films per Study for Stations 8 to 11 ........ 54
27 Q-Q plot for the Overall Request Rate to the Fileroom .................... 55
28 The Queueing Model for the System ............................................ 68
29 Flow Chart for the Program ....................................................... 69
30 Mean retrieval times vs. Bus speeds ............................................ 80
31 Device delays vs. Bus speeds ..................................................... 80
32 Utilization of devices vs. Bus speeds .......................................... 80
33 Mean retrieval times vs. Compression ratio .................................. 81
34 Device delays vs. Compression ratio .......................................... 81
35 Utilization of devices vs. Compression ratio .................................. 81
36 Mean retrieval times vs. Block size .......................................... 82
37 Device delays vs. Block size ..................................................... 82
38 Utilization of devices vs. Block size .............................................. 82
39 Mean retrieval times vs. Influx rate .................................................. 83
40 Device delays vs. Influx rate .............................................................. 83
41 Utilization of devices vs. Influx rate .................................................. 83
42 Mean retrieval times vs. Request/influx ratio ...................................... 84
43 Mean retrieval times vs. DBMS times ............................................... 84
44 Device delays vs. DBMS times .......................................................... 84
45 Utilization of devices vs. DBMS times ............................................... 85
46 Mean retrieval times vs. Diagnosis think times ................................. 85
47 Device delays vs. Diagnosis think times ........................................... 85
48 Utilization of devices vs. Diagnosis think times ................................. 86
CHAPTER 1

INTRODUCTION

1.1. Problem statement

There exists a need in the field of radiology today to develop a comprehensive system for digitally storing and retrieving large amounts of data [2]. More and more digital imaging machines are now being used in hospitals and mass storage media are looked on as more than just peripheral devices attached to these instruments. What is needed is a network of systems that would facilitate the communication of images among medical personnel. It would also be necessary to develop computerized patient records that organize these images so as to present the radiologist with the same options to which he is accustomed, i.e. a choice of images and patient information which can be quickly retrieved when requested. This would expedite diagnoses and procedures and also reduce mismanagement of films and film folders which is inherently part of manually handled information. High speed communication networks, mass storage devices and high resolution display stations must be designed in the course of developing such a Picture Archival and Communications System (PACS). New techniques in digital signal processing, data base management, and high speed computer networking are necessary [3][4][5].
A layered approach has been suggested to design the PACS [6]. This method has been implemented in most computer networks and Cox et al. are of the opinion that the same approach must be applied to picture networks [2][6]. The phased approach wherein the PACS is developed gradually in stages so as to interface with existing equipment seems to be popular since it does not involve any dramatic overhauling of the whole radiology department. This method has been proposed and implemented in part by Stockbridge [7] and Duerinckx [5].

This thesis concerns itself with the computer communications and architectural aspects of such a system as well as the requirements imposed on the data base management system. The PACS must be designed based on information about procedure volumes and the magnitude of data transfer rates at any given time, particularly at peak hours. Since significant fluctuations can be expected with respect to the demands during a twenty four hour period, the model must be able to predict the performance of the system under peak traffic conditions as well as quiescent conditions.

The impact of varying the characteristics of the individual components of the proposed system has been studied and the bottlenecks of the system under different conditions have been identified using a computer simulation model of the PACS.
1.2. Description of this thesis

This thesis consists of two parts. The first deals with the traffic characterization of procedures at three dissimilar radiographic facilities: an academic medical center, a large community hospital, and a small community hospital. The second part proposes an architecture for a PACS and a model of this system is given. The input parameters of this model are derived from the results in part one. The output parameters are compared for different configurations and varying equipment characteristics e.g. bus speeds, database management (DBMS) delays, protocols, etc.

An academic medical center, the North Carolina Baptist Hospital is a 701 bed institution associated with the Wake Forest Medical Center located in Winston-Salem, North Carolina. In addition to the inpatient population there are about 170,000 visits to the outpatient clinics each year, i.e. about 470 per day [1].

Forsyth Memorial Hospital is a 819 bed institution serving Winston-Salem and surrounding areas. There are no outpatient clinics in this hospital but a very busy emergency department with about 52,000 visits per year, i.e. an average of 140 per day.

The Northern Hospital of Surry County is a 120 bed hospital at Mt. Airy, North Carolina. It also does not have any outpatient clinics. The radiology
department of this hospital handles about 34,000 examination per year i.e. about 93 per day.

These three hospitals share the same radiological facility. Three surveys were taken at these hospitals which collected information about different aspects of current practices in the radiology department. The first survey was a log of all requests of films from the fileroom or storage area. The second survey dealt with the recovery of films and the information used in a diagnosis. The third survey was concerned with the influx of patients into the radiology department and the traffic in and out of the examination rooms.

In the analysis of all this information the data was subdivided by different criteria into groups or examination types. Graphs were obtained showing the idiosyncrasies of each examination type. Much of the data was concerned with determining how well the current (manual) filing system works. Since this data was not germane to this project it is not given here [22]. Furthermore there are many entries on the surveys which were too detailed to be incorporated into the current model. It was assumed that these details would not affect the performance evaluation significantly and are far too complicated to be incorporated in a prototype anyway.

The second part of this thesis is concerned mainly with the simulation model of the network. For simplicity the network has been represented by its basic components i.e. the bus, reading stations, acquisition nodes, a DBMS,
magnetic discs and their controllers, and optical jukeboxes. Some of the under-
lying design features of this hypothetical PACS were to a great extent
influenced by the system developed at AT&T. That is,

a) using a centralized DBMS which is accessed by all reading stations and
acquisition nodes.

b) assuming that images are retrieved in a single thread fashion.

c) the acquisition nodes themselves might be connected to several input sta-
tions.

Using this simulation model, the delays for film and folder retrieval have
been obtained for different conditions. Optimal block sizes have been found for
the storage of data and its retrieval. The effect of bus speed and other factors,
such as DBMS times, block size, request and input rates, on the retrieval time
has been studied.
CHAPTER 2

THE PACS ENVIRONMENT: A DESCRIPTION

2.1. Organization of the radiology department

2.1.1. Admissions

In a radiology department each patient who needs to have an examination done is associated with an order form. This is true for inpatients as well as outpatient. This order form which bears the relevant request for a radiological examination, is submitted at the admissions desk. A computer currently stores the location of each patient's existing film folder and the admissions desk immediately initiates a search for the relevant folder at the fileroom. Simultaneously an appointment is set up for the examination. Upon completion of the examination, the films created are sent to the reading room. Figure 1. gives a graphic description of the above procedure.

2.1.2. Diagnosis

The film folder containing the patient's past films is sent directly to the reading room. The radiologist then makes his diagnosis based on the new films taken as well as the old films in the folder. In addition, the radiologist may
use clinical or lab studies in the course of making his diagnosis. At the end of
the diagnosis the new films containing important information are chosen for
long term storage. These are inserted into the folder which is then returned to
the fileroom. Sometimes the case is a 'STAT' case i.e. the referring physician
has to be notified immediately.

2.1.3. The Filerooms

The fileroom is somewhat hierarchically organized. The upper level con-
tains recent information and hence recovering film folders is faster. The
filerooms are described in Table 1 below.

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>house folders kept for 8 days</td>
</tr>
<tr>
<td></td>
<td>waiting reports kept for 3 days</td>
</tr>
<tr>
<td>D2</td>
<td>1985 &amp; 1986 folders kept (two years)</td>
</tr>
<tr>
<td>D3</td>
<td>1984 folders stored</td>
</tr>
<tr>
<td>D4</td>
<td>1979 - 1983 folders stored</td>
</tr>
<tr>
<td>Warehouse</td>
<td>1978 folders stored</td>
</tr>
</tbody>
</table>

Table 1: List of Current Filerooms

The lower levels are organized by the creation time of the films they hold.
This presumably simplifies the management of these films. Every summer the
films are manually transported to the warehouse. Films to be moved must be
keypunched into the computer to record their transfer.
2.1.4. The Referring services

The referring services are those that a patient initially goes to for treatment. The referring physician may decide that an ultrasound study needs to be done, for example, and sends the patient to the radiology department. In a fully implemented PACS network the referring physician would be able to access the resultant images directly on his reading station. Thus there exists a traffic of images to the referring services which ought to be taken into account in designing a PACS. However, a prototype would deal with relatively few reading stations and the several terminals necessary to connect the referring physicians is something that must be dealt with only after a basic system has been implemented.

2.2. Objectives of the surveys

The objectives of the surveys were to analyze statistical information on film traffic to be used as input for the development of the communication network and for the marketing aspects of the PACS. The three hospitals mentioned earlier, Baptist Hospital, Forsyth Memorial Hospital, and Mt. Airy Community Hospital, were to be analyzed to provide the data for this study.

The objectives were as follows:
1. Establish the total number of patients as well the number of patients divided into exam category and referring service admitted to the hospital each day.

2. Determine the arrival rate for examinations per hour as well as the average waiting time for an examination.

3. Ascertain the arrival pattern of requests to the fileroom and the traffic of films between filerooms and reading rooms.

4. Determine the number of unavailable films requested per hour and the number of films that were already in use when requested.

5. Determine the number of requests for new folders, new radiographs in old folders, and old radiographs made within a twenty four hour period.

6. Find patterns in diagnostic practices i.e. the percentage of old films used in a diagnosis, the fraction of films chosen for storage, the images sent to the referring services, the clinical and lab studies used for the diagnosis.
Figure 1: Current Procedure in a Radiology Department
CHAPTER 3

THE SURVEYS

The three surveys were conducted from May 1985 to July 1985. A description of each survey is included and the results drawn from each. All the results submitted to Bowman Gray are not given here but only those that pertain to this project.

3.1. The Fileroom survey

3.1.1. Aims and objectives of the survey

The purpose of this survey was to determine the arrival rate and arrival pattern to the fileroom. Whenever an exam or diagnosis is done a request occurs at the fileroom to recover the patient’s folder. This traffic is largely independent of the implementation of the storage and retrieval system and hence can be used in designing the PACS.

The questions on the survey can broadly be classified as addressing two problems: What are the average and peak request rates to the fileroom and when does the peak occur?; how well does the current system work i.e. what fraction of requests are successful and what percentage of films are lost or misplaced due to various reasons? The peak rates will determine the performance
of the network under heavy traffic conditions. The survey entries are included in Table 2.

<table>
<thead>
<tr>
<th>Fileroom Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Subsection/unit#: A 5 digit code representing the patient's ID preceded by a single digit code for the examination type.</td>
</tr>
<tr>
<td>2. Exam Category: A single or double character code denoting the examination performed.</td>
</tr>
<tr>
<td>3. Film requested: This concerns the patient's hospitalization status i.e. new patient, returning patient, or old patient.</td>
</tr>
<tr>
<td>4. Film Status: Whether the folder in question was in use, unaccountable, or found.</td>
</tr>
<tr>
<td>5. Time request received: Self explanatory.</td>
</tr>
<tr>
<td>6. Total Time to find Film: Self explanatory.</td>
</tr>
<tr>
<td>7. Day: Day of the week.</td>
</tr>
<tr>
<td>8. Hospital: Baptist, Forsyth, or Mt. Airy.</td>
</tr>
</tbody>
</table>

Table 2: Questions on the Fileroom Survey

3.1.2. Results from the Fileroom survey

There were 5658 data collected for this survey. Of these 4215 were from Baptist, 1221 from Forsyth Hospital, and 222 from Mt. Airy. The survey was taken over a two week period.

When we look at the plot of requests vs. time of day (Figure 2a) we are presented with something similar to a telephone traffic load profile. The busy hour was found out to be between 7 am and 8 am. It is interesting to note a
peak at 7 pm. This is an anomaly representing the number of film folders being pulled out for the next day's clinics. It can be seen that Forsyth (Figure 2b) and Mt. Airy (Figure 2c) do not have this peak since they do not maintain clinics. These give an estimate of the prescheduled examinations. This number is significant in reducing the peak traffic by requesting in advance the folders for scheduled examinations.

The graph of the variation of the average time to locate a folder in a twenty four hour period (Figure 4) shows clearly that the location time is maximum during the peak hours. One percent of the folders had a location time of greater than 50 minutes and 5% had a location time of greater than 20 minutes. These times represent only the time between the arrival of the request and the presentation of the folder at the window of the fileroom. They do not include the time taken to deliver the folder to its destination.

On examining the plot of particular examinations types vs. time of day (Figure 5), one notices that 71.5% of these are of plain films. Since plain films are still in analog form and require at least a 4 megabytes per film, it is obvious that one of the major hurdles of designing the system is the incorporation of plain films in the PACS both from a communications as well as a mass storage point of view.

Figure 6 shows the overall traffic load as a function of time of day and patient status. One sees that most of the requests during the peak hours are
attributed to old or returning patients. These patients already have films in storage which must be recovered for diagnosis. Figure 3 indicates how well the existing system works. It is a plot of the number of requests for film folders satisfied and those that were not due to the fact that they were in circulation or were misplaced. It is important to note that between 2% and 18% of the film folders were unaccountable and many (upto 23%) were already in use by someone else when requested. The latter was prevalent in the morning.
Figure 2: Variation in the number of requests for folders
Figure 3: Success of recovery of Film Folders from the Fileroom
Figure 4: Variation of Location Times for Folders with Time of Day

Figure 5: Distribution of Number of Studies per Exam Category
Figure 6: Number of Studies vs. Time of Day for Patient Categories
3.2. The Radiological Survey

3.2.1. Description of the survey

This survey was intended to find out the diagnostic practices of radiologists and perhaps detect some pattern in their requests for information. In particular, we would like to know what the number of films (that the radiologist asks for) is likely to be as well as the clinical and lab information that he might need. We would also like to know what fraction of the total films requested are new (and hence easily recoverable) and how many of them are old and have to be fetched from the archives. Also of interest is the traffic from the referring services.

The survey is not temporal in the sense of finding the times at which diagnoses occur (that information was contained in the fileroom survey) but rather is a collection of samples which hopefully represents all the different situations that would be encountered in the department. For this cross section 2208 samples were taken. Table 3 gives a listing of the questions on this survey.
<table>
<thead>
<tr>
<th></th>
<th>The Radiological Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Doctor:</td>
</tr>
<tr>
<td>2</td>
<td>Patient:</td>
</tr>
<tr>
<td>3</td>
<td>Subsection:</td>
</tr>
<tr>
<td>4</td>
<td>Referring Service:</td>
</tr>
<tr>
<td>5</td>
<td>Exam:</td>
</tr>
<tr>
<td>6</td>
<td># new images:</td>
</tr>
<tr>
<td>7</td>
<td># to make diagnosis:</td>
</tr>
<tr>
<td>8</td>
<td># old images:</td>
</tr>
<tr>
<td>9</td>
<td># for storage:</td>
</tr>
<tr>
<td>10</td>
<td>Other rad. studies:</td>
</tr>
<tr>
<td>11</td>
<td>Clinical Info.:</td>
</tr>
<tr>
<td>12</td>
<td>Lab Studies:</td>
</tr>
<tr>
<td>13</td>
<td>Properly exposed films for enhancement:</td>
</tr>
<tr>
<td>14</td>
<td>Improperly exposed films for enhancement:</td>
</tr>
<tr>
<td>15</td>
<td>Is this study normal?:</td>
</tr>
<tr>
<td>16</td>
<td>STAT Report:</td>
</tr>
<tr>
<td>17</td>
<td>Did the radiologist consult with other radiologists?:</td>
</tr>
</tbody>
</table>

Table 3: Questions on the Radiological Survey
3.2.2. Results obtained from the Radiological survey

The radiological survey was the most fruitful as far as recovering usable information was concerned. The diagnoses of the radiologist were grouped under certain criteria e.g. exam category, referring service, and this uncovered several interesting facts.

The number of old films used in a diagnosis is an important parameter in simulating the PACS. The recovery of these films from the archives takes the longest with the current system. Obviously this will be also be true for a digital PACS too. There are several more old films than new and these will have to be stored in slower memory devices. Figure 7 shows that 39\% of all images required for a 'plain films' diagnosis are old films. This represents a significant fraction of the total. At the other end of the spectrum only 2\% of all contrast images requested are old. This fact is however of very little consequence since there are comparatively very few contrast studies being done. On an average around 30\% of all images recovered are old films.

We now move on to the films per study that we are dealing with and what happens to them. Referring to Figure 8 we see that the number of new films per study greatly varies depending on the examination type. Computed tomography generates the maximum number of films per study. However CT accounts for only 12.8\% of the number of examinations conducted and this is not such a major problem. Also, as will be shown in later chapters, CT does not require
high resolution and this simplifies the storage and communications problem. While only six diagnostic radiology films are generated per exam, this exam will put the heaviest load on the PACS. This is because of two reasons: the first is that plain films account for 71.5% of all examinations; the second is that plain films require the highest resolution.

From the same graph (Figure 8) we also see the numbers of films that are chosen for long term storage. This figure is important when estimating the amount of memory needed for the archive. It also gives the traffic of films that return to the database after the diagnosis is made and the radiologist has annotated them. Of course, there is no necessity to physically restore the film to the database - only the annotation need be 'restored'.

On certain occasions, the referring physician is to be notified immediately on completion of the diagnosis. These are labeled STAT reports. From Figure 9 we see that this traffic to the referring doctors is between 2% and 23% of all new images.

Coming to the issue of other information necessary to make a diagnosis we look at Figures 10, 11, & 12. The information that might be of use are lab studies, clinical information, and other radiological studies. Of these the first two are textual and do not contribute to the load on the system. However, the last is pictorial data and the data contained within is of the same order of magnitude as the regular flow of data in the network.
Looking at Figures 10 & 11 we see that though the necessity of lab studies and clinical information is not an architectural problem as such due to the relatively small amounts of data involved, it does prove that the PACS system needs to be comprehensive in that it should be possible to retrieve all kinds of textual data for the radiologist.

Figure 12 shows the usage of other radiological studies. Prominent are plain films which are referred to in almost every type of examination. This is of use when one considers a distributed PACS in which images of different examination types are stored at different locations.

Figure 13 is of interest for academic reasons. It shows the percentage of studies that are normal i.e. there was nothing wrong with the patient. There has been some discussion on the advisability of storing the films of healthy patients. If this were not necessary then the storage requirements are reduced by a factor of about three.

Figure 14 is of no concern to this project but is interesting anyway. The survey questioned the radiologist as to how many films he would like to reprocess digitally if it were possible. We see that this figure is between 0.1% and 13% depending on the examination type.

Thus we see that the radiological survey has much to reveal about the way radiologists do business. The figures included here are only a small part of those generated in reply to a request from Bowman Gray. Many of these
figures have been used in designing the system discussed later in this thesis report.
Figure 7: Number of Old Images among Images Used for Diagnosis

Figure 8: Images per Study, for Storage, for Referring Services

Figure 9: Fraction of Images for Storage, to referring Services
Figure 10: Usefulness of Lab Studies

Figure 11: Usefulness of Clinical Information
Figure 12: Usefulness of other Radiological Studies

Figure 13: Fraction of Normal Studies per Exam Category

Figure 14: Reprocessing of Films: Probability of Occurrence
3.3. The Admissions survey

3.3.1. Description of the survey

The purpose of the admissions survey was to gather data about the admissions process and the generation of films. It would be very useful to know the arrival rate of new films to be stored and the peak hours of this process. It is also necessary to know, in designing the PACS, how many of such examination rooms can be expected. We would also like to find out what traffic can be expected from the referring services. In addition, we would also like to know what the number of examinations being performed of each type is.

This survey was taken over a period of two weeks and a record was made of every patient that entered and the examination that was conducted. The admissions survey had 6611 samples collected over that two week period. Table 4 lists the questions in the admissions survey and a description of each.

<table>
<thead>
<tr>
<th>Question</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Patient unit number</td>
<td>The subsection/patient ID number</td>
</tr>
<tr>
<td>2. Examination</td>
<td>The examination conducted</td>
</tr>
<tr>
<td>3. Time examination requested</td>
<td></td>
</tr>
<tr>
<td>4. Time examination started</td>
<td></td>
</tr>
<tr>
<td>5. Shift examination started</td>
<td>in case the previous two were not recorded</td>
</tr>
<tr>
<td>6. Day</td>
<td></td>
</tr>
<tr>
<td>7. Hospital</td>
<td>Baptist, Forsyth, or Mt. Airy</td>
</tr>
</tbody>
</table>

Table 4: Questions on the Admissions Survey
3.3.2. Results obtained from the Admissions survey

The admissions survey provided data on the times that films were generated as well as the waiting times for a patient requiring an radiological examination.

Monday is the most active day for admittances. Figure 15 shows that bulk of the films being generated are from diagnostics. These films can take upto 4 megabytes each (without compression) and represent a heavy load on the system. For example, there were 740 diagnostic examinations done on Monday, roughly 85% of the total number of examinations on that day.

Looking at the distribution of diagnostic examinations throughout the day (Figure 16) we see that the peak is between 8 am and 9 am (520 studies). This continues through the morning until it slows down in the afternoon. This unfortunately coincides with the peak for films being requested for diagnosis.

Looking at the admissions pattern for the other examination types we see that the peak arrival rate for nuclear medicine images (Figure 17) is between 9 am and 10 am (53 studies). The peak arrival rate for CT_body (Figure 18) was also between 9 am and 10 am (35 examinations).

The peak for MRI (Figure 19) was between 8 am and 9 am (10 studies). The peak for ultrasound (Figure 20) was between 2 pm and 3 pm (30 studies). There were two peaks for special procedures (Figure 21) - between 8 am and...
Thus we see that the peak traffic can be up to 4 times the average traffic during working hours. There is much emphasis in PACS on fast retrieval times especially in emergency cases and to design the network we simply considered the peak hours and ignored the average traffic which, due to the lack of activity at night and on weekends, is deceivingly low.

Also of interest is the traffic from the referring services. In a fully implemented PACS, the referring services would be connected to the database and it would be necessary to know the traffic that can be expected from each. The traffic would be two-way. The referring service would send patient information and clinical and lab studies to the radiologist to aid in the diagnosis. The radiologist would send images and results from the diagnosis to the referring physician when requested. This turnaround time represents the total efficiency of the radiology department, not only of the PACS but also of the examination times and the diagnosis times.

Figure 22 gives the average number of procedures per patient for each referring service. The number of procedures per patient can be approximated as an exponential distribution as is shown in the table below.
<table>
<thead>
<tr>
<th>Visits/patient</th>
<th>Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2332</td>
</tr>
<tr>
<td>2</td>
<td>366</td>
</tr>
<tr>
<td>3</td>
<td>121</td>
</tr>
<tr>
<td>4</td>
<td>53</td>
</tr>
<tr>
<td>5</td>
<td>27</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 5: Distribution of the number of procedures per patient
Figure 15: Number of Admittances on Monday: A plot vs. exam category

Figure 16: Distribution of Diagnostic Examinations with Time

Figure 17: Distribution of Nuclear Medicine Examination with Time
Figure 18: Distribution of CT-Body Examinations with Time

Figure 19: Distribution of MRI Examinations with Time

Figure 20: Distribution of Ultrasound Examinations with Time
Figure 21: Distribution of Special Procedure Examinations with Time
Figure 22: Number of Procedures per Patient per Referring Service
CHAPTER 4

DESCRIPTION OF THE SYSTEM UNDER STUDY

4.1. Hardware Description

A centralized architecture was adopted for this study (Figure 23). The alternative was to have separate subnetworks for each examination category. However, the traffic between exam categories is considerable i.e. Figure 12 shows that a radiologist may require several different types of films to make a diagnosis. One is thus led to question the advisability of a decentralized system. Hence in this configuration there is a single DBMS and storage system which is accessible by all reading stations and acquisition nodes as is implemented in the PACS developed at AT&T. Following is a discussion of the components of the system and the requirements imposed on each.

4.1.1. Storage Requirements and Devices

4.1.1.1. Requirements of the Mass Storage System

Storage of the large amounts of data generated each day may prove to be the Achilles heel of the PACS system if the capacity of the system is not estimated correctly.
The first step was to find out how much data was being generated each day. To do this we went back to the surveys to gauge how much memory would be required. Studies at the Mallinckrodt Institute of Radiology indicated that 254.81 Megabytes and 502.76 Megabytes of data would be generated per day for a radiology department serving a 614 bed hospital [4]. The figure we got from the Bowman Gray surveys was considerably higher due to the assumption that plain films would eventually be incorporated in the design (plain films would preferably be digitized to a 2048 x 2048 x 12 bit image according to doctors at Bowman Gray). Thus at Bowman Gray 17 Gigabytes are generated each day. However, these images can frequently be digitally compressed. The compression ratios can be as high as 1:20 but a more reasonable value would be 1:10 in the average case. Thus this figure is reduced to 1742 Megabytes per day. If data has to be stored over a seven year period then this results in an overall storage volume of 4454 Gigabytes.

In the current filing system at Bowman Gray, there are five different filerooms to store data (Chapter 2). This organization is not hierarchical in the strictest sense of the word since one of the reasons for its configuration is to ease administration. We thus decided to reduce the number of levels to two for simplicity. The number of levels of recovery are determined most by capacity of storage at each level and the speed of the device. The storage options available are solid-state RAM, fixed magnetic disks, optical disks, optical disks in
jukeboxes, and magazines of optical disks [3]. In this study we simulate two types of storage devices: magnetic disks and optical disk jukeboxes. The magnetic disks have an access time of around 16 milliseconds per record and a storage capacity of the order of megabytes. The optical disc jukeboxes have an access time around 5 seconds (for considerably more information) and a storage capacity of the order of hundreds of gigabytes [8].

4.1.1.2. Description of Optical Disks and Jukeboxes

The optical jukebox used for this study is the one developed by the Radio Corporation of America [9][10]. This device is not commercially available but a prototype has been developed and is functional. An optical disk represents an important advance in storage technology. Storage of previously unmanageable sizes of data is made possible due to high density recording and high transfer rates. Also of significance is the low cost per megabyte. An optical disk can store up to 12.5 Gigabytes of data on a 12 inch disk with very low error rates. The RCA design combines disks into a "jukebox" which has the capability of randomly accessing any one of up to 128 such disks. The total storage capacity is thus around 1600 Gigabytes. Three optical jukeboxes would supply a storage capacity of 4800 Gigabytes which is more than the 4454 Gigabytes necessary.

The jukebox is made up of the following components: an optics platform, a laser, one or more turntables, storage for the disk cartridges, and a
translation mechanism. For a disk to be accessed it must first be loaded onto the turntable which is then spun up to speed before it can be read from or written to. From [9] we obtain the following analysis of the optical jukebox.

\( T_{\text{disk}} \): Before a disk is read it must be loaded onto the turntable. This step takes 4.07 seconds to load the disk and accelerate it.

\( T_{\text{track}} \): The time taken for the head to move to a particular track = 230 mS.

\( T_{\text{fer}} \): Time taken to read a track once the head has been positioned = 34.7 mS.

\( T_{\text{galvo}} \): Time taken for the head to tilt from one track to another = 3 mS.

\( T_{\text{shift}} \): Time taken for the head to shift once the range of the galvo is exceeded = 17.4 mS.

The range of the galvo is ten tracks. A very fast movement will enable the head to read the next track. However when more than ten tracks must be read then a translation shift of the head must be done and this takes longer. Thus the average time to read a track and move to the next is

\[
T_{\text{read}} = T_{\text{fer}} + T_{\text{galvo}} \times \frac{10}{11} + T_{\text{shift}} \times \frac{1}{11}
\]

\[
T_{\text{read}} = 34.7 + 3 \times \frac{10}{11} + 17.4 \times \frac{1}{11} = 39\, \text{mS}
\]

If the data is distributed at random among the disks then

\[
p(\text{data is on current disk}) = \frac{1}{128}
\]

\[
p(\text{data is not on current disk}) = \frac{127}{128}
\]
We also assume that if the disk is being loaded then the head is already positioned and $T_{\text{track}} = 0$.

$$T_{\text{opt}} = \frac{1}{128} * T_{\text{track}} + \frac{127}{128} * T_{\text{disk}} + T_{\text{read}} \frac{(\text{size of data})}{.228 \text{ megabytes per track}}$$

$$T_{\text{opt}} = 4.04 + 0.171 * (\text{size of data})$$

where the size of the data being retrieved is in megabytes and the time is in seconds.

4.1.1.3. The Magnetic Disk Subsystem

A magnetic disk subsystem is included for the higher level short term storage. In the current system short term storage is required for eleven days after which the films are transferred to an archive. Thus we can calculate the amount of storage required i.e. 1742 Megabytes per day x 11 days = 19162 Megabytes. Assuming each magnetic disk stores 600 Megabytes we will need 32 disks for short term storage. In this configuration several disks are controlled by a channel for data transfer [11][12]. The controller initiates a seek at the disk and then is free to do other tasks while the seek is in progress. When the seek is done the disk is blocked until the channel is free to transfer the data from or to that disk. Each track on the magnetic disk can contain 48 kilobytes.

For maximum efficiency the record size should hence not exceed 24 kilobytes (the address field for a record allows for 15 bit encoding which defines the maximum block size as 32 kilobytes). The seek time is assumed to be 16 ms and the
scan time, 7 ms. The transfer rate is taken as 17.4 Megabits/second. Since a request for a megabyte size file results in the retrieval of several blocks, one method of minimizing the retrieval time is to distribute these blocks on all the available disks so as to concurrently execute seeks on each one [11]. The arrival process to the I/O configuration can hence be characterized as a bulk arrival process. It would thus be reasonable to assume that increasing the number of channels available would greatly enhance the performance of the system. In this architecture we assume four channels, i.e. eight disks per channel.

4.1.2. The Communications Link

Several alternatives present themselves. One possibility would be a point-to-point link from each reading station and input modality to the database [7]. Another alternative is to have a bus architecture the protocol of which can be specified at the time of implementation. A high speed fiber optic token ring protocol presents itself. Preliminary studies showed that the bottleneck is not likely to be the communications link but the database. Simulation results have confirmed these findings. A bus architecture is simulated here with two classes of customer. A high priority 1 kilobyte request for an image and a low priority record size data block for transfer to a reading station from the database or from the input modality to the database. The method of scheduling is non-preemptive. Different bus speeds and their effect on the end-to-end delay have
been considered.

4.1.3. The Reading stations

It was outside the scope of this thesis to include detailed specifications of the reading stations. It is assumed that a reading station has enough local storage to handle one complete study. The recovery of images at a reading station are considered independent of each other and the buffer at that reading station, infinite, as far the simulation is concerned. Single threading is used to recover an image from a reading station i.e. the images of a study are requested in sequence and each image can only be requested after the previous image has arrived at the reading station [7]. The reading stations were defined by the radiologists at Bowman Gray and are given in Table 6 below.

<table>
<thead>
<tr>
<th>Station number</th>
<th>Primary usage</th>
<th>Size of Film (bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Nuclear Medicine</td>
<td>256x256x8</td>
</tr>
<tr>
<td>2</td>
<td>Ultrasound</td>
<td>256x256x8</td>
</tr>
<tr>
<td>3</td>
<td>Orthopedics</td>
<td>2000x2000x12</td>
</tr>
<tr>
<td>4</td>
<td>Pediatrics</td>
<td>2000x2000x12</td>
</tr>
<tr>
<td>5</td>
<td>CT</td>
<td>512x512x12</td>
</tr>
<tr>
<td>6</td>
<td>MRI</td>
<td>512x512x12</td>
</tr>
<tr>
<td>7</td>
<td>Emergency room</td>
<td>2000x2000x12</td>
</tr>
<tr>
<td>8</td>
<td>Special Procedures</td>
<td>2000x2000x12</td>
</tr>
<tr>
<td>9</td>
<td>GI, GU</td>
<td>2000x2000x12</td>
</tr>
<tr>
<td>10</td>
<td>Orthopedics</td>
<td>2000x2000x12</td>
</tr>
<tr>
<td>11</td>
<td>Orthopedics</td>
<td>2000x2000x12</td>
</tr>
<tr>
<td>12</td>
<td>CT</td>
<td>512x512x12</td>
</tr>
</tbody>
</table>

Table 6: Description of the Reading Stations
4.1.4. Database Management System

In addition to handling vast number of images that need to be stored, the DBMS will have to be able to locate several records for each image. Many mass storage systems [13][14][15] contain a minicomputer and a high speed data buffering subsystem to handle the location and flow of large amounts of data. A separate hard disk is commonly used just to store the directory of entries. However, simulating each of these devices would not only lengthen the time of simulation but also commit oneself to a certain DBMS implementation. An arbitrary value of 100 mS is used as a DBMS delay in this simulation [9].

4.1.5. The Acquisition Nodes

The acquisition nodes digitize the images from the examination rooms and send them to the database [16]. The rate of input of images may depend on the arrival rate of patients to the department and on the service time of the frame grabber. In this simulation, the acquisition nodes are treated exactly like reading stations. The delays experienced by the data to be stored are approximately the same as those which must be retrieved. There are three acquisition nodes in the system and the system response time can be found for different input rates of new images.
4.2. Applications Software

One of the primary requirements of the software is that the radiologist does not have to dramatically change his procedures and acquire new skills to be able to use the PACS [17]. That is to say the communications and the database retrieval involved in recovering images or studies must be totally transparent to the user.

The diagnosis will probably begin thus. The radiologist is presented with a menu of images which he can access as well as textual information pertaining to the patient. He then selects those that he wants and sits back for hopefully a short period of time before the images arrive. When the study arrives he has the option of flipping through the images before he is able to study particular images in detail.

Each reading station has its own storage capable of storing several images and simultaneously display them on a graphics terminal [16][17].

One other aspect of the diagnosis is that the radiologist will frequently annotate new images for long term storage. However, only the annotation need be stored and not the whole image itself. The textual information has been ignored in this simulation for two reasons. The first is that textual data is much smaller than pictorial data and hence will not contribute to the load of the system. The second is that the textual data can be handled by an entirely
different system which will use the same physical channel but will otherwise be completely independent. This makes sense when we consider the fact that textual data will not have to be subject to the same delays as image data.

4.3. Data Used in the Simulation Model

The simulation uses both estimates as well as survey data as input. The input data is classified in Table 7. As far as possible the survey data has been used as input data to accurately simulate the PACS environment. In some cases, e.g. the number of acquisition nodes, it was impossible to simulate each node since this would lengthen the simulation time considerably. Instead, the number of acquisition nodes have been reduced to three with variable influx rates.

<table>
<thead>
<tr>
<th>Diagnosis think times</th>
<th>This is the amount of time that the radiologist thinks after getting all the films and before going on to the next case.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of films per study</td>
<td>The number of images required for a diagnosis</td>
</tr>
<tr>
<td>Size of an image</td>
<td>The resolution of the image</td>
</tr>
<tr>
<td>DBMS time</td>
<td></td>
</tr>
<tr>
<td>Seek &amp; scan time for the magnetic disk</td>
<td></td>
</tr>
<tr>
<td>Optical disk attributes</td>
<td></td>
</tr>
<tr>
<td>Bus speeds</td>
<td></td>
</tr>
<tr>
<td>Influx rate of new images</td>
<td>From the acquisition nodes</td>
</tr>
</tbody>
</table>

Table 7: Input Data to the Simulation Model
The simulation was run for several values of the input parameters so as to give an idea of how the system would behave under different conditions. All the parameters correspond to peak loading conditions.

4.3.1. Estimates

4.3.1.1. Diagnosis Think Times

The main estimate of considerable importance is the diagnosis think time for each reading station. This determines the arrival rate of image requests from the reading station. A parametric analysis was done around the original values given in Table 8 below. The way this was done was to proportionately scale each think time by a factor which could be inputted to the program at the command line.

<table>
<thead>
<tr>
<th>Station No.</th>
<th>Minutes</th>
<th>Station No.</th>
<th>Minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5.0</td>
<td>6</td>
<td>1.5</td>
</tr>
<tr>
<td>1</td>
<td>5.0</td>
<td>7</td>
<td>5.0</td>
</tr>
<tr>
<td>2</td>
<td>1.5</td>
<td>8</td>
<td>3.0</td>
</tr>
<tr>
<td>3</td>
<td>3.0</td>
<td>9</td>
<td>1.5</td>
</tr>
<tr>
<td>4</td>
<td>10.0</td>
<td>10</td>
<td>1.5</td>
</tr>
<tr>
<td>5</td>
<td>10.0</td>
<td>11</td>
<td>10.0</td>
</tr>
</tbody>
</table>

Table 8: Diagnosis Think Times used for the Simulation

It must be mentioned that the reading stations are assumed to operate under heavy traffic conditions i.e. there is no idle time between diagnosis. This
represents the worst situation in terms of imposed workload.

4.3.1.2. Influx of Images from the Acquisition Nodes

This value was also an input to the program. The default value used was an overall rate of 1.5 images/minute, i.e. one image every 40 seconds. Though is a real life situation the acquisition node input is bursty, particularly for CT. in this simulation the influx is a steady rate right through the simulation run.

4.3.1.3. Frequency of Archive Access

The retrieval time of an image is naturally affected by which level of storage it resides in. The more frequently accessed data is stored in the upper levels. Frequency of access of the existing filerooms were obtained from Bowman Gray. These are given in Table 9.

<table>
<thead>
<tr>
<th>Fileroom</th>
<th>Probability of Access</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>0.72</td>
</tr>
<tr>
<td>D2</td>
<td>0.18</td>
</tr>
<tr>
<td>D3, D4 &amp; Warehouse</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Table 9: Frequency of Access of Existing Filerooms

In the simulation model D1 and D2 have been combined into one layer and the remaining three filerooms, a second layer.
4.3.2. Data obtained from the Surveys

4.3.2.1. Distribution of Number of Images per Study

The number of images that is requested for a diagnosis was directly obtained from the Radiological Survey. The distributions for each is given in Figures 24 to 26. These distributions show that it is quite probable for several images to be requested for a diagnosis. For example, the probability of station 0 requesting 76 films per study is as high as 0.12. Station 1 requests 48 films 14% of the time. Station 7 requests 28 films in 18% of the studies.

However, we can assume that many of these numbers represent isolated cases. There were 2208 samples in the Radiological survey. This averages to 184 samples per station. This is insufficient to accurately predict the probability of a certain number of films being required for a diagnosis. However, much can be inferred from the average number of films per requested per diagnosis. These are given in the table below.
Table 10: Average Number of Images per Diagnosis per Station

<table>
<thead>
<tr>
<th>Reading Station</th>
<th>Average Number of Images requested for a Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>22.7</td>
</tr>
<tr>
<td>1</td>
<td>19.2</td>
</tr>
<tr>
<td>2</td>
<td>4.3</td>
</tr>
<tr>
<td>3</td>
<td>10.0</td>
</tr>
<tr>
<td>4</td>
<td>13.9</td>
</tr>
<tr>
<td>5</td>
<td>4.9</td>
</tr>
<tr>
<td>6</td>
<td>36.3</td>
</tr>
<tr>
<td>7</td>
<td>14.8</td>
</tr>
<tr>
<td>8</td>
<td>10.3</td>
</tr>
<tr>
<td>9</td>
<td>4.9</td>
</tr>
<tr>
<td>10</td>
<td>18.2</td>
</tr>
<tr>
<td>11</td>
<td>10.0</td>
</tr>
</tbody>
</table>

It must be understood that the above numbers are based on current procedures. It is extremely difficult to predict what the effect of PACS on the diagnosis procedures will be [17].

4.3.2.2. Overall Arrival Pattern of Requests

Though the heavy traffic assumption was used in the model it is of interest to examine the arrival process of requests for film folders to the fileroom. The interarrival times of requests to the fileroom were calculated from data from the Fileroom survey. A Q-Q plot was made of this data to check the distribution of interarrival times. The Q-Q plot is a tool to find the probability distribution of a variable based on collected data [19]. The manner of finding out the distribution is as follows.
Q-Q plots (or probability plots) are used to determine the empirical distribution function $F_n$ for a variable defined at $X_{(i)}$ points. The method is to compare the given points to a known distribution function $F$ and assess if it falls in that category. The data is divided into quantiles pairs for $q = (i - 0.5)$ for $i = 1, 2, ..., n$. The $(i - 0.5)/n$ quantile of $F_n$ is $X_{(i)}$ and the $(i - 0.5)/n$ quantile of the known distribution function $F$ is $F^{-1}[(i - 0.5)/n]$. To compare the empirical distribution function $F_n$ to $F$ we thus plot the points

$$\left(X_{(i)}, F^{-1}\left(\frac{i - 0.5}{n}\right)\right)$$

For the interarrival times to the fileroom, the exponential distribution function is perhaps the most common and the data was checked against it. Hence the pairs

$$\left(x_{(i)}, -\ln\frac{n - i + 0.5}{n}\right)$$

were plotted. The Q-Q plot is shown in Figure 27.

From the figure above we can see that the plot is approximately a straight line. This shows that the request for film folders to the fileroom can be reasonably approximated to be a Poisson process.
Figure 23: Diagram of PACS Architecture
Figure 24: Distribution of Number of Films per Study for Stations 0 to 3
Figure 25: Distribution of Number of Films per Study for Stations 4 to 7
Figure 26: Distribution of Number of Films per Study for Stations 8 to 11
Figure 27: Q-Q plot for the Overall Request Rate to the Fileroom
CHAPTER 5

DESCRIPTION OF THE SIMULATION MODEL

The queueing model of the system is shown in Figure 28. The DBMS, optical disks, magnetic disks, and magnetic disk controllers had infinite buffers. At the DBMS a single request for an image is converted into several request for blocks if the image was found to reside in the magnetic disk system. These blocks are completely independent until they reach the reading station which waits until all the blocks of a particular image have arrived before requesting the next image.

The simulation was run on a VAX 11/750 with UNIX 4.2 BSD. All the input variables were supplied at the command line level. The output had three formats. The first was a trace through the programs execution. The second was a detailed listing of each device's characteristics and parameters. The third format (Appendix B) was a summary of the second with only the mean parameters of each type of device listed. A description of the input and output parameters follows as well as a description of the logic driving the program and the estimation techniques used to validate it.
5.1. Inputs to the model

The program is executed with the following command:

\%

\texttt{sim2 [options]}

The options were simulated time, bus speed, block size, compression ratios, influx rate, DBMS time, and the scaling factor for the diagnosis think times. The simulated time is the length of time in the radiology department that we wish to simulate. We found that since the diagnosis think times were of the order of 10 minutes it was necessary to simulate at least four hours in the hospital. The default value was four hours. Steady state was reached after about two hours of simulated time.

5.1.1. Speed of the bus

The speed of the bus greatly affects the image retrieval time. It is obvious that for some low value of the bus speed, the bus will start to be the bottleneck rather than the I/O subsystem. The bus speeds were varied from 0.1 megabytes/second to 10 megabytes/second. The default value is 2 megabytes/second.

5.1.2. Block sizes

When stored on the magnetic disk subsystem and transmitted via the bus, an image is broken up into several blocks. The block size used is hence the
packet size for the communication link as well as the record size in the mag-
netic disk subsystem. This simplifies management of the different blocks and
also allows independent recovery of each block and reassembly at the reading
station. The magnetic disk system can store with maximum efficiency two
blocks of 23476 bytes per track. The address mechanism can handle blocks up
to 32 kilobytes i.e. 15 bits. The block sizes considered here were from 4 kilo-
bytes to 1024 kilobytes (the smaller blocks were of practical importance, the
larger, academic interest). The default was 24 kilobytes per block.

5.1.3. Compression ratio

It was assumed that the images would not be stored in their entirety but
some compression algorithm would be used. Compression ratios vary depend-
ing on the type of data being compressed. An image of \( x \) megabytes can be
stored in only \( x/n \) megabytes of memory without loss of information where \( 1:n \)
is the compression ratio. At an average \( 1:10 \) is a reasonable figure and this
value is used as the default. The simulation program was run for compression
values ranging from 1.0 (no compression) to 0.05 (\( 1:20 \) compression).
5.1.4. Influx rate

The influx rate is a measure of the rate of storage of images. What it actually determines is the acquisition node think time. For \( n \) acquisition nodes and an influx rate of \( r \) per minute, the acquisition node think time \( t_a \) is effectively \( \frac{n}{r} \) minutes. The default value used was 1.5 images per minute i.e. a think time of 2 minutes per acquisition node. However this value was varied through a wide range: from 0.2 ( \( t_a = 900 \text{sec} \) ) which is basically a no input situation, to 160 ( \( t_a = 1.1 \text{sec} \) ).

5.1.5. DBMS time

The Data Base Management System (DBMS) must, from the time it receives a request, locate each block of the image and send this information to the magnetic disk controller. if the image resides in the upper level, or to the optical disk if it exists on the lower level. This operation would require looking up an index which would itself reside on a magnetic disk. A typical value for the DBMS time is 100 milliseconds [14]. This is the default value used. The program was run with different values of the DBMS time ranging from 50 mS to 1 sec.
5.1.8. Diagnosis think times

The diagnosis think times for each station were estimated by the radiologists at Bowman Gray. The diagnosis think times basically govern the frequency of diagnoses and indirectly the arrival rate of image requests. These figures were given in Table 8. Each diagnosis think time was scaled by some factor which was made an input variable. The default of this factor was 1.0. For parametric analysis this factor was varied from 0.2 through 100. The upper value represents the system at zero utilization.

5.2. Outputs from the model

The outputs consisted of measures of performance like queueing times, mean queue length, utilization as well as the overall retrieval time and the time spent in each component of the system. The format of the output that was used to create the plots of performance under different conditions is shown in the sample output in Appendix B.

5.2.1. Mean retrieval time for an image

The mean retrieval time for an image is probably the critical measure that one would look for. It is the time elapsing from the moment a reading station issues a request to the moment the image is fully loaded in the reading station. This time consists of all the queueing times, for the bus, DBMS, disk I/O
subsystem or optical jukebox, and bus again plus all the service times. In particular it consists of the time it takes for the request to reach and get processed by the DBMS plus the time required for all the blocks to be transmitted back to the reading station and reassembled.

5.2.1.1. Confidence interval calculation

The two important parameters of an endogenously created variable, i.e. a variable whose value is determined by the simulation program, are the mean and the standard deviation [18]. Also of importance are the percentiles of the variable. For instance, one may be interested in the 95% percentile which means that only 5% of the samples are greater than \( t_{95} \). To calculate the confidence interval \( t_{95} \) one must know the variance of the random variable, the expression for the confidence interval is,

\[
(\bar{x} - 1.96 \frac{s}{\sqrt{n}}, \bar{x} + 1.96 \frac{s}{\sqrt{n}})
\]

where \( \bar{x} \) is the mean of the random variable, and \( s^2 \) is the variance.

where the variance of a random variable \( X \) is defined as follows,

\[
s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2
\]

where \( \bar{X} \) is the mean and defined by

\[
\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i
\]

We thus reduce the problem to calculating the mean and variance of the
random variable. To reduce the confidence interval one either increases the number of samples or reduces the variance using a variance reduction technique. Autocorrelation affects the estimate of the variance. This can be circumvented using the batch means method.

5.2.1.2. The batch means method

While it is not necessary to store all the observations of $X_1, X_2, X_3, \ldots$ to calculate the mean, one sees that computation of the variance does need all the observations to be stored in memory. Due to autocorrelation the variance calculated by the above expression is much higher than that of an uncorrelated variable. The first step of the batch means method is to plot the autocorrelation coefficients for the variable vs. the lag $k$. From the intersection of this plot with the x-axis we get the intervals at which we can collect samples to calculate the variance. Since the autocorrelation at this interval is zero we effectively reduce the variance. However simply picking out these samples would affect the calculation of the mean of the variable. Hence we collect the data in batches only the means of which are stored. As can be seen below the mean of $X$ is unaffected. We define the variable $Z$ such that

$$z_i = \frac{x_m + x_{m+1} + x_{m+2} + \ldots + x_{m+b-1}}{b}$$

where $b$ is the batch size. Thus we see that $E(z) = E(z)$. In this program the batch is of size 50. The sample output (Appendix B) was obtained using the
default values of the input variables and shows that 254 batch means were collected i.e. a total of 12700 observations. The mean retrieval time was 0.9394 seconds. The standard deviation was 0.1982 secs and the confidence interval, 0.0244 secs. This confidence interval was sufficiently small at 95% confidence.

5.2.2. Delay, mean queue length, and utilizations

The queueing delay, mean queue length and the utilization of the DBMS, bus, magnetic disk, magnetic disk control unit and the optical disk were computed. The throughput of images was also calculated for a given set of input data.

5.3. The logic driving the program

The program is an event driven simulation. The events that can occur are a request occurs, the bus delivers a request at the DBMS. the DBMS finishes searching the index, a seek is done on a magnetic disk, the magnetic disk control unit transfers a block from the disk, the optical disk transfers an image, the bus delivers a block back at the reading station.

5.3.1. Flow chart for the program

A future event list is maintained in an ordered linked list. The only primary events are the requests that are randomly initiated at the beginning of
the program. Every succeeding event is conditional since the model is a closed queueing system. Figure 29 is a flow chart of the program execution.

5.3.2. Estimation procedures

Since the system itself does not exist internal validation was conducted. The first method was to check if the sum of the delays and service times of each component was comparable to that of the overall retrieval time. The second method was to increase certain input parameters and see if the output parameters behaved as expected. As can be seen from the results in the next chapter, the output parameters satisfy the second criterion.

To validate by the first method we consider the retrieval time $t_r$.

$$ t_r = t_p + q_{bus} + t_{bus} + q_d + t_d + p(m)t_{mag} + p(o)t_{opt} + n_b \times t_{bus}^{bl} $$

where,

- $t_p$: time for the reading station to create the request packet = 100 mS,
- $q_{bus}^{r}$: queueing time for a request to access the bus,
- $t_{bus}^{r}$: time to transmit the request packet over the bus = 0,
- $q_d$: queueing time for the DBMS,
- $t_d$: DBMS location time = 100 mS,
- $p(m)$: probability that the image is on magnetic disk = 0.9,
- $p(o)$: probability that the image is on optical disk = 1 - $p(m)$ = 0.1,
- $t_{mag}$: time for the average film in the magnetic disk system,
\( t_{\text{opt}} \) : time for an image to be retrieved from the optical disk system,

\( n_b \) : number of blocks in an image \( \approx 20 \),

\( t_{bus}^{bl} \) : time to transmit a block over the bus.

We now wish to calculate the waiting time for a request to access the bus. Since requests are high prioritized the average waiting time \( q_{\text{bus}} \) will be half the transmission time for a block multiplied by the bus utilization. Getting each of these values from the sample output we have,

\[
q_{\text{bus}}^r = t_{\text{bus}}^{bl} \times 0.5 \times \rho_{\text{bus}}
\]

where \( \rho_{\text{bus}} \) is the utilization of the bus. Hence we have from the above,

\[
q_{\text{bus}}^r = 12 \text{mS} \times 0.5 \times 0.215 = 1.29 \text{mS}.
\]

Also from the sample output, \( q_d = 3.7 \text{mS} \).

In calculating \( t_{\text{mag}} \) we consider what actually happens in the magnetic disk system. The blocks of a particular image can reside on any magnetic disk. When a request arrives at the magnetic disk the disk controller initiates a \textit{seek}. After the seek is completed, the disk queues up for use of the controller. There are 8 disks per controller. The magnetic disk is then blocked until it can access the channel and transfer the data. When it is assigned the channel a \textit{scan} is initiated and the data transferred out. Thus it is the magnetic disk and the controller with the most blocks which decides the time that the image will take to transfer out. Since calculating the expected value of the maximum is extremely difficult we find the bounds for \( t_{\text{mag}} \).
The lower bound is when all the blocks are evenly distributed among the disks. There are 32 disks and an average of 20 blocks per image. In the best case the blocks will be evenly distributed over all 4 channels. Thus the time taken will be that for 5 blocks per channel = 5 × 10.8 mS. The lower bound of $t_{mag}^l = 16 \text{ mS} + 54 \text{ mS} = 70 \text{ mS}$.

The worst case will occur when all the blocks happen to be on the same disk. The variable $t_{mag}^u$ will then be $20 \times (16 \text{ mS} + 10.8 \text{ mS}) = 536 \text{ mS}$. We also add the queueing times at the magnetic disk and the magnetic disk controller in the worst case. Thus $t_{mag}^u = 536 \text{ mS} + 26.1 \text{ mS} + 11.3 \text{ mS} = 573.4 \text{ mS}$.

We calculate $t_{opt}$ using the equation in Section 4.1.1.2. For an image of 0.5 megabytes we have $t_{opt} = 4.04 + 0.171 \times 0.5 = 4.125 \text{ secs}$. Thus we get from the equation above for the retrieval time, the lower bound is

$$t_{i}^l = 100 + 1.29 + 3.7 + 100 + 0.9 \times 70 + 0.1 \times 4125 + 20 \times 11.8 = 916.5 \text{ mS}$$

and the upper bound is

$$t_{i}^u = 100 + 1.29 + 3.7 + 100 + 0.9 \times 573.4 + 0.1 \times 4125 + 20 \times 11.8 = 1378.6 \text{ mS}.$$  

The simulated value of the retrieval time $t_i^m$ from the sample output is 939.4 mS. Since $916.5 \text{ mS} \leq 939.4 \leq 1378.4 \text{ mS}$, we consider the model validated.

5.3.3. Programming methodology

The language used for the program was C under the UNIX environment. A hierarchical programming approach was adopted and each file calls functions
defined in the file immediately below it. The program is contained in six files
the contents of which is as follows: *declare.c* contains the declarations and
structure definitions for the queues and the arrays; *arrays.c* contains the book-
keeping functions to calculate the mean, variance and other parameters;
*queue.c* holds all the queue handling routines like ordered insertion for the
event list, *fifo* insertion for the node queues, and the delete routines; *rand.c*
contains the routines for the generation of the random variables such as
number of images per diagnosis and other functions returning parameters such
as the size of images, the optical disk time, etc; *funcs.c* stores all the conditional
events functions which are triggered off in the order described in the flow chart;
*sim2.c* sits above *funcs.c* and runs the overall simulation program.

A listing of the program is given in Appendix A. The six files are listed in
decreasing order i.e. *sim2.c, funcs.c, rand.c, queue.c, arrays.c, and declare.c.*
Figure 28: The Queueing Model for the System
Figure 29: Flow Chart for the Program
CHAPTER 6

RESULTS

The aim is to try to find general behavioral patterns in the PACS. Certain performance characteristics are required of the components of the PACS i.e. bus speed, DBMS time. An attempt was made to define these critical parameters. The organization of the results is a series of graphs showing the behavior of the system under various stimuli. The conclusions that can be drawn from these graphs are shown in the next section.

All the optimum points have been obtained by keeping the remaining parameters fixed while varying one or two parameters.

6.1. Mean retrieval times under different conditions

The input parameters have been varied and for each parameter three graphs are shown. The first shows the effect on the graph of changing a second input parameter, most frequently the bus speed, the second shows the breakup of the mean retrieval time into its different components, and the last gives the utilizations of the bus, DBMS, magnetic disk subsystem, and the optical disk subsystem. Wherever possible, analytical derivations have been included confirming the results arrived at through simulation.
We introduce the concept of the average arrival rate of requests for images, which will be used later in this chapter to verify some of the simulation results, and consider the arrival rate of requests from station $i$, $\lambda_i$

$$\lambda_i = \frac{\text{number of images requested per study}}{\text{cycle time for station } i}$$

where the cycle time is basically the time to do one diagnosis, and the number of images requested per study are those given in Table 10. Thus,

$$\lambda_j = \frac{n_j}{D_j - n_j \times t_i}$$

where $D_j$ is the diagnosis time for station $j$, and $t_i$ is the average time to retrieve an image. Thus the average arrival rate $\lambda$ to the system is 0.88276 images/second with $t_i = 1.0$ sec.

$$\lambda = \sum_{j=0}^{11} \lambda_j = \sum_{j=0}^{11} \frac{n_j}{D_j + n_j \times t_i}$$

This is not a closed form solution for $\lambda$ since $t_i$ depends on $\lambda$ itself. However, for small $t_i$ we have the approximation,

$$\lambda = \sum_{j=1}^{11} \frac{n_j}{D_j}$$

Substituting $D_j$ and $n_j$ from Tables 8 and 10 we get $\lambda = 63.4533$ images/minute (1.05755 images/second). This is approximately 15% higher than the more accurate value calculated previously.
6.1.1. Bus speed variations

Figures 30 and 31 show the behavior of the system with changing bus speeds. As expected we see that the mean retrieval time decreases with higher bus speeds. However, we wish to find out what the minimum bus speed is that will result in acceptable retrieval times. We also know that increasing the speed beyond a certain threshold will only marginally improve the response time of the system. Identifying that threshold is of primary importance.

From Figure 30 we see that at a compression ratio of 0.8 the retrieval times remain virtually unchanged beyond 4 MBps (32 Mbps). This speed is very difficult to achieve especially when software delays are taken into account. However with a compression ratio of 0.1 i.e. effectively reducing the image size to one-eighth the previous value we see that the necessary bus speed to achieve subsecond retrieval times is 2 MBps (16 Mbps). Inspection of Figure 31 supports this conclusion. We see that at 2 MBps the delays at the magnetic disk subsystem and the bus are equal. Looking at the utilization curves (Figure 32) we see that at 0.1 MBps the utilization of the bus is almost unity identifying the bus as the bottleneck of the system. At 2 MBps the utilization drops to 0.2 which is acceptable. Taking a look at the utilizations of the other devices we find that their utilizations drop at very low bus speeds. This is due to the fact that the simplification for \( \lambda \) no longer holds since \( t_r \), the retrieval time for an image is of the same order as \( D \), the diagnosis times. Hence the value of \( \lambda \)
drops significantly.

Let us now derive an expression for the utilization of the bus and the DBMS and compare it to the figures on the graphs. The DBMS utilization $\rho_{dbms}$ is given by

$$\rho_{dbms} = \frac{\lambda}{\mu_{dbms}}$$

where $\mu_{dbms}$ is the service rate of the DBMS = 10 images/sec. Thus we get $\rho_{dbms} = 0.882759/10.0 = 0.088$. This corresponds to the simulated value of 0.088.

For the bus, taking the example of 2 MBps,

$$\rho_{bus} = \frac{\lambda}{\mu_{bus}} = \lambda \times \text{transmission time/block} \times \# \text{ blocks/image} = 0.21769$$

The simulated value $\rho_{bus}$ from the sample output is 0.215122.

6.1.2. Compression ratio variation

We now consider the effect of compression of images on the PACS. Compression is used to store the same amount of information in less memory. Not only is the storage and retrieval from memory more efficient but the communication of images is also faster as a result. In the terminology used a higher compression ratio implies better compression and a smaller amount of memory used for the same image.

We consider the effect of varying the compression ratio from 0.05 (1:20) (which probably is the upper bound for radiographic images) to 1.0 (no
74 compression). Figures 33 to 35 show how the system responds to this variation. The compression of images effectively changes the number of blocks per image since the block size is fixed and the effective image size is reduced with compression. Figure 33 displays the effect of compression on the mean retrieval time for an image at three bus speeds. As can be seen, the average response times for images is almost linear with compression ratio. The nonlinearity is due to increased queueing delays as the number of blocks per image increases.

From the same figure we see that a 0.2 compression ratio (or higher) is necessary for subsecond retrieval times when the bus speed is 2 MBps. With a 5 MBps bus lesser compression ratios (≤0.5) can be tolerated. Figure 34 gives the delays at each device with varying compression ratios. Naturally, the queueing delays at the magnetic disk system and the bus increase. However, more can be construed from Figure 35 which shows the utilization of each device. The compression ratio has maximum effect on the bus which turns out to be the bottleneck as the compression ratio is decreased.

6.1.3. Block size variation

Some very interesting results appear when we compare the mean retrieval times for different block sizes. The block is used as the record size on the magnetic disk subsystem as well as the packet size for the communications system. The overhead of each record is one seek and one scan to locate the block. The
overhead of each packet is a header of length 100 bytes containing address, sequence and parity information. When we look at Figure 36 which shows the mean retrieval time for an image at different bus speeds it appears that there is an optimum block size of 32 kB. This size is independent of the bus speed. Fortunately, this is very close to the optimum storage size of 24 kB (see Section 4.1.1.3) as far as efficiently utilizing magnetic storage space is concerned.

To explain this behavior one must look at Figure 37 which gives the mean retrieval time in terms of component delays. We can explain the U-shape of the magnetic retrieval time as follows. For small block sizes the seek and scan times are of the same order as the block transfer time. For example, it takes 0.9 milliseconds to transfer a block of size 2 kB at the rate of 17.4 Mbps. Compared with this the scan time is of the order of 7 milliseconds and the seek time, 16. Thus it is easy to understand why small block sizes result in increases delays.

On the other hand, as we increase the block size the concurrency of transferring the image (through different channels via the disk control unit) is reduced. This is as a result of uniformly distributing the records of an image among the magnetic disks. For example, if the block size is 512 kB then we have an average of 1.78 blocks per image. Whereas with 24 kB blocks we were able to concurrently transmit the image through four channels (with four disc control units), we now have to send the entire image through one or two chan-
nels.

Another factor which affects the retrieval time is the fragmentation of the last block. Since the image size is a continuous variable depending on the compression used, we assume that the last block is only half utilized. Therefore there is an overhead of $B/2$ where $B$ is the block size. Naturally this overhead increases with larger blocks.

To explain the behavior of the "bus" curve in Figure 37 we have to keep in mind that the plot represents the queueing time of one block. With smaller block sizes there will be a correspondingly increased number of blocks. The upward curve around 2 kB is the effect of having comparable packet sizes and header information.

Corresponding to the other two graphs we look at Figure 38 which shows the utilization of each device with block size variation. The block size does not affect either the DBMS or the optical disk system (images on the optical disk are stored in their entirety and not divided into records for reasons mentioned in Section 4.1.1.2).

6.1.4. Influx variation

There are three acquisition nodes in this architecture each of which sends an image to be stored every $t$ seconds. Figures 39 to 42 show how different aspects of the performance are affected by the overall influx of images to the
system. All the graphs belonging to this section are derived with standard diagnosis think times i.e. a constant rate of requests for images from the reading stations. Figure 39 shows the effect of the influx rate on the average retrieval time for an image. One notices that an influx rate up to 30 images/minute (1800 images/hour) will only marginally affect the performance of the system. Looking at Figure 40 we see that this increase is due to the queueing time to access the bus. Thus under heavy traffic conditions the bus speed is of critical importance. Figure 41 shows how the utilization of each device is affected by the influx rate. The utilization of all the devices naturally increases but noticeable is the rapid increase in the utilization of the bus.

Figure 42 is very interesting twist to Figure 39. It shows the retrieval times as plotted against the request/influx ratio. In a real life situation, the request rate is always much greater than the influx rate since each film is stored only once but read several times. We see that the influx/request ratio can go as high as 0.6 without substantially affecting the system's performance. This is encouraging since it shows that the acquisition side of a PACS does not affect the retrieval process to any substantial degree.

8.1.5. DBMS time variation

Figure 43 gives a plot of Mean retrieval times vs. DBMS times. As we can expect the retrieval time increases with the DBMS time. Surprisingly, the
retrieval times turn out to be independent of the bus speed as the DBMS times exceed 500 mS. This we can attribute to a decreasing queuing delay to access the bus as well as increased magnetic disk retrieval times (Figure 44). Thus the transmission time, though linearly dependent on the bus speed, turns out to be insignificant when compared to the other device times in the system. The utilization of the DBMS increases with the service time of the DBMS. However, we see that though the queueing delays of the bus and magnetic disk subsystem change, the utilization of these devices remain unchanged. This leads us to the conclusion that increasing the DBMS time changes the arrival process to the other devices while keeping the average arrival rate constant.

8.1.6. Diagnosis think time variation

When we look at the variation of retrieval times vs. think times we notice some extremely interesting results. The x-axes on Figures 46 to 48 represent a ratio to the normal think times. Thus a ratio of 1 implies that the think times are as defined in Table 8 while a think time of 0.5 means that the think times used in the simulation are half those defined in Table 8. Taking this reasoning to its limit we see that a ratio of 0 will result in continuous requests for films from a particular station while a ratio of \( \infty \) would imply that the studies are so infrequent that we can assume that the PACS is at any given moment handling at most one study i.e. one image with single threading.
Figure 46 gives the variation of mean retrieval times with think times. Obviously there exists some minimum retrieval time for the PACS to retrieve one image without interference from other requests. As we can see from Figure 46 for think time ratios greater than 0.6 the average response time remains basically unchanged. Hence we come to the conclusion that the architecture proposed can accommodate $66\% \left(1/0.6\right)$ more studies per hour without noticeable degradation. We see that this is true for any bus speed.

Considering Figure 47 we see again that the device delays are constant for ratios greater than 0.6. The queueing delays are due to the fact that from the DBMS onwards through the magnetic disk system and the bus the arrival process is a bulk arrival process. The image is broken up into several blocks which are retrieved from the magnetic disk subsystem and transmitted via the bus independently. These blocks are identical, and uniformly distributed among the magnetic disks.

Figure 48 demonstrates how the utilization of each device tends to zero as the average arrival rate of film requests goes to zero. This behaviour of the system further validates the model.

Thus we are presented with several graphs which characterize the behaviour of this system. The results given in this chapter are summarized in the next chapter.
Figure 30: Mean retrieval times vs. Bus speeds

Figure 31: Device delays vs. Bus speeds

Figure 32: Utilization of devices vs. Bus speeds
Figure 33: Mean retrieval times vs. Compression ratio

Figure 34: Device delays vs. Compression ratio

Figure 35: Utilization of devices vs. Compression ratio
Figure 36: Mean retrieval times vs. Block size

Figure 37: Device delays vs. Block size

Figure 38: Utilization of devices vs. Block size
Figure 39: Mean retrieval times vs. Influx rate

Figure 40: Device delays vs. Influx rate

Figure 41: Utilization of devices vs. Influx rate
Figure 42: Mean retrieval times vs. Request/influx ratio

Figure 43: Mean retrieval times vs. DBMS times

Figure 44: Device delays vs. DBMS times
Figure 45: Utilization of devices vs. DBMS times

Figure 46: Mean retrieval times vs. Diagnosis think times

Figure 47: Device delays vs. Diagnosis think times
Figure 48: Utilization of devices vs. Diagnosis think times
CHAPTER 7

DISCUSSION AND CONCLUSION

In this thesis the PACS environment was examined and the traffic was quantified. The traffic can broadly be represented as requests for retrieval of films and storage of films. The fileroom is the node to which these activities are directed. Requests may come from radiologists diagnosing a case or a referring physician. The influx of films is from examination rooms. Thus three surveys were taken to characterize the traffic in the radiology department: the Fileroom survey, the Radiological survey, and the Admissions survey.

To computerize this system one must consider the archival, processing, and storage of digital images. In this report we presented the traffic that can be expected from each of these processes. Of interest is the fact that the request rate is prone to fluctuation and the peak (Monday: 7 - 8 am) can be as much as 4 times the average during working hours.

The Fileroom survey was further subdivided into exam categories and the components due to each examined. Plain films account for a majority of the film requests (71.5%). A peak in the number of request was found between 7 pm and 8 pm. This figure represents the number of prescheduled diagnoses. Also of interest is the fact that 1% of the folders had a location time greater that 50 minutes and 5%, greater than 20 minutes. This is a substantial amount
In conclusion, the PACS is eminently viable as far as capacity allocation is concerned. However, what has not been considered here and will probably be a major problem is software that will run at the speeds shown to be necessary if the PACS is to speedily deliver images. For instance, if the communications software results in an effective data transmission rate of 500 kilobytes/second the retrieval time increases to 2.5 secs. An effective speed of 100 kilobytes/second results in colossal retrieval times of 32.7 seconds. When considering the fact that around 25 of these images must be retrieved for a diagnosis we are then faced with a study recovery time of around seven minutes.

Other problems concern interfacing the PACS with current equipment, and creating applications software that will create an environment for the radiologist that will not greatly differ from what he is used to. Efficient data structures have been assumed. These will have to be developed. However, all things taken into account, it has been shown that a PACS as described is definitely feasible and is something of the very near future.
LIST OF REFERENCES


PACS IV, 1986.


APPENDIX A

APPENDIX A. PROGRAM LISTING FOR THE SIMULATION
#include "funcs2.c"

/*
 * This is a simulation program to simulate the
 * environment of a Picture Archival and Communication System.
 * The program is event driven with the event codes listed below.

This program was written in March to May 1986.

Programmer: Vijay D'Silva (vgd@ncsu)

0 to STAT - requests from stations
25 to 25+ STAT diagnosis
50 to 50+OPT optical disc retrieval
55 to 55-MAG magnetic disc retrieval
100: bus
101: dbms
125 to 125 - CONT controller operation

main(argc,argv)
int argc;
char *argv[];
{
    int code,i;
double iter=7200.00;
    if (argc>1) iter=(atoi(argv[1])-7200.0);
    if (iter==0.0){
        printf("sim2 time(7200) SPEED(10) BLOCK(24) COMP(0.1) 
        INPUT(1.5) DBMS(0.1) DIAGSC(1.0)\n");
        exit(1);}
    if (argc>2) SPEED=atof(argv[2]);
    if (argc>3) BLOCK=atof(argv[3])/1024.0;
    if (argc>4) COMPACT=atof(argv[4]);
    if (argc>5) INPUT=180.0/atof(argv[5]);
    if (argc>6) dbms_time=atof(argv[6]);
    if (argc>7) DIAGSCALE=atof(argv[7]);

    init();
    while (master iter){
        firstevent(&master,&code,&eventhead);
        if (code == 25)
            request(code);
        else if (code < 50)
            diagnose(code - 25);
        else if (code < 55)
            optical(code - 50);
        else if (code < 100)
            magnetic(code - 55);
        else if (code == 100)
            bus();
        else if (code==101)
            dbms();
        else controller(code - 125);
    }
    results();
}
funcs.c

#include "rand.c"

/* called when a request from a reading station occurs */
int code;

/* printf REQUEST %d films %d master %d n", code, numfilms[code], master); */
if (busq.queue == NULL)
  insert(master + bus_time(-1), 100, &eventhead);
/* printf "busp " */
  update(0.0, &bspecs);
/* printf "bus "; */
  addq(code + 100, &busq); /* code for request packet */
}

/* called when a diagnosis is done and the next batch of films must be requested. */
diagnose(term)
int term;

/* printf DIAGNOSE %d n", term); */
numfilms[term] = num_img(term); /* # images
update(master - diagask(term) - diag_time(term), &dspecs[term]);
diagask[term] = master; /* for total diag time */
insert(master - prot_time, term, &eventhead); /* 1st image */

/* called when an optical disk finds a film */
optical(disc)
int disc;

int term = last(&time, optq[disc], i); /* which station
/* printf OPTICAL %d n", disc); */
rmlast(&optq[disc]);
for (i = 0; i < blockcount[term]; i++){
  if (busq.queue == NULL){
    insert(master - bus_time(100), 100, &eventhead);
    /* printf "busp "; */
    update(0.0, &bspecs);
    /* printf "bus "; */
    addq(term, &busq);
  } else if (optq[disc].queue != NULL){
    insert(master - optseek(last(&time, optq[disc]), 50 + disc, &eventhead);
    /* printf "osp%d ", disc); */
    update(master_time, &ospecs[disc]);
  }
}

/* called when a magnetic disk finishes a "seek" */
magnetic(disc)
int disc;

int term = last(&time, magq[disc], contra = (int)(disc / SHARE);
/* printf MAGNETIC %d n", disc); */
if (contq[contra].queue == NULL){
  insert(master + magxfer(100) + SCAN, 125 + contra, &eventhead);
  update(0.0, &cspecs[contra]);
  addq(disc, &contq[contra]);
}

/* called when the disk controller finishes a transfer */
controller(code)
int code;

int disc = last(&time, contq[code], term;
```c
/* controller */

if (code == 3)
  printf("%d
", disc),traverse(contq)); /*
/prnflf "CONTROLLER %d
",code); */
rmlast(&contq[code]);
term = last(&time, maqq[disc]);
update(master-enterdisk[term], &dsksubsys);

/channel
% d
% d
,code, disc); */
rmlast(&maqq[disc]);
if (busq.queue == NULL){
  update(0.0, &bspecs);
}

/*
bus */
addq(term, &busq);
if (maqq[disc].queue != NULL){
  update(master + SEEK, disc), &mspecs);
}

if (code[queue] == NULL){
  update(master + magxfer(100), &eventhead);
}

int term = last(&time, busq);

/BUS %d
,term); */
rmlast(&busq);
if (term >= 100){
  term -= 100;
  if (dbmsq.queue == NULL){
    update(master + dbms_time, 101, &eventhead);
  }
}

bus() {

int term = last(&time, busq);

/BUS %d
,term); */
rmlast(&busq);
if (term >= 100){
  term -= 100;
  if (dbmsq.queue == NULL){
    update(master + dbms_time, 101, &eventhead);
  }
}
}
```

---

The function appears to handle a controller for a bus system, likely part of a larger system or software, and is designed to manage various operations related to disk and bus operations. The code snippet includes functions for updating disk and bus statuses, managing queues, and inserting operations into queues. It also includes conditionals for handling different types of operations based on the input code values.
/* called when the DBMS finishes locating an image */
dbms()
{
 int disc,type=hierarch(),term=last(&time,dbmsq);i;
 /*printf("DBMS %d \n",term); */
 rmlast(&dbmsq);
 blockcount[term]=(int)(sz_imag(term) /BLOCK+1.0);
 if (type<=1){
 /*printf("number of blocks %d, blockcount[term]), */
 enterdisk[term]=master;
 for (i=0;i<blockcount[term];i++){  
disc=(int)(randm() *MAG);
 count[disc]+=;
 if (magq[disc].queue==NULL){
 insert(master+SEEK+i,0.05*dbms_time,55+disc,&eventhead);
 /*printf("%d ",disc),*/
 /*printf("%d
",count[disc]);
 if (dbmsq.queue!=NULL){
 insert(master+dbms_time,101,&eventhead);
 term=last(&time,dbmsq);
 /*printf("disp ");
 update(master-time,&aspecs);
 }

 /* initialization stuff
init()
{
 int i;
 printf("fileinit\n ");
 fileinit();
 for (i=0;i<SAMPLES;i++)
 fillr[i]=1000.00;
 for (i=0;i<STAT;i++){
 initspecs(&sspecs[i]);
 initspecs(&dspecs[i]);
 blockcount[i]=0;
 insert([statask[i]=(diagask[i]=randm() *2.5)],i,&eventhead);
 numfilms[i]=num_imag(i);
 for (i=0;i<OPT;i++){
 initspecs(&ospecs[i]);
 initstation(optq[i]);}
 for (i=0;i<MAG;i++){
 count[i]=0;
 initspecs(&mspecs[i]);
 initstation(magq[i]);}
 for (i=0;i<CONT;i++){
 initspecs(&c.Specs[i]);
 initstation(contq[i]);
 initspecs(&bspecs);
 initstation(busrq);
 initspecs(&aspecs);
 initstation(dbmsq);
 initspecs(&disksubsys);
}
called at the end. recovers two kinds of output with a
page feed between, the first is a listing of each device
and its parameters (useful for checking consistency). the
second type gives the overview of the output statistics

```c
results()
{
  int i, num, numag = 0, numopt = 0, numcont = 0, reqfilms = 0, stofilms = 0;
  double mean, var, sigwait = 0, varopt = 0, mmag = 0, mopt = 0, mcont = 0, alength, util;
  double sigleng = 0, sigutil = 0;
  for (i = 0; i < MAG; i++)
    addq(0, &magq[i]);
  for (i = 0; i < OPT; i++)
    addq(0, &optq[i]);
  for (i = 0; i < CONT; i++)
    addq(0, &contq[i]);
  addq(0, &busq); addq(0, &dbsq);
  printf("R %7s %9s %3s %5s %s\n", "SERVER", "MEAN", "VARIANCE", "NO.", "BLKS");
  printf("R %7s %7s %8s %4s %5s %5s %5s\n", "BUSTIME", "OPTSEEK", "DIAGNS", "IMGS", "DIAG", "MDIAG", "VDIAG", "NDIAG");
  for (i = 0; i < STAT; i++)
    getspecs(sspecs[i], &mean, &var, &num);
    getstatsq(magq[i], &alength, &util);
    numag += num; mmag += mean - umag;
    printf("R %3s%2d %9.7f %9.7f %9d %9.7f %9.7f
", "stat", i, mean, var, num, (int)(size_imag(i) / BLOCK - 1.0));
  getspecs(ospecs[i], &mean, &var, &num);
    getstatsq(optq[i], &alength, &util);
    numopt += num; mopt += mean - umopt;
    printf("R %3s%2d %9.7f %9.7f %9d %9.7f %9.7f
", "opt", i, mean, var, num, alength, util);
  getspecs(bspecs, &mean, &var, &num);
    getstatsq(bsq, &alength, &util);
    printf("R %3s%2d %9.7f %9.7f %9d %9.7f %9.7f %9.7f
", "bus", i, mean, var, num, alength, util);
  getspecs(aspecs, &mean, &var, &num);
    getstatsq(dbsq, &alength, &util);
    printf("R %3s%2d %9.7f %9.7f %9d %9.7f %9.7f
", "dbm", i, mean, var, num, alength, util);
  for (i = 0; i < CONT; i++)
    getspecs(cspecs[i], &mean, &var, &num);
    getstatsq(contq[i], &alength, &util);
    numcont += num; mcont += mean - numcont;
    printf("R %3s%2d %9.7f %9.7f %9d %9.7f %9.7f
", "cont", i, mean, var, num, alength, util);
  printf("R time for request protocol %9.7f\n", prot_time);
    printf("R seek time on mag disc %9.7f\n", SEEK);
    printf("R scan time for mag disc %9.7f\n", SCAN);
    printf("R seek time on opt disc %5.3f\n", seek_opt);
    printf("R total number of magnetic disc blocks %d\n", numag);
    printf("R total number of optical disc files %d\n", numopt);
    printf("R average # blocks in a file: %7.2f\n", size_retrieve());
  printf("R")
  printf("R TIME SIMULATED: %f sec\n", *master);
  printf("R")
  printf("R")
  printf("R")
```
for (i=0;i<STAT;i++) {
    getspecs(sspees[i],&mean,&var,&num);
    if (i<12) reqfilms += num; else stofilsms += num;
    printf("station %2d %5.3f seconds",i,mean);
    getspecs(specs[i],&mean,&var,&num);
    printf("%6.3f seconds %6.3f mean,meanimag(i));
    getspecs(specs[i],&mean,&var,&num);
    printf("%6.3f mean,num %800 /master");
}
samprtrieve(&mean,&var,&num);
printf("\nOVERALL RETRIEVAL TIME FOR FILMS: MEAN %7.4f seconds ;VARIANCE %7.4f seconds ; NUMBER OF SAMPLES %d\n",mean,1.96 * var / sqrt(num * 1.0));
for (i=0;i<MAG;i++) {
    getstatsq(magq[i],&length,&util);
    sigleng=sigleng+length:sigutil+=util;
    printf("\nMAGNETIC DISC MEAN WAITING TIME: %7.4f sees MEAN NUMBER IN DISC: %7.4f MEAN UTILIZATION: %7.4f\n",mean,numag/mag,sigleng/MAG,sigutil/MAG);
    sigleng=(sigutil=0);
}
for (i=0;i<OPT;i++) {
    getstatsq(optq[i],&length,&util);
    sigleng=sigleng+length:sigutil+=util;
    printf("\nOPTICAL DISC MEAN WAITING TIME: %7.4f sees MEAN NUMBER IN DISC: %7.4f MEAN UTILIZATION: %7.4f\n",mean,numopt/numopt,sigleng/OPT,sigutil/OPT);
}
for (i=0;i<CONT;i++) {
    getstatsq(contq[i],&length,&util);
    sigleng=sigleng+length:sigutil+=util;
    printf("\nDISC CONTROL UNIT MEAN WAITING TIME: %7.4f sees MEAN NUMBER IN CONT: %7.4f MEAN UTILIZATION: %7.4f\n",mean,numcont/numcont,sigleng/CONT,sigutil/CONT);
}
getspea(dsksubsy,$mean,&var,$num);
printf("\nAVE.TIME FOR A BLOCK IN MAG DISK SUBSYST %7.4f sees\n",mean);
getspea(bspe,$mean,&var,$num);
getstatsq(busq,&length,&util);
printf("\nBUS mean waiting time for a block: %7.4f secs\n",mean);
printf("\n mean number in bus system: %7.4f utilization: %9.7f\n",alength,util);
getspea(&mean,&var,$num);
getstatsq(dbsq,&length,&util);
printf("\nDBMS mean waiting time for an image: %7.4f secs\n",mean);
printf("\n mean number in dbms system: %7.4f utilization: %9.7f\n",alength,util);
printf("\nTRANSMISSION TIME FOR A BLOCK OVER THE BUS: %7.4f\n",mean);
printf("\nTRANSFER TIME FROM DISC (READ / WRITE): %7.4f\n",mean);
printf("\nINFLUX %7.4f \n",mean);
for (i=0;i<100;i++) {
    printf("\nrequests %7.4f /min ; REQUEST /INFLUX %7.4f \n",reqn.lms/reqn.lms /master,reqn.lms /master,reqn.lms /1.0 /reqn.lms);
    printf("\n\fW);
#include "queue.c"

/* return a random number between 0 and 1 */
float randm()
{
    return(random()/(32788.0*32788.0*2.0));
}

/* return the diagnostic time for reading station 'term' */
float diag_time(term)
int term;
{
    switch (term){
    case 0 : return(5.0*0.0*DIAGSCALE);
    case 1 : return(1.5*0.0*DIAGSCALE);
    case 2 : return(3.0*0.0*DIAGSCALE);
    case 3 : return(5.0*0.0*DIAGSCALE);
    case 4 : return(10.0*0.0*DIAGSCALE);
    case 5 : return(1.5*0.0*DIAGSCALE);
    case 6 : return(10.0*0.0*DIAGSCALE);
    default: return(INPUT);
    }
}

/* returns 1 if the station is an acquisition node and a random number based on the distribution for the station if the station is a reading station */
int num_imag(term)
int term;
{
float f=randm();
int i=0:
    if (term>11) return(1);
    while(imagcdf[term][i]<f) i++;
    return(imagnum[term][i]);
}

/* size of the image in megabytes */
float siz_imag(term)
int term;
{
float mb=8.0*1024.0*1024.0 /COMPACT;
    switch (term){
    case 0 :
    case 1 : return(256.0*256.0*8.0 /mb);
    case 9 :
    case 10 :
    case 2 :
    case 3 :
    case 8 :
    case 7 :
    case 8 : return(2000.0*2000*12.0 /mb);
    case 11 :
    case 4 :
    case 5 : return(512.0*512.0*12.0 /mb);
    default: return(2000.0*2000*12.0 /mb);
    }
/* randomly selects the hierarchy of storage */
int hierarch()
{
    float r = rand();
    if (r <= .72)
        return(0);
    else if (r <= .9)
        return(1);
    else return(2);
}

/* storage transfer rate in MB/sec */
float rate(h)
int h:
{
    switch (h){
        case 1:
        case 0: return(17.4/8);
        default: return(15.0);
    }
}

/* storage parameters in seconds */
double optseek(term)
int term:
{
    return(TD - (siz_imag(term) /228)* TR);
}

/* time to transfer an image through the channel */
double imagxfer(term)
int term:
{
    switch (term){
        case 100: return(BLOCK /rate(0));
        default: return(siz_imag(term) /rate(0));
    }
}

/* time to transfer a block through the bus */
double bus_time(term)
int term:
{
    switch (term){
        case 1: return(1e-3/SPEED); /* 1 kB for a request */
        case 100: return((BLOCK+1.0e-4)/SPEED);
        default: return(siz_imag(term)/SPEED);
    }
}

/*initializing the parameters for the number of images the program reads from files cdf[0-11] */
fileinit()
{
    FILE *p;
    int i,m;
    char *name;
    float f;
    /* sprintf(name,"cdf%d",i); */
    /* p=fopen("cdf","r");
    */
    fscanf(p,"%f
",f);
    fscanf(p,"%f
",f);
    fscanf(p,"%f
",f);
}
```c
BLOCK = BLOCK / 1024;0;
scanf(fp, "%8.2f", &COMPACT);
scanf(fp, "%8.2f", &INPUT);

for (i=0;i<12;i++)
  for (m=0;m<20;m++)
    scanf(fp, "%8.2f", &imagcdf[i][m], &imagnum[i][m]);
  fclose(fp);

for (i=0;i<12;i++)
  for (m=0;m<20,m++)
    printf("%8.2f", imaged[i][m], imagnum[i][m]);

/*
   float meanimag(term)
   int term;
   {
     int i;
     float f=imagcdf[term][0], imagnum[term][0];
     if (term > 12) return(1); /* input modalites */
     for (i=1;i<20;i++)
       f=f+imagnum[term][i]*imagcdf[term][i]-imagcdf[term][i-1];
   }
   return(f);
*/
```
```c
#include "arrays.c"

function to delete the tail of the queue of type LINK

rmlast(pq)
STATION *pq;
{
    LINK temp = (*pq).queue;
    (*pq).milestone = master; (*pq).length --;
    if ((*pq).queue == NULL)
        printf("CANNOT DELETE FROM AN EMPTY QUEUE\n");
    else if (temp->nexts == NULL)
        free (temp);
        (*pq).queue = NULL;
    else while (temp->nexts->nexts != NULL)
        temp = temp->nexts;
        free (temp->nexts);
        temp->nexts = NULL;
    
    if ((*pq).queue == NULL)
        if ((*pq).length != 0) printf("Aaaaaaah!!!\n");
}

function to get the last element of a queue (return 0 if empty)

int last(timep,q)
STATION q;
{ 
    LINK temp = q.queue;
    if (temp == NULL)
        return(0);
    else while (temp->nexts != NULL)
        temp = temp->nexts;
    *timep = temp->get;
    return(temp->ask);
}

function to insert an element in a queue

insert(time,i,qp)
double time;
int i;
LINK *qp;
{ 
    LINK temp,curr = *qp,temp1;
    /*printf("insert %f , %d
",time,i);*/
    temp = (LINK)malloc(sizeof(REQUEST));
    temp->ask = time; temp->get = i *.0; temp->nexts = NULL;
    if (*qp == NULL)
        *qp = temp;
        return(0);
    if ((*qp)->ask >= time)
        temp->nexts = *qp;
        *qp = temp;
        return(0);
    if ((*qp)->nexts == NULL)
        (*qp)->nexts = temp;
```

return(0);
while ((curr->nexts->nexts!=NULL)&(curr->nexts->ask < time))
curr = curr->nexts;
if (curr->nexts->ask >= time){
temp = curr->nexts;
curr->nexts = temp;
return(0);
}
else {
curr->nexts->nexts = temp;
}

/\* remove the head of the queue \*/
firstevent(timep,temp,qp)
double *timep;
LINK *qp;
int *temp;
{
LINK temp=*qp;
if (*qp==NULL)
*temp = -1:
else{
while (*temp != NULL){
timep = (*qp)->ask;
temp = (*qp)->get;
qp = (*qp)->nexts;
}/
printf("%d,%d
",timep,*temp);
}
free(temp);
}

/\* function to add a new element in a queue in the ith position \*/
addq(val,qp)
STATION *qp;
int val;
{
LINK temp,t1 = (*qp).queue;
(*(ap).aven=(*(ap).aven+master-(*(ap).milestone))(*(ap).length);
if ((*ap).queue==NULL){
if ((*ap).length!=0)printf("Ooooooh\n");
*(ap).sigidle=(*(ap).sigidle-master-(*(ap).milestone);
*(ap).milestone = master; *(ap).length +=;
temp = (LINK)malloc(sizeof(REQUEST));
temp->ask = val*1.0;
temp->get = master;
if (val < 100){
temp->nexts = (*ap).queue; (*ap).queue = temp;
}
else if ((*ap).queue==NULL){
temp->nexts = (*ap).queue; (*ap).queue = temp;
}
else if ((*ap).queue->nexts==NULL){
temp->nexts = (*ap).queue; (*ap).queue = temp;
}
else{
while ((t1->nexts->ask < 100)&((t1)->nexts->nexts!=NULL))
t1 = t1->nexts;
temp->nexts = t1->nexts;
t1->nexts = temp;
}
}

traverse(q)
LINK q;
{
LINK temp=q;
while (temp!=NULL){
printf("%d,%2f->*,(int)(temp->get),temp->ask);
temp=temp->nexts;
printf("\n");
}
}
initstation

STATION stat;
{
    stat.avelen = 0;
    stat.length = 0;
    stat.sigidle = 0;
    stat.milestone = 0;
    stat.queue = NULL;
}

getstatsq

STATION stat;
double *utilp, *avelenp;
{
    *utilp = 1 - stat.sigidle / master;
    *avelenp = stat.avelen / master;
}
```c
#include "declare.c"

/* this routine updates the totals of the first and second moment of the variable */

update(wait, specs)
  double wait;
  specs (*specs);
  {
    (*specs).fmoment = (*specs).fmoment + wait;
    (*specs).smoment = (*specs).smoment + wait * wait;
    (*specs).number = (*specs).number + 1.0;
  }

/* initspecs initializes the totals above */
initspecs(specs)
  specs (*specs);
  {
    (*specs).fmoment = 0.0;
    (*specs).smoment = 0.0;
    (*specs).number = 0.0;
  }

/* this function recovers the mean, variance and number of samples */
getspecs(aspects, pmean, pvar, pnum)
  specs aspects;
  double *pmean, *pvar;
  long *pnum;
  {
    if (aspects.number > 0.0)
      *pmean = aspects.fmoment / aspects.number;
    *pnum = (int)(aspects.number);
    *pvar = sqrt(aspects.smoment / aspects.number - (*pmean) * (*pmean));
    else
      *pnum = 0;
      *pmean = 0.0;
      *pvar = 0.0;
  }

/* enters a point of data in a batch (for mean retrieval time calculation) */
sampleupdate(time)
  float time;
  {
    int i;
    double mean;
    batchmean[batchcount++] = time;
    /* print */
    if (batchcount == MEANS)
      mean = 0.0;
      for (i = 0; i < MEANS; i++)
        mean = mean + batchmean[i];
      fitlra[samplecount++] = mean / (1.0 * MEANS);
    /* print */
    batchcount = 0;
  }

/* recovers the mean retrieval time from the batch means also recovers the variance and the number of batches */
sampleretrieve(pmean, pvar, pnum)
  float *pmean, *pvar;
  long *pnum;
  { double tmean = 0.0, tvar = 0.0;
    int i;
  }
```
arrays.c

*pnum = samplecount;
if (samplecount > SAMPLES) samplecount = SAMPLES;
if (samplecount==0) samplecount = 1;
for (i=0;i<samplecount;i++)
    tmean = tmean + filmra[i];
*pmean = tmean /samplecount;
for (i=0;i<samplecount;i++)
    tvar = tvar + (filmra[i]- *pmean) * (filmra[i]- *pmean);
if (samplecount==1) samplecount = 2;
*pvar = sqrt(tvar /(samplecount-1));

long totalfilms=0,totalblocks=0;

/* routine that keeps track of the total number of films requested
   to calculate the average arrival rate of requests for films
   and also the average number of blocks in a film */
sizeupdate(number)
int number:
{
    totalfilms--; 
    totalblocks -= number:
}

/* retrieve the average number of blocks in a film */
sizeretrieve()
{
    return((totalblocks*.1) /(totalfilms*.1));
}
/ * this file includes the declaration of all the variables and * the structure definitions. */

#include <stdio.h>
#include <math.h>
#define STAT 15
#define MAG 32
#define SEEK 0.016
#define SCAN 0.007
#define OPT 3
#define CONT 4
#define SHARE 8
#define NULL 0
#define TD 4.03
#define TR .039
#define SAMPLES 3000
#define MEANS 50

int numfilms[STAT],blockcount[STAT];

/* link1 is used for the queues */
struct link1{
    float ask; /* struct for station q's as well as */
    float get; /* the dbusy q ask=busy, get=idle */
    struct link1 *nexts;
};

/* specs defines a structure to keep track of random variable statistics */
typedef struct { double fmoment,smoment,number; } specs;

typedef struct link1 REQUEST; /* REQUEST is the record name */
typedef REQUEST *LINK; /* LINK is the type that points to REQ */

/* keeps track of the average length of the queues and */
/* also the utilization statistics */
typedef struct{
    int length;
    double avelen, sigidle, milestone;
    LINK queue;} STATION;

LINK eventhead=NULL;
STATION optq[OPT],dbmsq;
STATION contq[CONT];
double master=0,prot time=0.1,dbms time=0.1,time,magdbtime=0.05,termin;
double SPEED=10.0,BLOCK=24.0/1024,INPUT=120.0,COMPACT=0.1,DIAGSCALE=1.0;
double enterdisk[STAT];
int q_max;

specs ospecs[OPT],bspecs[STAT],aspecs,cspecs,dspecs[CONT],dspecs[STAT];
specs dissubsys;
double statask[STAT],diagask[STAT];
long samplecount=0,batchcount=0;
float imagdf[12][20],films[SAMPLES],batchmean[MEANS];
int imagnum[12][20];
long count[MAG];
STATION busq,magq[MAG]; /* queues for requests to mag discs */
specs mspecs[MAG];
APPENDIX B

APPENDIX B. SAMPLE OUTPUT OF THE PROGRAM
**INPUT PARAMETERS**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BUS SPEED</td>
<td>2.00 megabytes/sec</td>
</tr>
<tr>
<td>BLOCK SIZE</td>
<td>24.00 kilobytes</td>
</tr>
<tr>
<td>NUMBER OF INPUT MODALITIES</td>
<td>3</td>
</tr>
<tr>
<td>RATE OF INFLUX OF NEW FILMS</td>
<td>0.0250 /second</td>
</tr>
<tr>
<td>NUMBER OF MAGNETIC DISCS</td>
<td>32</td>
</tr>
<tr>
<td>NUMBER OF CHANNELS</td>
<td>4</td>
</tr>
<tr>
<td>NUMBER OF OPTICAL JUKEBOXES</td>
<td>3</td>
</tr>
<tr>
<td>DBMS TIME</td>
<td>0.1000 seconds</td>
</tr>
<tr>
<td>COMPACT RATIO</td>
<td>0.100</td>
</tr>
</tbody>
</table>

**OUTPUT PARAMETERS**

<table>
<thead>
<tr>
<th>Station</th>
<th>Mean Waiting Time Per Film</th>
<th>Mean Study Recovery Time</th>
<th>Mean Number of Images Diagnoses Per Hour</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.724 seconds</td>
<td>14.464 seconds</td>
<td>22.707</td>
</tr>
<tr>
<td>1</td>
<td>0.795 seconds</td>
<td>13.475 seconds</td>
<td>19.170</td>
</tr>
<tr>
<td>2</td>
<td>0.967 seconds</td>
<td>3.829 seconds</td>
<td>4.283</td>
</tr>
<tr>
<td>3</td>
<td>0.975 seconds</td>
<td>9.897 seconds</td>
<td>9.955</td>
</tr>
<tr>
<td>4</td>
<td>0.720 seconds</td>
<td>9.008 seconds</td>
<td>13.933</td>
</tr>
<tr>
<td>5</td>
<td>0.558 seconds</td>
<td>4.701 seconds</td>
<td>4.905</td>
</tr>
<tr>
<td>6</td>
<td>0.948 seconds</td>
<td>29.709 seconds</td>
<td>36.243</td>
</tr>
<tr>
<td>7</td>
<td>1.085 seconds</td>
<td>18.229 seconds</td>
<td>14.762</td>
</tr>
<tr>
<td>8</td>
<td>0.981 seconds</td>
<td>11.056 seconds</td>
<td>10.269</td>
</tr>
<tr>
<td>9</td>
<td>1.004 seconds</td>
<td>5.353 seconds</td>
<td>4.907</td>
</tr>
<tr>
<td>10</td>
<td>1.023 seconds</td>
<td>18.662 seconds</td>
<td>18.169</td>
</tr>
<tr>
<td>11</td>
<td>0.766 seconds</td>
<td>4.698 seconds</td>
<td>9.965</td>
</tr>
<tr>
<td>12</td>
<td>0.908 seconds</td>
<td>0.911 seconds</td>
<td>1.000</td>
</tr>
<tr>
<td>13</td>
<td>1.036 seconds</td>
<td>1.041 seconds</td>
<td>1.000</td>
</tr>
<tr>
<td>14</td>
<td>0.791 seconds</td>
<td>0.793 seconds</td>
<td>1.000</td>
</tr>
</tbody>
</table>

OVERALL RETRIEVAL TIME FOR FILMS: MEAN 0.9394 seconds; VARIANCE 0.1982 seconds; NUMBER OF SAMPLES 254

CONFIDENCE INTERVAL: 0.9394 +/- 0.0244

MAGNETIC DISC MEAN WAITING TIME: 0.0261 sec; MEAN NUMBER IN DISC: 0.0552; MEAN UTILIZATION: 0.0417

OPTICAL DISC MEAN WAITING TIME: 0.2745 sec; MEAN NUMBER IN DISC: 0.1238; MEAN UTILIZATION: 0.1161

DISC CONTROL UNIT MEAN WAITING TIME: 0.0113 sec; MEAN NUMBER IN CONT: 0.1198; MEAN UTILIZATION: 0.0730

AVERAGE TIME FOR A BLOCK IN MAG DISK SUBSYS: 0.1067 sec

BUS MEAN WAITING TIME FOR A BLOCK: 0.1041 sec; MEAN NUMBER IN BUS SYSTEM: 2.2117 UTILIZATION: 0.2161229

DBMS MEAN WAITING TIME FOR AN IMAGE: 0.0037 sec; MEAN NUMBER IN DBMS SYSTEM: 0.0816 UTILIZATION: 0.0882936

TRANSMISSION TIME FOR A BLOCK OVER THE BUS: 0.0118 sec

TRANSFER TIME FROM DISC (READ/ WRITE): 0.0108 sec

INFLUX: 1.4958 min: REQUESTS 61.5542 /min; REQUEST/INFLUX: 34.4892
and is indicative of how well the current manual system works. It was also
found that between 2% and 18% of the film folders were lost and up to 23% were already in use when requested. The former are as a result of manually handled information and should not exist with a PACS. The latter should also disappear since copies of each film will always exist in the database even when the folders are 'lent' out for diagnosis.

The Radiological survey contained data about the diagnosis practices by radiologists. Important to note was the fact that 39% of plain films requested were old but only 2% of all contrast films requested belonged to the same category. Averaged, 30% of all images recovered are old films. We found that the radiologist would like to reprocess up to 13% of the images if possible. This reflects on the need for digital image processing for the enhancement of images.

The Admissions survey listed the requests for examination and was a measure of the 'influx' into the department and consequently, the fileroom. Following are the results from this survey. Monday is the most active day for admittances, a majority of these are for plain film examinations. The peak for most exam categories was between 8 am and 9 am. The distribution of the number of times a patient visits the radiological facility is listed in Table 5. This allows us to predict whether or not a patient will return for another examination and could be of use in deciding which level of storage that patient's
records must be kept. However this distribution can be approximated to an
Poisson arrival process (one of the properties of which is the memoryless pro-
PERTY - the next arrival is independent of previous arrivals).

A simulation program was written to evaluate the performance of a
centralized architecture for the PACS. Twelve reading stations were defined
and three acquisition nodes were included. These stations and acquisition
nodes were connected to the centralized system via the bus. The description of
the reading stations is contained in Table 6. The capacity of the mass storage
system was estimated on the basis of 17 GBytes of data being generated per
day and designed to accommodate 32 magnetic disks and three optical
jukeboxes. These figures were arrived at by assuming that the short term
storage stored data for 11 days while it was necessary for the long term storage
to store data for at least seven years. It was calculated that the access time
from the optical disk system is given by

\[ T_{opt} = 4.04 + 0.171 \times (\text{size of data}) \]

where \( T_{opt} \) is in seconds and the data size, in megabytes. The seek time for the
magnetic disk was assumed to be 16 mS, and the scan time, 7 mS. Four chan-
nels were assumed for the 32 disks, i.e. 8 disks per controller. The communica-
tions link was a FIFO scheduled server.

The simulation program was run for several values of the input param-
eters: the bus speed (0.1 MBps - 10 MBps), the block size (1 kB - 1 MB), the
compression ratio (1.0 to 0.05), the influx rate (0.2 images/minute to 160
images/minute), the DBMS time (50 mS to 1 sec), and the diagnosis think time (0.2 to 100 times the original diagnostic think times). A parametric analysis was done by varying a parameter along with the bus speed or the compression ratio while keeping the others constant.

Thus we are presented with several graphs (Figures 30 to 48) which characterize the behaviour of the system under different conditions. It was shown that an effective transmission rate of 2 MBps is sufficient to achieve sub-second delays (Fig. 30). The ideal block size with the architecture defined was found to be 24 megabytes (Section 6.1.3). The system’s performance was almost proportionately dependent on the compression ratio (Fig. 33). Predictably, the compression ratio had a more drastic affect at low bus speeds (Fig. 33). The dependence of the system’s performance on the extent of compression used on the images reaffirms the need for efficient compression algorithms. It turned out that the retrieval time for an image was almost linearly dependent on the DBMS time (Fig. 43). The retrieval time was also largely independent of the influx rate as defined. Finally it was shown that most of the queuing delays experienced by a block was due to internal interference between blocks of the same image (Fig. 46). Consequently, the system can handle 66% more studies per hour without an appreciable difference in the average image retrieval time. With respect to influx variation it was shown that the system can handle up to 30 images/minute without substantial increase in retrieval time (Fig. 39).