ACHIEVING HIGH DEPENDABILITY OF AN ENDOSCOPY RECOMMENDER SYSTEM (ERS)

JAN ZIELIŃSKI AND HENRYK KRAWCZYK

Faculty of Electronics, Telecommunications and Informatics,
Gdansk University of Technology,
Narutowicza 11/12, 80-952 Gdansk, Poland
jan.zielinski@vega.eti.pg.gda.pl, hkrawk@pg.gda.pl

(Received 12 September 2003; revised manuscript received 30 January 2004)

Abstract: The paper presents a strategy for achieving high dependability of a computer-based system devoted to endoscopic examination (ERS). Two levels of replication are used: hardware (3 computers) and database (redundant copies). Archivisation of documents describing patient examinations and films made during such examinations is described. The influence of the used techniques on performance and dependability of the replicated system is estimated.

Keywords: telemedicine, security, dependability, replication

1. Introduction

Computer-based systems designed for medical purposes constantly grow in importance as they lead to significant improvements in the quality of medical services. Such systems must meet very high requirements not only in terms of security, reliability, fault tolerance, but also user friendliness. Thus they tend to be highly complicated and expensive. Saving costs connected with the introduction of a new computer system can also help hospitals to better serve patients, which is of great importance in the environment of Polish health service. Of course, cost-saving cannot be at the expense of quality of the system. The Department of Computer Architecture (KASK) of the Gdansk University of Technology (PG) cooperates with the Department of Endcoscopic Examinations of the Medical Academy of Gdansk (AMG) in developing such systems. The first computer program designed to support administrative tasks connected with examination of patients was created in 1993. The system, called ENDO, was written in the Clipper programming language and worked in MS-DOS and Windows operating systems. After some time it became clear that the system had to be replaced by a new implementation. In 2002 a new system, called an Endoscopy Recommender System (ERS), was built thanks to the grant number 8 T11C 001 17 of National Committee of Science [1]. The main assumption was to improve the quality of the system, its user-friendliness, reliability and security, while maintaining all the functions of the previous system. Additionally, support for acquiring films from endoscopic equipment and their archivisation was introduced in the new system. A very innovative possibility is recommending a diagnosis to the physician on the basis of images captured during examinations. The system can be used remotely (in the Internet or the intranet of AMG) for consultations or teaching. Descriptions of patient examinations can conform to the Minimal Standard Terminology (MST) [2] norm. In the present paper, the basic and improved architectures of the ERS are presented (the latter ensuring greater reliability of the system). To achieve described functionality while maintaining high dependability of the system, replicated architectures have been used on the hardware and software (database) levels.

2. Basic architecture of the ERS

The ERS was built in a classic three-tier architecture [3] using the PHP and Java languages. This choice of a programming platform enables the system to be easily maintained, allows one to reuse stable and secure database engines and web servers for data storage and publishing, and ensures flexibility of the system. The logical architecture of the system is shown in Figure 1.

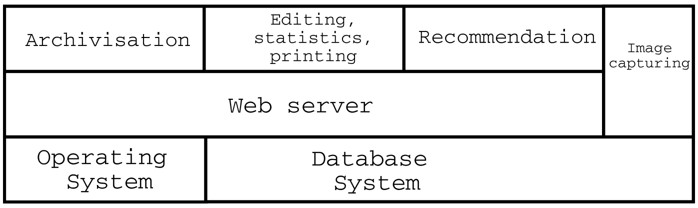

Figure 1. Logical architecture of the ERS

The ERS consists of a few basic modules, developed independently. The presence of some of them in an actual system installation is optional. The main functionality is provided by the data editing module [1]. It allows one to collect data about patients, to describe examinations both in the classical (text) and in the MST form, and to print descriptions and statistics of examinations. This module is written entirely in the PHP language and is thus very flexible; it can be easily modified to conform to the changeable requirements of medical personnel. The interpreter of the PHP language works as a module of an Open Source Apache web server and cooperates with a MySQL database (also Open Source). The film module, written in Java, uses Java Media Framework extensions for on-line capturing of images from the digital output of an endoscopic device. Films can be edited and stored in the film library. Frames that are the most representative for a particular disease can be selected and stored separately in the database. The description of an examination is based on the film and the images [1, 2]. Usually, the archivisation module is required in each installation of the system and its functionality have to be used regularly. It utilizes the possibilities of the PHP language and cooperates with the operating system, the database management system and the CD-Writer software to assure data security. It also manages the library of films made during examinations. The last module is the recommendation module. It uses the PHP and Java languages to run innovative algorithms of diagnosis recommendation based on data mining techniques [4].

3. Replicated architecture of the ERS

The architecture of the system had been redeveloped before it reached its final form. Initially, the system ran on a single computer with the Windows 98 SE operating system. The choice of the operating system was determined by the cooperation of JMF with a frame grabber. As the security of the system was unsatisfactory, its architecture was changed until it reached the form shown in Figure 2. To achieve a sufficient level of system quality three computers are now used. The main application server of the ERS is a computer referred to as *endo*. It is controlled by the Debian GNU Linux system, well-known for its high level of security. The migration from Windows to Linux has been easy due to the ERS's high flexibility. Both web and database servers run on the *endo* computer. The Windows computer (referred to as *gui*) is used as a Graphical User Interface terminal, for storing films from the examinations, and as a print server. A very important function of *gui* is keeping a spare on-line (replicated) copy of the database from *endo*. Besides, a *vega* computer at PG has joined the configuration as an additional spare-copy server. In the development of the ultimate design of the new architecture, the main goal has been to keep the system maximally secure [3]. The connection to the Internet is guarded by firewall set up on *endo*; it serves as a gateway to the Internet for *gui*. To improve security, an IP masquerade with network address translation (NAT) is used. The *vega* computer is also hidden behind a firewall. The connection between *endo* and *vega* utilizes an SSH channel, which

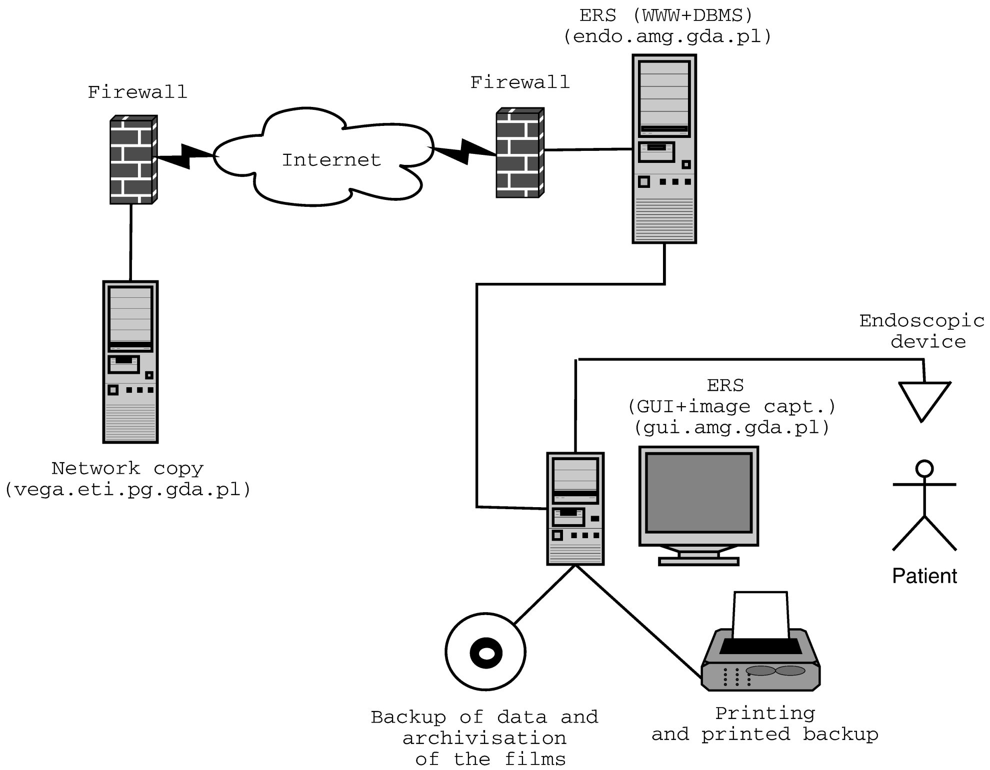

Figure 2. Architecture of the ERS

guarantees authorization and ciphering of communications. The control of privileges of system users is based on groups of users. The users are authorized using a standard login/password mechanism.

4. Archivisation of documents and films

Properly securing data kept by the ERS has allowed it to fulfil dependability and security requirements for medical data [3, 5]. Initially, the system used single-copy backup, created with a specialized module of the DBMS and saved to a CD-RW disc. Such a solution assured that only data from one day of examinations could have been lost. Unfortunately, the network at AMG was not fully secure and a destructive worm attacked the computer. It turned out that the backup subsystem did not work properly (due to an external software failure). We then decided to change the architecture of the system and the backup creation procedures. Having researched of data-securing techniques, we chose to utilize the data replication mechanism supported by a MySQL database. Data replication [6] involves the existence of at least two database servers in a system. One of them is a primary (master) copy, the others are secondary (slave) copies. All queries are directed to the master copy, where they are processed and passed on to the secondary copies. These execute all queries modifying the data (INSERT, UPDATE DELETE), keeping the database consistent. Thus, an on-line backup of the database is created. In the ERS *endo* serves as a primary server. The secondary copy is kept on the *gui* computer. Failure of one of the computers does not prevent the system from working. After a reconfiguration, the application can recover and work on a single machine, at the cost of losing some of its functionality. A whole-day copy of the database on CD-RW is still created after the examinations are finished. The backup creation is monitored and the system administrator is informed by e-mail about any failure. The network-backup mechanism is also used to further improve the system's reliability. The network copy is created automatically at 8.00, 10.00, 12.00, 14.00, 16.00 and 20.00 hours. The whole content of the database is sent via the secure SSH tunnel to the *vega* computer and stored there. This solution can also be considered as replication, with a long-term distance between synchronizations. The creation of the network backup is also monitored and the administrator is informed about unsuccessful attempts to create a copy. The last mean to secure data is a printed copy. It is done once a day, along with the CD backup. The most important informations about each examination performed during the day are printed in a manner that assures maximum paper-saving. All print-outs are kept, so that the system can recover even after a catastrophic loss of electronic data. The large size of films made during examinations makes it impossible to keep them in the DBMS. The archivisation subsystem of the ERS provides a mechanism to manage the library of films. Metadata about all films in the ERS are kept in the database. They contain a description of the film, its recording date, file size, the film's status (*e.g.* high importance) and an identification of the examination the film belongs to. Management of the films is possible due to an interface which allows one to operate on the film list. The list can contain all films, films connected with a particular examination or all un-backed-up films. The interface is shown in Figure 3 (patient names have been erased).

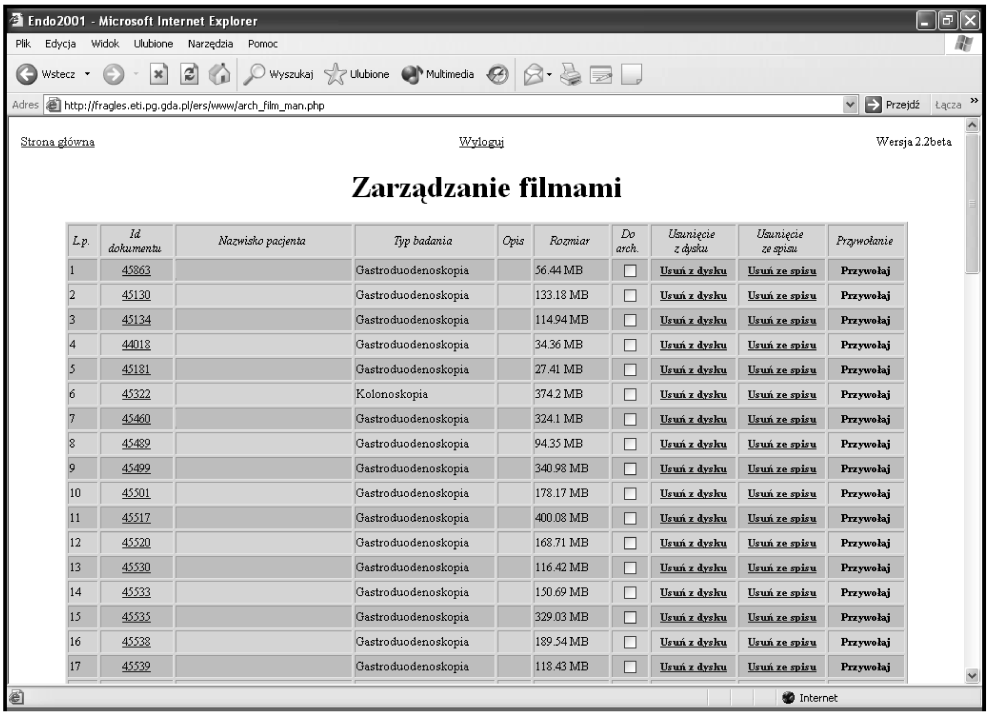

Figure 3. The film management interface

The operator can chose a number of films (the size of which does not exceed the size of the optical disk) and move them to the CD-R. Each CD has a unique identifier, which allows the operator to find the CD easily. If the film is then needed, it can be found in the list of films (thanks to an advanced search mechanism) and copied back to the hard disk of the computer. The system shows the ID of the CD-R that should be inserted into the CD drive and automatically copies the film. The film can be deleted when it is no longer needed.

5. Final remarks

The ERS was developed at the Department of Endoscopic Examinations of AMG and has worked well since October 2002. The use of innovative solutions has allowed us to fulfil all the functionality, flexibility, performance and security requirements of the system. It is hardly possible to represent user satisfaction with numbers, but past the initial phase the number of requests for modifications is very low (1 minor change of functionality per month), which suggests that the system meets the expectations of the personnel. When compared to the previous system, the ERS is described as "significantly better". The system's performance can be judged on the results shown in Table 1. Tests were performed for the most typical functions of the system. At that time, the database contained data of 25 619 patients and 46 073 examinations.

The dependability of the system is a function of data security and reliability of the system. A thorough analysis of security risks which the ERS faces is provided in [3]. Efficiency of the security means used is also analyzed there. If we assume that failures of the computer systems and destruction of backup copies are independent of

Table 1. Results of performance tests

Functionality	Result [s]
Statistics of examinations made by a particular physician in years 1993–2003	9.138
Summary of a number of patients examined in years 1993–2003	5.090
Searching for patients with a particular name given by the pattern (Zie*)	0.544

each other and that the probability of each of the following events: "data loss on the *endo* computer", "data loss on the *vega* computer", "data loss on the *gui* computer", "loss of the CD backup", "loss of the printed backup" equals p, we obtain that the probability of losing all data is p^5, which is significantly better than in the case of a single backup (p^2). So far, experience has shown that p can be estimated to be 0.03.

The future development of the system is connected with the idea of DREAMS, a Distributed, Replicated Architecture for Medical Systems. The idea, introduced in [3], is based on the creation of an universal platform for medical systems based on replication algorithms.

References

[1] Krawczyk H, Knopa R, Kruk S R, Mazurkiewicz A and Zieliński J 2001 *Proc. Int. Conf. on Medical Informatics and Technologies* **2**, Ustron, Poland, pp. 75–84

[2] Smoczyński M, Stankiewicz-Wróblewska J, Marek I, Mittlener S and Dubowik M 2001 *Proc. Int. Conf. on Medical Informatics and Technologies* **2**, Ustron, Poland, pp. 85–93

[3] Zieliński J 2003 *Replication of Information in Web Applications*, MSc Thesis, Gdansk University of Technology, Gdansk, Poland (in Polish)

[4] Mazurkiewicz A and Krawczyk H 2001 *Proc. Int. Conf. on Medical Informatics and Technologies* **2**, Ustron, Poland, pp. 95–100

[5] Zieliński J and Krawczyk H 2003 *Studia Informatica, Silesian University of Technology* **24** 2B(54) 183 (in Polish)

[6] Krawczyk H and Zieliński J 2002 *Studia Informatica, Silesian University of Technology* **23** 2B(49) 185 (in Polish)