Hospital database workload and fault forecasting

Paulo Silva*, Cesar Quintas†, Júlio Duarte*, Manuel Santos‡, José Neves§ , António Abelha§ and José Machado§
*Universidade do Minho, Department of Informatics, Braga, Portugal
†Centro Hospitalar do Porto, Oporto, Portugal
‡Universidade do Minho, ALGORITMI, Guimarães, Portugal
§Universidade do Minho, CCTC, Braga, Portugal

***Abstract*—With the growing importance of hospital information systems, databases became indispensable tools for day-to-day tasks in healthcare units. They store important and confidential information about patients clinical status and about the other hospital services. Thus, they must be permanently available, reliable and at high performance. In many healthcare units, fault tolerant systems are used. They ensure the availability, reliability and disaster recovery of data. However, these mechanisms do not allow the prediction or prevention of faults. In this context, it emerges the necessity of developing a fault forecasting system. The objectives of this paper are monitoring database performance to verify the normal workload for the main database of Centro Hospitalar do Porto and adapt a forecasting model used in medicine into the database context. Based on percentiles it was created a scale to represent the severity of situations. It was observe that the critical workload period is the period between 10:00 am and 12:00 am. Moreover, abnormal situations were detected and it was possible to send alerts and to request assistance.**

I. Introduction

The evolution of technology contributed to the exponential growth of data which needs to be stored for the proper functioning in organizations. Databases are powerful tools where this large amount of data can be stored and managed in a simple way. The database management system (DBMS) coordinates a set of hardware and software components, furthermore, DBMS is also responsible for managing the database users' requests [1].

Today, databases are considered essential for everyday organizations tasks. In healthcare units, due to hospital databases store very important information about the patients' clinical status, administrative information and other relevant information for the healthcare services, they are essential. Therefore, it is crucial to ensure the confidentiality, integrity and availability of data [1][2][3].

However, database availability is a complex feature. There are several problems that influence database availability, such as: the database can be inaccessible due to network problems or a virus; the database can be too slow and therefore not satisfy the user's requests [4].

The unavailability of the healtcare units databases is often related resource limitation faults. These faults occur when the physical resources are not enough to keep the database running. It is essential that the database must continue to operate despite the faults. For this reason, fault tolerant systems are already used. They are responsible to ensure the availability of data even a fault occurs. However, these systems do not allow the early detection of faults [5][6][7].

Therefore, the development of a fault forecasting system which allows taking early actions to solve problems is crucial. Forecasting models have been used in critical areas such as medicine. The objectives of this research is characterized the workload of the one database of Centro Hospitalar do Porto and a development of a model for forecasting and prevention of database faults by adapting an existing forecasting model in medicine (Modified Early Warning Score) [8].

The remaining paper is organized in the following sections: in the second section, we will address the issue of hospitalar information systems, showing the platform where our study was made ; in the third, we will present the forecast model (MEWS); At the fourth, will be described the methodology. In the fifth part, the results will be presented and discussed. At last, the seventh section are presented conclusions.

II. Hospital Information Systems

HIS can be defined as a subsystem hospital with a socio-technological development, which covers all hospital information processing[9]. Its main purpose is to contribute to the quality and efficiency of healthcare. This objective is primarily oriented to the patient after being directed to health professionals as well as the functions of management and administration [9][10]. The HIS also assumes much importance in relation to costs since the sector of communication technologies in healthcare is increasingly important [10]. There are four basic functional processes. This process list begins with the patient's admission and ends in discharge or transfer to another institution. The other categories will serve to support the healthcare with the primary objective of improving quality. The four functional categories are characterized as follows: care; clinical process management; work organization and resource planning; and hospital management [11][12].

EHR can be assumed as a HIS for excellence and has replaced the traditional manual recording in Paper Clinical Process (PCP). EHR may include all hospital areas with a need for registration information. This information can be clinical, administrative and financial [13][14].

A. AIDA

AIDA means Agency for Integration, Diffusion and Archive of Medical Information. It is a platform that consists of a Multi-Agent System (MAS) and it can be considered like the

978-1-4673-1666-8/12/$31.00 ©2012 IEEE

main HIS where it has been working. The AIDA main goal is to overcome difficulties in achieving uniformity of clinical systems, as well as medical and administrative complexity of different hospital information sources [14]. AIDA was created by a group of researchers from the University of Minho and is currently installed in many Portuguese hospitals. It is an electronic platform that holds intelligence electronic employees (agents). This platform promotes a pro-active behaviour in its main functions: communication between hospital heterogeneous systems; storage and management all hospital information; response to requests in time; sending and receiving information from hospital sources like laboratories (medical reports, images, prescriptions). Thus, AIDA enables interoperability between hospital subsystems, assuming a main role where it is installed [15][16]. AIDA has an easy access for your users, allowing the management of clinical information anywhere in the hospital. In addition, the platform enables the sending of messages via phone or e-mail. The same way, AIDA establishes connection with all others systems of patients information: EHR; Administrative Information System (AIS) used by administrative people; Medical Information System (MIS) used in medical record; and Nursing Information System (NIS) used by nurses [17].

B. AIDA-PCE

The AIDA-PCE is an EHR and was implemented in the Centro Hospitalar do Porto. It is working as a subsystem of the main HIS. The AIDA-PCE follows a problem-oriented organization suggested by Lawrence Weed in the 60s. This information organization is known as the Problem Oriented Medical Record (POMR) and it assumes that registration is a production of clinical scientific document. In this kind of organization, the clinical information (annotations, therapeutics and treatments, diagnostics, diaries) should be recorded for specific problem, creating a list of issues organized in a tree structure, where each new problem derives from the main branch [13][18][19]. These problems must be classified as active or inactive, in which active problems are those where the disease is still active or even when medical intervention is required immediately. On the other hand, inactive problems require no urgent action. In this EHR problems assets are monitored and recorded daily using a SOAP (Subjective, Objective, Assessment and Planning) framework. Thus, each record contains the patient's Symptoms, a doctor's Observation, an Analysis of diagnosis and a treatment Plan that the patient is subject to [13][19]. The AIDA-PCE has many common features with PCP but it has a response, which is fast, reliable and safe. The structure of this EHR allows seamless integration with existing HIS by promoting the ubiquity of records between different specialties and services. The ubiquity of the AIDA-PCE allows access to mechanisms for monitoring alarm systems and decision support. The electronic record allows generation of documents and customized reports for specific purposes. It becomes easier to configure interfaces for registration and more. The information contained herein is standardized and uniform [13][15][20]. In the hospitals where

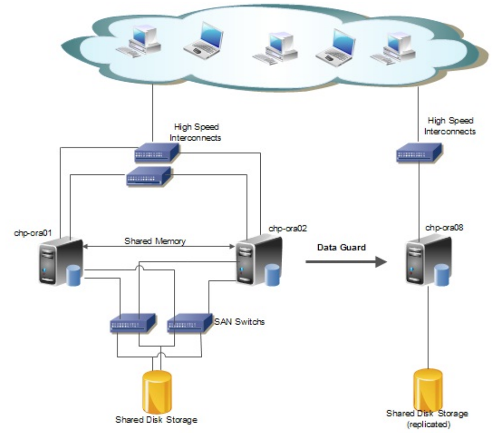


Fig. 1. AIDA and AIDA-PCE Architecture

AIDA and AIDA-PCE was installed, it was made a substantial investment to ensure the availability, reliability and scalability of the system.

C. AIDA and AIDA-PCE faults

As it was mentioned above, AIDA platform and AIDA-PCE are running at real behavior with real cases, because of this reason some problems have emerged. The problems are not responsibility of the AIDA and AIDA-PCE but these ones interfere directly with AIDA and AIDA-PCE operation and, consequently, the quality of the medical record and patients treatment. The main problems are due to communication faults in the network and are related to peak hours using it. Other situations are related to database faults, such as: bad management of datacenter; many users accessed at the same time; breaks on energy. Some of these crashes must to be avoided, but actually errors will always happen [1].

D. AIDA and AIDA-PCE fault tolerance mechanism

Fault tolerance mechanisms are used to ensure avaliability of data but also to load balancing and recovery [1][2].

AIDA and AIDA-PCE architecture, present in Figure 1, is composed by an Oracle Real Application Clusters System (RAC) and a dataguard solution. RAC is used for improving the availability and scalability. This architecture is composed by a shared repository of data which can be accessed through the server/computer that contains a database instance. If a server is down, it is possible access to the database easily by another server [1][6].

A data guard solution consists in one or more standby databases (replicas of the main database). When the main database is unavailable the replica can be used without the need to interrupt operation of the system. It is crucial that the main and the standby databases are synchronized and the access is read-only during recovering [2].

III. MEWS

The concept of Early Warning Score was been introduced in 1997 and the vital signs measured were: Temperature, Respiratory rate, Systolic BP, Heart rate and Neurological. However, in 1999 a new model appeared with the use of two additional variants: urine output, saturation of hemoglobin with oxygen SPO2 [21][22]. This model is the Modified Early Warning Score (MEWS). MEWS assumes that a serious problem of health is often preceded by physiological deterioration.

This model consist in a strict and continuous monitoring of the patient's vital signs. Then, using the decision table, see Table I, the scores are calculated to determine the level of risk of each patient, trying to understand when a serious problem will be occur [8][21] [22].

Normally, if any of the parameters have a score equal to two, the patient must be in observation. In the case of the sum of scores being equal to four, or there being an increase of two values the patient requires urgent medical attention. In a more extreme situation, if a patient has a score higher than four, his life is at risk. The use of this model allows identify the patients at risk and give them priority, improves the monitoring physiological parameters of the patients and thus support the medical decision [8][23].

IV. Methodology

A. Monitoring database performance

Monitoring is the first step for process of identify the source and symptoms of the faults. This is a complex process, however, Oracle provide several tools to help in this process. An example of these tools are the **performance views** which contain useful information for monitoring [24][25].

The most important views related to this paper are the ones which contain information about database component statistics. For global monitoring of the database must be used system-level views (v$sysstat, v$sys_time_model, v$system_event). For more detailed information about the sessions should be used to session level views (v$sesstat,v$session_event, v$sess_time_model [4][25]. It was also used an operating system command (*sar*) to gather relating to the memory and processor utilization [26].

There are others tools for database monitoring. However, the performance views were chosen for two reasons: they make the integration with the business intelligence tool used to do the analysis and presentation of data (Pentaho) easier; and it is easier to define the zone that is to be monitored, thereby reducing the cost associated with monitoring. A Java application was developed to monitoring AIDA-PCE database.

According to the objective of preventing faults related to the resource limitation have been selected the following statistics [2][24][27]:

DB time - This statistic gives information related to database response time. The response time is the period between an initial user request and the return of the results. In Oracle systems, this time is a sum of total time (including CPU time, IO time, Wait time). Therefore it is a good indicator of the workload of the system. Typically, this time increases with the number of simultaneous users or applications, but it also may increase due to large transactions [2][28].

Number of transactions - Database had more or less work according to the transactions which are performed. In Oracle databases, the number of transactions can be obtained by adding up the values of statistics "user commits" and "user rollbacks" due each transaction always ends with a "commit" command and any undo operation as a "rollback" command [2][4][26].

Number of Operations - One transaction may trigger a large set of operations (sum of user calls and recursive calls) depending the query. This means that there may be a large number of transactions and a low number of operations or otherwise. In Oracle databases this information can be obtained by collecting, the "execute count" statistic [26].

Number of sessions - The collection of this statistic is important because each session is associated with a piece of memory, so many simultaneous sessions can cause problems. In Oracle systems, the number of sessions can be obtained by statistic "logons current" [24].

Processor utilization - The processor is one of the most important components, so it is necessary to constantly monitor its utilization by the user processes. Low values of processor utilization may indicate problems at the level of I/O. If the values are too high, it can compromise the functioning of the database. The percentage of utilization can be obtained thought a command of operating system such as **sar -u**[24].

Memory utilization - The memory is a key component to the speed of the database systems. The speed of access to data depends on the place where they are: memory or disk. If the data are in memory then access to them is faster. This statistic is also accessible through the operating system commands such as **sar -r** [4].

Size of redo file - The redo files are used to store information about changes made to the database. These are very important for the recovery of faults. An increase in the size of these files, it indicates a higher number of operations and therefore a higher database load. In Oracle systems, the size (kb) of redo file can be obtained by statistic "redo size" [25].

Amount of I/O requests - The I/O operations are very time consuming often are associated with writing or reading data from memory to disk. An excess of this kind of operations can indicate problems in memory [24].

Amount of redo space requests - Indicates the lack of space to write in the redo buffer, this can lead to delays because it is necessary to write some data to disk to release memory. This can happen due to a poorly sized buffer, or excess entries generated simultaneously [29].

Volume of network traffic - Network is very important for database performance because several database components are connected through network, and all the user requests come from the network. If a volume there is a problem in network, the database can be slow and compromise users' requests [24].

Recursive Calls ratio - Calls to database can be of two types: user calls or recursive calls. When a user request can be resolved through a single SQL query, this is a call. A recursive call occurs when a user request need one query SQL that needs another SQL query. Ideally this ratio should be as low as possible. The recursive calls ratio is the fraction between the recursive calls and total calls (user calls + recursive calls).

TABLE I
MEWS SCORES

MEWS SCORE	3	2	1	0	1	2	3
Temperature (C)		< 35.0	35.1-36.0	36.1-38.0	38.1-38.5	> 38.6	
Heart rate (min -1)		< 40	41-50	51-100	101-110	111-130	> 131
Systolic BP (mmHg)	< 70	71-80	81 - 100	101 - 199		> 200	
Respiratory rate (min -1)		< 8		8-14	15-20	21-29	> 30
SPO2	< 85	85-89	90-93	> 94			
Urine output (ml/kg/h)	Nil	< 0.5					
Neurological		New confusion/agitation		Alert	Reacting to voice	Reacting to pain	Unresponsive

The high value can indicate problems with the design of tables or an excessive amount of triggers running at the same time. This ratio can be calculated by the equation [25][29].

Buffer cache ratio (BC) - This ratio shows the percentage of data that is in memory cache, rather than in the disk. Normally, the BC is very high so it is necessary to pay attention if BC decreases, this may indicate lack of memory problems. BC can be calculate by the the fraction between the number of disk accesses (physical reads) and the number of memory accesses (consistente gets + block gets).

The monitoring program collected values during a month. Using a Business Intelligence tool it is possible analyze the data collected and create graphs that demonstrate the normal behavior of the database. For each statistic, were calculated the upper (using the 75th percentile) and lower (using the 25th percentile) limits.

B. Business Intelligence (BI) tool

BI tools are useful because database monitoring provides a large quantity of data which is difficult to analyze. Thus, BI tools allow the management of the data and present the conclusions of a more clearly way.

The BI tool chosen for this research was Pentaho Community Version because it has all the desired features. The Pentaho Community version provides tools that allow the creation of reports, dashboards/charts, application of data mining, integration techniques and data modeling [30][31]. There is a main application, the bi-server, where it is possible to perform reporting and analysis. In this application, a plug-in which allows the development of dashboards can be added. However, other applications that can be easily integrated with the main server in order to add new features. The applications used in this paper are [31]:

- **CDE** - Allows the user to develop the updateable dashboards, in order to show more clearly the data collected/treated. It was developed and maintained by Webdetails. The front-end is based on HTML, and dashboards can be populated with data coming from a variety of sources, such: SQL querys, xml files, mondrian cubes, spoon transformations [31].
- **Pentaho Data Integration (Spoon)** - Delivers powerful Extraction, Transformation and Loading (ETL) capabilities. This tool is very useful to integrate information from different sources and also to perform some mathematical operations on data. The division of a multi-step transformation reduces the complexity of the SQL query used to obtain the desired information [31].

After represented the normal behavior of the database, a new program was developed for detecting the abnormal situations. Depending on the value of the deviation, for each statistic, is assigned scores to abnormal situations such as is done in MEWS. Two situations can happen (see Table II): if the score is greater than 0 and less than 3, a visual warning will be issued on the dashboard; if the score is greater than 3, warnings will be sent via email to the database administrator, allowing he to take speedy action to prevent the occurrence of a fault in the database.

This program is upgradeable, as new limits calculated at the end of the day based on new measurements, because it is possible that exists an increase of load and the database remains operational.

V. RESULTS

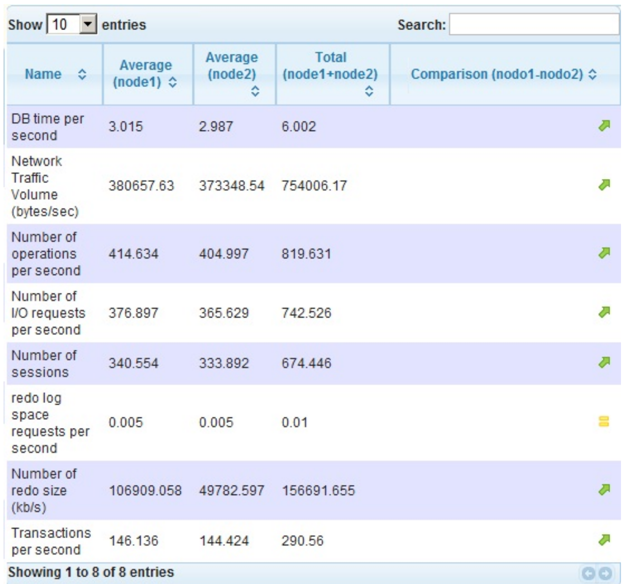
As a first step will be examined workload of the database. In Figure 2, is represented a table that contain the mean values of each statistical measure for each node. It is possible observe that node 1 (chp-ora01) has more workload than node2 (chp-ora02). Only in the "redo log space requests per second", the mean of node1 is not higher to the node 2.

Overall, the AIDA database has an average 674 sessions and about 290 transactions per second, which shows that is a database with a high workload. The high value of use of the network, shown that it is also a very important characteristic for evaluating the performance of the database. Concerning the use of two main components, memory and processor, through the Figure 3 is possible to observe that the average is practically the same in both nodes. Memory presents high values unlike the processor. This may be due to take into account only the processor utilization by user processes. To identify the most critical points of the day, graphs and tables of Figures 4, 5 were constructed. In Figure 4, it is possible identify that the period between 11:00 and 12:00 is the most critical period in node1. In this period, there is three important peaks: DB time, percentage of processor usage and number of sessions.

Figure 5, presents the workload peaks for node two (chp-ora02). In this node, the distribution of peaks is more diverse.

TABLE II
SEVERITY SCORES

SCORE	0	1	2	3
Value	$< p75$	$> p75 < p80$	$> p80 < p90$	$> p90$
Severity	Normal	Low severity	Severe	Critical



Show 10 entries Search:

Name	Average (node1)	Average (node2)	Total (node1+node2)	Comparison (nodo1-nodo2)
DB time per second	3.015	2.987	6.002	
Network Traffic Volume (bytes/sec)	380657.63	373348.54	754006.17	
Number of operations per second	414.634	404.997	819.631	
Number of I/O requests per second	376.897	365.629	742.526	
Number of sessions	340.554	333.892	674.446	
redo log space requests per second	0.005	0.005	0.01	
Number of redo size (kb/s)	106909.058	49782.597	156691.655	
Transactions per second	146.136	144.424	290.56	

Showing 1 to 8 of 8 entries

Fig. 2. AIDA Workload

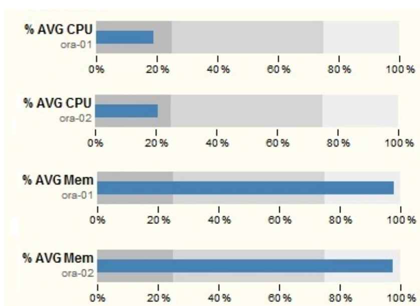


Fig. 3. AIDA Workload

However, in the period between 10:00 and 11:00, there is three important peaks: Network traffic, percentage de processor usage and recursive call ratio. In general one can consider the period from 10:00 to 12:00, as the period most favorable for the occurrence of faults. For each statistic of each node limits were calculated by using percentiles. The Figure 6, shows the normal behavior of the variable “number operations per second” throughout the day. For this example only the period between 08:00 and 19:00 is considered because in this range there is the highest number operations per second. It is possible to see that the database has more activity in the morning, due to the higher number of users connected. It is expected that 50% of the collected data are present between the two limits defined by the 25th and 75th percentiles. In order to identify the abnormal situations scores were assigned

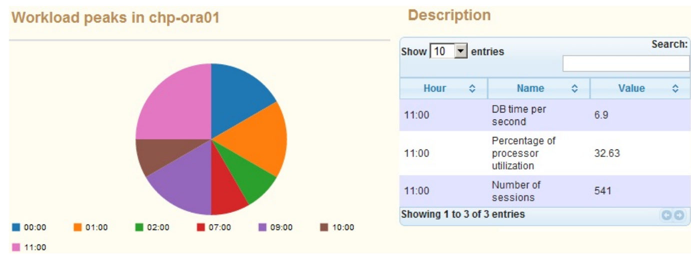


Fig. 4. Workload peaks chp-ora01 node

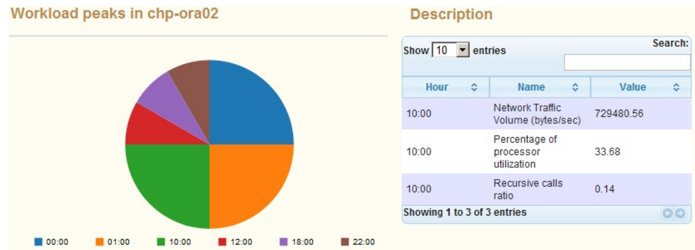


Fig. 5. Workload peaks chp-ora02 node

to percentiles. The data above the percentile 75 is considered a low severity situation, above 80 a grave situation and above 90 a critical situation (see Table II). Figure 7 presents abnormal situations that occur along one day. In the period between 12:00 and 14:00 there is two abnormal situations. The first is a severe situation and the other is a critical situation, according to Table II. However, these situations do not cause database fault. For this reason it is necessary to update the limits. Limits are update in the end of day taking account of all measured values which do not cause fault.

VI. CONCLUSIONS

A model was proposed to characterize and represent the normal workload for AIDA database. It was observed that the node1 of the database is more used than node2. Moreover, workload peaks were identified and it was possible to observe that the critical period is the period between 10:00 am and 12:00 am.

With percentiles, it was possible to determine the limits that characterize the abnormal situations. A decision table, with percentiles and scores was created. Therefore, the methodology of MEWS were adapted to database reality through an upgradeable and learning forecasting fault system. The Pentaho Community allowed to perform the data analysis and development of dashboards which makes it easy the interpretation of results these results.

Due to the heterogeneity of the database workload it was found that the limits will be updated every day.

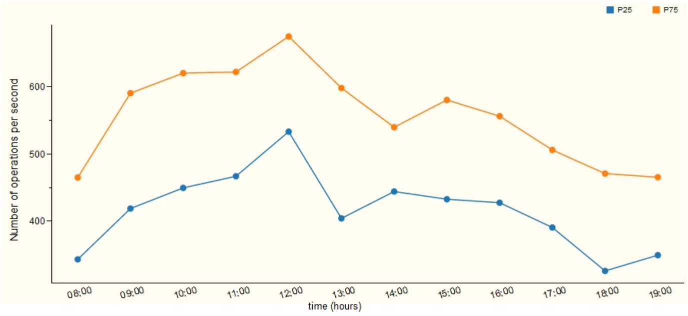


Fig. 6. Limits of number of operations per second in chp-ora01

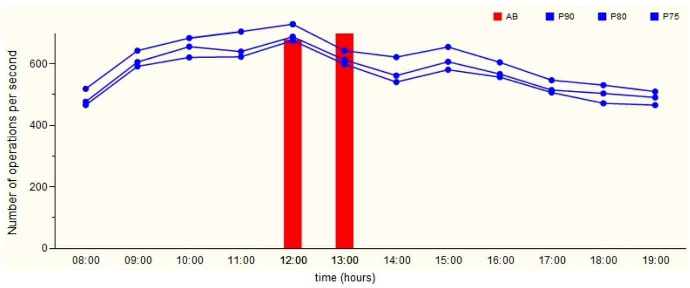


Fig. 7. Limits and abnormal situations in chp-ora01

ACKNOWLEDGMENT

This work is financed with the support of the Portuguese Foundation for Science and Technology (FCT), with the grant SFRH/BD/70549/2010 and within project PEst-OE/EEI/UI0752/2011.

REFERENCES

[1] S. Drake, W. Hu, D. McInnis, M. Skold, A. Srivastava, L. Thalmann, M. Tikkanen, O. y. Torbjornsen, and A. Wolski, "Architecture of Highly Available Databases," in *Service Availability*, ser. Lecture Notes in Computer Science, M. Malek, M. Reitenspie, and J. Kaiser, Eds. Springer Berlin / Heidelberg, 2005, vol. 3335, pp. 1–16. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-30225-4_1

[2] R. Godinho, "Availability, Reliability and Scalability in Database Architecture," MSc Thesis, Universidade do Minho, 2011.

[3] A. Rodrigues, *Oracle 10g e 9i:Essentials For Professionals (in portuguese)*. Lisboa: FCA, 2005.

[4] R. Schumacher, *Oracle Performance Troubleshooting With Dictionary Internals SQL & Tuning Scripts*. Kittrell: Rampant TechPress, 2003.

[5] O. Yi, "Highly Available Database Systems," in *Integrated information systems and databases seminar*, 2006.

[6] M. Cyran, P. Lane, and J. Polk, *Oracle Database Concepts, 10g Release 2 (10.2)*. Oracle, 2005. [Online]. Available: http://docs.oracle.com/cd/B19306_01/server.102/b14220.pdf

[7] W. Hodak, S. Kumar, and A. Ray, "Oracle Database 11g High Availability," Oracle, Tech. Rep., 2007.

[8] C. P. Subbe, M. Kruger, P. Rutherford, and L. Gemmel, "Validation of a modified Early Warning Score in medical admissions," *QJM*, vol. 94, no. 10, pp. 521–526, 2001. [Online]. Available: http://qjmed.oxfordjournals.org/content/94/10/521.abstract

[9] R. Haux, A. Winter, E. Ammenwerth, and B. Brigl, *Strategic Information Management in Hospitals: An Introduction to Hospital Information Systems*. Springer-Verlag, 2004.

[10] J. Machado, A. Abelha, P. Novais, J. Neves, and J. a. Neves, "Quality of service in healthcare units," *Int. J. Computer Aided Engineering and Technology*, vol. 2, no. 4, pp. 436–439, 2010.

[11] J. Duarte, C. F. Portela, A. Abelha, J. Machado, and M. F. Santos, "Electronic health record in dermatology service," in *Communications in Computer and Information Science, 221 CCIS (PART 3)*, 2011.

[12] H. AmmenwerthE., BuchauerA., "Arequirementsindexfor information processing in hospitals," *Methods of Information in Medicine*, vol. 41, pp. 282–288, 2002.

[13] M. Miranda, G. Pontes, P. Gonçalves, H. Peixoto, M. Santos, A. Abelha, and J. Machado, "Modelling intelligent behaviours in multi-agent based hl7 services," in *Studies in Computational Intelligence*, Springer, Ed., vol. 317, 2010.

[14] E. Coiera, *Guide to Health Informatics (2nd ed.)*, 2nd ed. London: Hodder Arnold, 2003.

[15] J. Duarte, J. Neves, A. Cabral, M. Gomes, V. Marques, M. F. Santos, A. Abelha, and J. Machado, "Towards intelligent drug electronic prescription," in *European Simulation and Modelling Conference*, Guimarães, Portugal, 2011.

[16] F. Portela, M. Vilas-Boas, M. F. Santos, A. Abelha, J. Machado, A. Cabral, and I. Aragao, "Electronic health records in the emergency room," in *ACIS-ICIS 2010*, pp. 195–200.

[17] V. Slee, D. Slee, and J. Schmidt, *The Endangered Medical Record: Ensuring Its Integrity in the Age of Informatics*. Saint Paul, Minnesota: Tringa Press, 2000.

[18] M. Miranda, J. Duarte, A. Abelha, J. Machado, and J. Neves, "Interoperability and healthcare," in *European Simulation and Modelling Conference*, Leicester, UK, 2009.

[19] K. Hyrinen, K. Saranto, and P. Nyknen, "Definition, structure, content, use and impacts of electronic health records: A review of the research literature," *International Journal of Medical Informatics*, vol. 77, pp. 291–304, 2008.

[20] C. Bossen, "Evaluation of a computerized problem-oriented medical record in a hospital department: Does it support daily clinical practice?" *International Journal of Medical Informatics*, vol. 77, pp. 592–600, 2007.

[21] A. Albino and V. Jacinto, "Implementation of early warning score - EWS," Centro Hospitalar do Barlavento Algarvio, EPE, Portimão (in portuguese), Tech. Rep., 2009.

[22] J. Gardner-Thorpe, N. Love, J. Wrightson, S. Walsh, and N. Keeling, "The value of Modified Early Warning Score (MEWS) in surgical in-patients: a prospective observational study." *Annals of the Royal College of Surgeons of England*, vol. 88, no. 6, pp. 571–5, Oct. 2006.

[23] G. Devaney and W. Lead, "Guideline for the use of the modified early warning score (MEWS)," Outer North East London Community Services, Tech. Rep., 2011.

[24] I. Chan, *Oracle Database Performance Tuning Guide, 10g Release 2 (10.2)*. Oracle, 2008.

[25] K. Rich, *Oracle Database Reference, 10g Release 2 (10.2)*. Oracle, 2009.

[26] C. Shallahamer, *Forecasting Oracle Performance*. Berkeley: Apress, 2007.

[27] L. Ramos, "Performance Analysis of a Database Caching System In a Grid Environment," MSc Thesis (in portuguese), FEUP, 2007.

[28] K. Dias, M. Ramacher, U. Shaft, V. Venkataramani, and G. Wood, "Automatic performance diagnosis and tuning in Oracle," in *Proceedings of the 2005 CIDR Conf*, 2005. [Online]. Available: http://crmondemand.oracle.com/technetwork/database/focus-areas/manageability/cidr-addm-134903.pdf

[29] Hoopoes, "Oracle Database Tuning Statistics," 2007. [Online]. Available: http://www.hoopoes.com/cs/oracle_tune.shtml

[30] M. Tereso and J. Bernardino, "Open source business intelligence tools for SMEs," in *Information Systems and Technologies (CISTI), 2011 6th Iberian Conference on*, 2011, pp. 1–4.

[31] Pentaho, "Welcome to the Pentaho Community." [Online]. Available: http://community.pentaho.com/