An Open Source Health Care Management System for Pakistan

Nadeem G. Chaudhry  
chnadeem@ciitlahore.edu.pk

Saira Ilyas  
sairailyas82@yahoo.com

Asif Shahzad  
asifsh7@yahoo.com

Adnan Saleem  
aden2007@gmail.com

Mian Rashid  
oficerpk@yahoo.com

Tahir A. Chaudhry  
tahirakramch@gmail.com

COMSATS Institute of Information Technology, Defence Road, Off Raiwind Road, Lahore, Pakistan

Paper ID- ICOST2006-10

Abstract

Healthcare is of great importance to any country, and especially so for Pakistan. Unfortunately, majority of health care domains in Pakistan are operating at a sub-optimal level, as the use of Information Systems is almost non-existent. There is an utmost needs to improve the current healthcare services by introducing Information Technology. COMSATS Institute of Information Technology, Lahore, is designing and developing a comprehensive, open source, Healthcare Management System in collaboration with Allama Iqbal Medical College & Jinnah Hospital Lahore, and Lahore General Hospital. In this paper, we propose COMSATS OpenSource Healthcare Management System (COS-HMS) an international standards compliant healthcare information system including its architecture and technology specifications.

Keywords: Healthcare Information System (HIS), LINUX, MySQL, JSF, JSP, J2EE, ICD-10, HL7, DICOM, LOINC.

1. Introduction

In the absence of proper Healthcare Management System, most of the health care amenities, particularly in government sector, are lacking the best possible outcome. The development of an exhaustive healthcare system involves complex issues like finance, performance, security, scalability, and adherence to standards. A number of efforts have progressed to address such problems for instance vistA and care2x. Yet, it is not realistic for the healthcare institutions in Pakistan to find an existing open source information system and adapt it to its local needs, as these systems do not fulfill our local requirements.

Our goal is to build a healthcare application covering the necessities of the hospitals in Pakistan and allowing them to communicate clinical data internationally. We have used an open source technology described in section 5. Our application incorporates Health Level 7 that provides standards for the exchange of data among worldwide healthcare computing applications. Several other standards are also used in conjunction with HL7 such as ICD-10 disease codes, ICD-10 procedure codes, LOINC, DICOM, ISO 3166-1 etc; details are discussed in section 6.

2. Present Healthcare Information Systems in Pakistan

At present, most, if not all, of the government healthcare organizations do not have any HIS or it is of limited scope. The situation is a little better in private sector. Some, institutions, such as, The Aga Khan University Hospital, Karachi, and Shaukat Khanum Cancer Memorial Hospital, Lahore, have developed their own HIS.

One of the best examples is of HOPe (Hospital Operations Program) at Shaukat Khanum Cancer Memorial Hospital. A team of about 20 staff members developed the system over a period of 5 years. Its first module was implemented about 4 years ago, since then it has been in constant use. The team has kept on adding new modules since the initial implementation. It uses international standards in most of the areas. For example ICD9 coding scheme is used for disease encoding. The system is fully developed in Oracle 8i/9i at backend and Developer 6/6i for the front end. This approach is appropriate from software engineering point of view, as the integration problems are reduced to a minimum and development is relatively easy.

Another important aspect that should be taken into account is the cost. Purchasing Oracle license for
the first time, costs approximately 5 million rupees. Subsequent yearly renewal cost is approximately 1 million rupees which is not feasible for every organization to afford.

Therefore, we have used an open source technology: Linux Operating System, JBOSS Application Server and MySQL Database Server.

3. Open Source Healthcare Information Systems

All possible open source options were reviewed. The first was Veterans Health Information Systems and Technology Architecture (VistA) which was originally developed and maintained by the U.S. Department of Veterans Affairs (VA), based on the systems software architecture and implementation methodology developed by the U.S. Public Health Service jointly with the National Bureau of Standards. It was designed to provide a high-quality medical care environment for the country's military veterans. VistA has a proven record of supporting a large variety of clinical settings and medical delivery systems. VistA is widely believed to be the largest integrated HIS in the world. VistA is in production today at hundreds of healthcare facilities across the country from small outpatient clinics to large medical centers. The software is currently used by the Indian Health Service and a number of other healthcare organizations around the world.

An important problem with VistA is MUMPS, the language used to implement it. Massachusetts General Hospital Utility Multi-Programming System (MUMPS) is a programming language that came into being by the efforts of Dr. Octo G. Barnett and his team around 1967. For reasons, beyond the scope of this paper, MUMPS never became as popular and widely used as other languages of that era, such as C. Very few professional developers have the experience in developing applications in MUMPS. Moreover, not many people are interested in learning MUMPS that means that any modifications/enhancements to VistA would be extremely difficult.

Another open source, generic, multi-language HIS is Care2x. The project was started in May 2002 with the release of the first beta version of Care2x by a nurse who was dissatisfied with the HIS in the hospital where she was working. Since then, the development team has grown to over 100 members from over 20 countries. Care2x is a web-based HIS that is built upon other open-source projects: the Apache web server from the Apache Foundation, the scripting language PHP and the relational database management system mySQL. Several source code branches exist that try to integrate the option to choose from other RDBMS like Oracle andpostgresql. Care2x is a very feature rich HIS, that is fully configurable for any clinical structure. It is built upon different modules which include e.g. in- and out-patient administration, admission, pharmacy, radiology (including DICOM image uploads), laboratories, ambulatories, nursing, medocs, DRG (Diagnosis Related Group), etc. Online help is available for some clinical paths.

There are several minor problems with Care2x such as lack of structured documentation, unorganized code and poor naming conventions. Most of these problems can be fixed by modifications but there are two serious issues with Care2x. Firstly, it does not encapsulate the broad scoped, widely accepted healthcare standard. Broad scoped standard in this context means a standard embracing almost the entire domain in the healthcare arena (as hl7). For example DICOM and ICD 9 are implemented for medical images communications and disease codes respectively but admission, discharge and transfer of patients (ADT) domain is not standardized and so is the case with several other domains. Secondly, it uses PHP. There is great controversy about PHP’s capabilities in large scale enterprise applications including scalability, security and reliability. So, we decided to use Java which has all of the above mentioned abilities needed for very large applications and has been in use for longer than PHP.

4. Our Proposed Solution

After analyzing the above-mentioned systems, we reached the conclusion that the best course of action is to design and develop a totally new system from scratch.

In our solution, we have separated the whole processing into layers so that at each layer the best technologies and tools can be applied to get maximum performance. Based on our research we believe that our chosen technologies will deliver excellent results.

5. System Architecture

Web based applications are basically two tier applications consisting of a web server tier and a database server tier. We have split the web server tier into two, web tier and an application server tier. To implement this layered architecture, we selected Java J2EE 1.4 (Java 2 Enterprise Edition) that has a huge installed base in enterprise industry. At the web tier, it uses Servlets, Java Server Pages (JSPs) and newly introduced rich UI framework known as Java Server Faces (JSF) technology. At the Application server tier
it uses EJBs (Enterprise Java Beans). The detailed architectural diagram is show below.

5.1 Web Server Tier

The web tier receives the user request, generates contents and responds to users. The application tier manages the user sessions and performs all the business processing required by web tier to generate content. It also interacts with the database. Our web server contains JSF components that enhance the user interaction experience over a web interface making it a Rich Internet Application (RIA).

5.1.2 JSF and Its RIA Capabilities. JSF technology simplifies building the user interface by providing a higher-level framework for working with web application, representing the page as event-aware components rather than raw markup. The promise of Java Server Faces is to bring rapid user-interface development to server-side. It is particularly suited by design, for use with applications based on the MVC (Model View Controller) architecture [7].

5.2 Application Server Tier

Considering our data intensive application and reviewing J2EE architecture, we decided to use session
beans to perform the business logic processing. We have split the load on different servers such that the web contents generation processing is performed by web container and business logic processing by the EJB container. Distributing them on different servers improves performance and fault tolerance when load is high.

5.3 Database Tier

Keeping in view the complexity of business logic processing, it is was decided to process database specific logic entirely on the database server. Stored procedures of the database server are used that eventually provide optimal performance for data-centric application.

We selected MySQL RDBMS, which provides excellent performance. Its current version (5.0) also supports views, stored procedures and triggers; features that were absent in the previous versions. MySQL Cluster is designed to provide 99.999% availability using distributed node architecture with no single point of failure. The system consists of multiple nodes that can be distributed across machines and regions to ensure continuous availability in case of node or network failure. MySQL Cluster uses a storage engine, consisting of a set of storage nodes to store data, which can be accessed using standard SQL with MySQL Server. With these added features MySQL can now be compared with proprietary database engines. As a result several significant applications with very large datasets have been and are being developed using MySQL. Following are some examples of such applications. Cox Communications is the fourth largest cable-television provider in the United States, serving approximately 6.3 million customers. To maintain optimum performance and customer-service levels, Cox has developed a huge data warehousing application. At the heart of this business-critical system is a 2.5 terabyte MySQL database. Given the size and complexity of the application, the MySQL database is under constant pressure to perform. Already the database includes more than 3,600 tables and two billion rows of data. Overall, it handles approximately 4 million inserts every two hours.

Friendster (an American Company) has over 17 million registered users and is handling 1.5 billion MySQL queries per day. According to Tim Denike, Senior UNIX Administrator at Friendster, open source was the only solution to conquer the site's massive growth problems. "Open source tools allowed us to scale a massively complex application across a system that required very little administrative overhead compared to those of other companies". They have Dual AMD Opteron (26 Servers) with database size of 21 TB.

Alamos National Laboratories built its “library without walls” application using the MySQL open source database to create a robust, secure distributed database containing more than 55 million scientific journal articles including more than 7 terabytes of data [8].

5.3.1 eWeek Database Benchmark Test Of the five databases tested, only MySQL and Oracle9i were able to run our Nile application as originally written for 8 hours without problems. The Oracle and MySQL drivers had the best combination of a complete JDBC feature set and stability. MySQL was the easiest to tune, and Oracle9i was the most difficult because it has so many separate memory caches that can be adjusted [9].
The following table shows the actual results of the test measuring pages per second at different thresholds of users taken from above graph.

<table>
<thead>
<tr>
<th>Users</th>
<th>MySQL</th>
<th>Oracle</th>
<th>SQL Server</th>
<th>DB2</th>
<th>Sybase ASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>94.735</td>
<td>100.700</td>
<td>99.724</td>
<td>102.382</td>
<td>95.753</td>
</tr>
<tr>
<td>200</td>
<td>186.594</td>
<td>199.653</td>
<td>181.406</td>
<td>203.859</td>
<td>191.071</td>
</tr>
<tr>
<td>300</td>
<td>270.824</td>
<td>293.318</td>
<td>208.900</td>
<td>302.783</td>
<td>280.535</td>
</tr>
<tr>
<td>400</td>
<td>361.812</td>
<td>384.671</td>
<td>205.335</td>
<td>398.688</td>
<td>370.335</td>
</tr>
<tr>
<td>500</td>
<td>443.559</td>
<td>476.241</td>
<td>206.676</td>
<td>484.065</td>
<td>448.888</td>
</tr>
<tr>
<td>600</td>
<td>523.524</td>
<td>544.665</td>
<td>204.429</td>
<td>322.888</td>
<td>475.829</td>
</tr>
<tr>
<td>700</td>
<td>578.082</td>
<td>594.806</td>
<td>207.359</td>
<td>237.747</td>
<td>471.294</td>
</tr>
<tr>
<td>800</td>
<td>599.612</td>
<td>615.624</td>
<td>200.518</td>
<td>215.476</td>
<td>471.118</td>
</tr>
<tr>
<td>900</td>
<td>601.788</td>
<td>631.388</td>
<td>203.829</td>
<td>207.247</td>
<td>478.718</td>
</tr>
<tr>
<td>1000</td>
<td>590.488</td>
<td>624.386</td>
<td>203.870</td>
<td>204.682</td>
<td>475.934</td>
</tr>
</tbody>
</table>

6. Standards

6.1 Why do need standards in healthcare informatics?

Healthcare applications when built without standards cannot communicate well with each other. This lack of standard hinders collaboration as organizations exchange information with national and international healthcare related organizations. For example, if two applications identify the same disease with a different name, then it is almost impractical to transfer patient clinical data among them. Similarly, if the laboratories identify tests by means of their own internal code values, receiving healthcare informatics systems cannot fully "understand" the results they receive unless they either adopt the producer's laboratory codes (which is impossible if they receive results from multiple source laboratories, e.g.; the
hospital lab, the local commercial lab, and a nursing home lab), or invest in the work to map each laboratory's code system to their internal code system. Indeed, a framework must be developed for minimizing incompatibility and maximizing the useful exchange of information between systems.

6.2 Standards Adopted in COS-HMS

6.2.1 Health Level 7 (HL-7) Health Level Seven is one of several American National Standards Institute (ANSI) -Accredited Standards Developing Organizations (SDOs) operating in the healthcare arena. Most SDOs produce standards (sometimes called specifications or protocols) for a particular healthcare domain such as pharmacy, medical devices, imaging or insurance (claims processing) transactions. HL7 is for complete clinical and administrative data.

The mission of HL7 is to provide a comprehensive framework and related standards for the exchange, integration, storage, and retrieval of health information. It supports clinical practices, the management, delivery and evaluation of health services. Basically, it is an Application Protocol but after years of implementing HL7 v2.x, HL7 standard developers realized it was virtually impossible to develop a comprehensive standard without including other levels of the International Organization for Standardization (ISO) Open Systems Interconnect (OSI) Model. The standard provides the layout of messages that are exchanged between two or more applications based upon a particular trigger event. It specifies which data elements are to be sent, the data type and suggested length of each element. Briefly stated, HL-7 defines what is to be shared, How, When and by whom. More than 93% of all organizations in the US with Health IT systems use HL7.

6.2.1.1 Primary Reasons to adopt HL7

The benefits of using HL7 standard are:
- Open – it is a platform and technology independent standard
- Widely accepted standard, having affiliates across 27 countries, that proves a well-defined framework to be used as a starting point (saving time and other resources, not “re-inventing the wheel”, reducing the custom programming required.
- Enables information exchange between computer applications developed by different (often competing) vendors.
- Reduces paper work, improves decision-support and allows you to integrate health information over time and across health service delivery systems
- Reduces costs and project delivery time in:
  - Analysis
  - Development
  - Implementation
- Implementation using XML document structure
- Ultimately, facilitates better patient care

6.2.1.2 External Coding Schemes and Standards used with HL7. HL7 has defined its internal coding schemes. For example, Marital Status Codes, Gender Codes etc. It also allows many of the external coding schemes and standards relating to different areas for example drugs, diseases, laboratory observations, clinical images etc to be used within the messages. Brief overview of some of the major standards incorporated in our application is given below.

6.2.2 ICD-10-Clinical Modification (ICD-10-CM)

The ICD-10-CM is a morbidity classification published by the United States for classifying diagnoses and reason for visits in all health care settings. The ICD-10-CM is based on the ICD-10, the statistical classification of disease published by the World Health Organization (WHO). The National Center for Health Statistics (NCHS), one of the Centers for Disease Control and Prevention (CDC), an agency within the United States Department of Health and Human Services (DHHS), has developed the guidelines for coding and reporting using the International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM). It is designed to promote international comparability in the collection, processing, classification, and presentation of morbidity and mortality statistics. It is also used for the extensive analysis of the general health situation of population groups and monitoring of the incidence and prevalence of diseases and other health problems.[11]

6.2.3 ICD-10 Procedure Coding System (ICD-10-PCS)

The International Classification of Diseases 10th Revision Procedure Classification System (ICD-10-PCS) has been developed by 3M Health Information System under contract awarded and funded by U.S. Centers for Medicare and Medicaid Services (CMS). ICD-10-PCS has a multi-axial seven-character alphanumeric code structure that provides a unique code for all substantially different procedures, and particularly allows new procedures to be easily incorporated as new codes. Each character contains up to 34 possible values. Each value represents a specific
option for the general character definition (e.g., stomach is one of the values for the body part character). The ten digits 0-9 and the 24 letters A H, J-N and P-Z may be used in each character (the complete coding structure is beyond the scope of this paper). This system has evolved during its development based on extensive input from many segments of the healthcare industry. The multi-axial structure of the system, combined with its detailed definition of terminology, permit a precise specification of procedures for use in health services research, epidemiology, statistical analysis and administrative areas [12].

6.2.4 DICOM (Digital Imaging and Communications in Medicine). DICOM was developed by the National Electrical Manufacturers Association (NEMA) in conjunction with the American College of Radiology (ACR). This standard comprises of image formats for not only radiology but for all of medicine, with additional specification for messaging and communication between imaging machines, with PACS, and with hospital information systems. It specifies a network protocol utilizing TCP/IP, defines the operation of Service Classes beyond the simple transfer of data. DICOM covers the following medical imaging:

- CT, MRI
- Nuclear Medicine & PET
- Digital X-Ray, Mammography, Ultrasound
- Electrocardiograms, Endoscopy

DICOM will soon be used by every medical profession that utilizes images within the healthcare industry across the globe [13].

6.2.5 LOINC (Logical Observation Identifier Names and Codes). The Logical Observation Identifiers Names and Codes [14] data base provides a standard set of universal names and codes for identifying individual laboratory results (e.g. Hemoglobin, Serum Sodium concentration), clinical observations, (e.g. PR-interval, Cardiac echo left ventricular diameter, Chest x-ray impression). It is to be used in the context of existing ASTM E1238, HL7, CEN TC251, and DICOM observation report messages employed in the various sub-domains of healthcare informatics. The ultimate goal is that these "universal" identifiers, when used in the context of the messaging standards, will allow the exchange of laboratory and clinical test results between heterogeneous computing environments.

7. Conclusion

This paper is more of an invitation to discussions/suggestions/ideas for improvement than a groundbreaking research into a new field. Our aim is to develop this system primarily for the health care industry of Pakistan, but being an Open Source system it would be free for any one to use it. We hope that it would be able to join ranks with other popular Open Source products and be of great value to people in health care industry all over the world particularly in Pakistan.

8. References

[4] Largest Open Source Database used by many renowned leading organizations http://www.mysql.com