International Journal of Engineering and Applied Sciences (IJEAS)
ISSN: 2394-3661, Volume-3, Issue-1, January 2016

Building a Distributed Decision Support System Architecture for National Health Care

Sinyinda Muwanei, Douglas Kunda, Gilbert Sibajene

Abstract— **There are several benefits for developing a distributed decision support system for national health care. Decision makers especially in the health sector of the nation will make timely, informative and accurate decisions as both detailed and aggregated data from different health centres or facilities in the nation will be at their disposal together with various tools for analysis of data. The paper describes the architecture and prototype of distributed decision support system for national health care.**

Index Terms— **Distributed Systems , Decision Support, Datawarehouse, Middleware,ETL**

I. Introduction

The healthcare sector is one of the largest sectors in every nation, catering to thousands of people on a regular basis. Patients receive medical treatment from health facilities. Nowadays, in most hospitals the health data is distributed across several heterogeneous and autonomous information systems whose interconnection is difficult to achieve. Integration of such system may bring about many advantages [1] .The accumulation of this information is beneficial both for subsequent treatment of same patients as well as statistical analysis by health management authorities of nation. This accumulated information hence needs to be securely and accurately stored for future analysis. This information at various health facilities is stored in online transactional processing systems. These transaction processing systems are not meant for analysis purposes. Therefore various what if scenario queries are not adequately catered for. Further each of these systems of various health facilities are independent of each other. The only viable solution is to use the data warehouse as the repository of detailed data from all the health facilities and using concepts of distributed systems to integrate these transactional systems with the data warehouse repository.

II. Technologies for Decision Support Systems

Developing a National Distributed Decision support health care system requires the use of various technologies. These are data warehousing for storage of detailed and aggregated data and distributed systems.

Muwanei Sinyinda is currently lecturing in the Centre for ICT Education at Mulungushi University in Kabwe, Zambia

Dr Douglas Kunda is currently the Director for the Centre for ICT Education at Mulungushi University in Kabwe, Zambia.

Mr Sibajene Gilbert obtained bachelors degree in Computer Science at Mulungushi University in Kabwe, Zambia.

2.1 Data Warehouses

According to bill inmon a dataware house is a "subject-oriented, integrated, time variant and non-volatile collection of data in support of management's decision making process" [2]. Data Warehouses (DW) systems represent a single source of information through which to analyze the status and the development of an organization [11]

Several other authors attempted to provide functional definitions for data warehouse. According to paulraj ponniah "The data warehouse is an informational environment that

- Provides an integrated and total view of the enterprise
- Makes the enterprise's current and historical information easily available for decision making
- Makes decision-support transactions possible without hindering operational systems
- Renders the organization's information consistent
- Presents a flexible and interactive source of strategic information" [3].

According to Ralph Kimball "a data warehouse is a system that extracts, cleans, conforms, and delivers source data into a dimensional data store and then supports and implements querying and analysis for the purpose of decision making".[4]

Traditionally, data warehouses (DW) have been counted among the most powerful problem-solving tools to enable easy access to information and enhance the effectiveness of decision-making processes[10].

Data warehouses are extensively used in various fields that include financial institutions, insurance, health and education. Over time there has been accumulation of data in online transaction processing systems and this data now has to be used to the benefit of organization. The greatest benefit is decision making. Top management of organization would make decisions not based on experience or intuition but based on analysis of accurate accumulated data in there information systems. Operational information systems are not developed to enable complex analysis but rather transaction processing with very limited analysis. Hence the need for a data warehouse that would be used sorely for analysis.

There are various data base management systems that that data warehousing. The notable ones are Mssql Server and Oracle. In this prototype presented in this paper, Oracle 11g has been chosen as the target database management system.

2.2 Distributed Systems

According to Coulouris and Dollimore "a distributed system is one in which components located at networked computers communicate and coordinate their actions only by passing messages"[5].

According to Ajay D. Kshemkalyani and Mukesh Singhal a distributed system is a collection of independent entities that cooperate to solve a problem that cannot be individually solved and the motivation for using a distributed system is some or all of the following
requirements:

- **Inherently distributed computations** In many applications that require reaching consensus among parties that are geographically distant, the computation is inherently distributed.
- **Resource sharing** Resources such as peripherals, complete data sets in databases, special libraries, as well as data (variable/files) cannot be fully replicated at all the sites because it is often neither practical nor cost-effective. Further, they cannot be placed at a single site because access to that site might prove to be a bottleneck.
- **Access to geographically remote data and resources** In many scenarios, the data cannot be replicated at every site participating in the distributed execution because it may be too large or too sensitive to be replicated. For example, payroll data within a multinational corporation is both too large and too sensitive to be replicated at every branch office/site. It is therefore stored at a central server which can be queried by branch offices. Similarly, special resources such as supercomputers exist only in certain locations, and to access such supercomputers, users need to log in remotely. Advances in the design of resource-constrained mobile devices as well as in the wireless technology with which these devices communicate have given further impetus to the importance of distributed protocols and middleware.
- **Enhanced reliability** A distributed system has the inherent potential to provide increased reliability because of the possibility of replicating resources and executions, as well as the reality that geographically distributed resources are not likely to crash/malfunction at the same time under normal circumstances.
- **Scalability** As the processors are usually connected by a wide-area network, adding more processors does not pose a direct bottleneck for the communication network.[6]

The choice of distributed computing for the national health care decision support is ideal due to the above requirements that distributed systems achieve. The various components that are heterogenous in nature will require a middle ware layer for the sake of interoperability. Large-scale distributed systems, such as e-healthcare systems, are difficult to develop due to their complex and decentralized nature. The Service Oriented Architecture facilitates the development of such systems by supporting modular design, application integration and interoperation, and software reuse. With open standards, such as XML, SOAP, WSDL and UDDI, the Service Oriented Architecture supports interoperability between services operating on different platforms and between applications implemented in different programming languages [7].

There are various programming languages that could be used for distributed computing. The most notable ones are java, cSharp and php. Java will be used for this prototype and has been chosen due to its features that favour distributed computing.

1. Architecture and Implementation of national health care decision support system

1.1 Data warehouse Design

1.2 An Entity Relationship Diagram is a diagrammatic view of all the database due to be developed.

The Entity Relationship Diagram below represents the database schema that lies at each respective health centre or clinic scattered across the country.

The OLTP databases found at all health centres or clinics will hold detailed data for patients, conditions , appointments, diagnosis , doctors , treatments as well as health care centres.

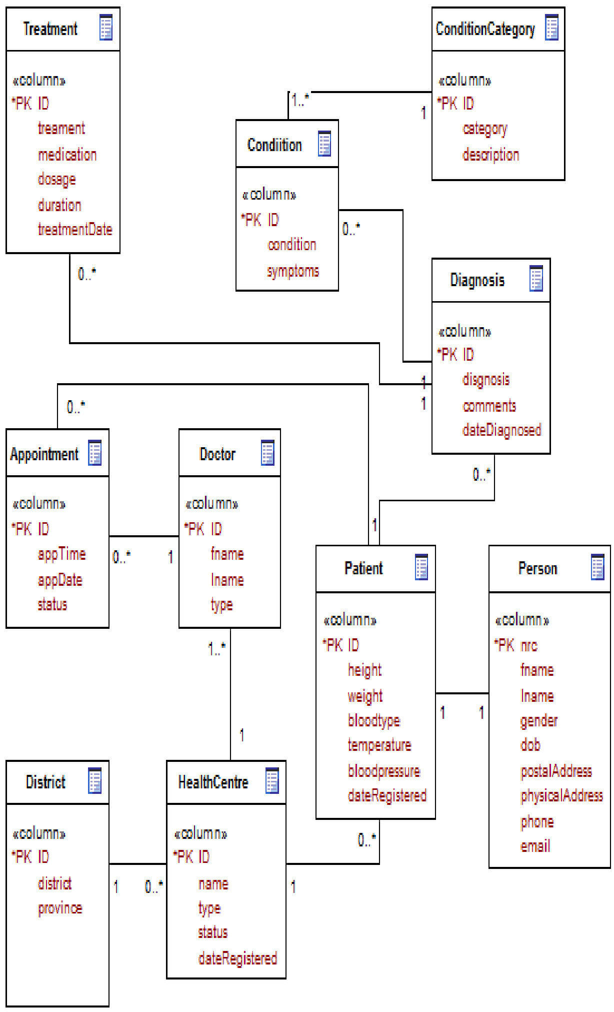

Figure 1 Entity relationship diagram

After resolving all cardinalities and constraints of the entities and relationships
in the above diagram relational schema is produced. These tables show the
final database structure in every health facility or clinic nation wide that will be repository of operational data.

Person	NRC(Primary key)
	Fname
	Lname
	Gender
	DoB
	postalAddress
	physicalAddress
	Phone
	Email
Patient	ID(primary key)
	NRC(foreign key)
	Height
	Weight
	Bloodtype
	Temperature
	Bloodpressure
	dateRegistered
	HCCID(foreign key)
Treatment	ID(primary key)
	Treatment
	Medication
	Dosage
	Duration
Doctor	ID(primary key)
	Fname
	Lname
	Type
	HCCID(foreign key)
Appointment	ID(primary key)
	AppTime
	AppDate
	Status
	patientID(foreign key)
	doctorID(foreign key)
Diagnosis	ID(primary key)
	Diagnosis
	Comment
	PatientID(foreign key)
	Datediagnosed
Condition	ID(primary key)
	Condition
	Symptom
	categoryID(foreign key)

Table 1

The data warehouse for the decision support system that will be used for online analytical processing will be held at the headquarters of health ministry or branch in the country. Data from all the OLTP databases in each clinic or health center nationwide would be extracted,cleaned and loaded into datawarehouse at the headquarters. The Data warehouse structure is generally represented with the help of the multi-dimensional schema which formulates information as facts and dimensions[13]. The star schema design is the chosen multi-dimensional schema and is derived from the entity relationship diagram above.

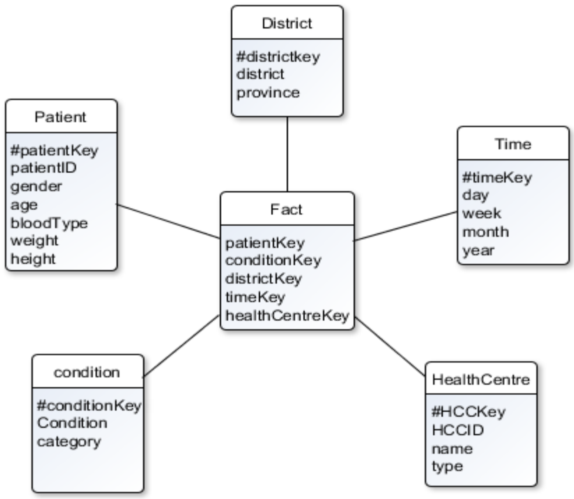

Figure 2 Star schema

Table	Attribute
Health Center	ID(primary key)
	Name
	districtID
	Status
	Type
	dateRegistered
District	ID(primary key)
	district
	Province
Fact	Patientkey
	Districtkey
	Healthcentrekey
	Conditionkey
	Timekey
Time	Key
	Date
	Days
	Weeks
	Month
	Year
Patient	Key
	patientID
	Height
	Weight
	Bloodtype
	Age
	gender
	dateRegistered
Condition	Key
	conditionID
	Condition
	category

Table 2: Data warehousing Tables

III. Extraction, Transformation and Loading Design

The Extraction, Transformation and Loading(E.T.L) procedure described in figure 3 below performs its extraction using an online procedure, also instead of flat files it directly runs a query on all participating databases by using multiples selects and appending them with the UNION keyword of SQL data manipulation language. After extracting data into the single staging area, the transformation procedure carries out a series of validation checks and updates or deletes records that do not satisfy this. The transformation of data also includes an aggregation process[8].The loading procedure is a series of *"INSERT */APPEND*/ AS SELECT"* statements to insert into the Dimensions. After the ETL processes are completed the analytical processes may begin. OLAP is meant to provide quality information to the end user by allowing the user to navigate through the dimensions of the data warehouse[9].

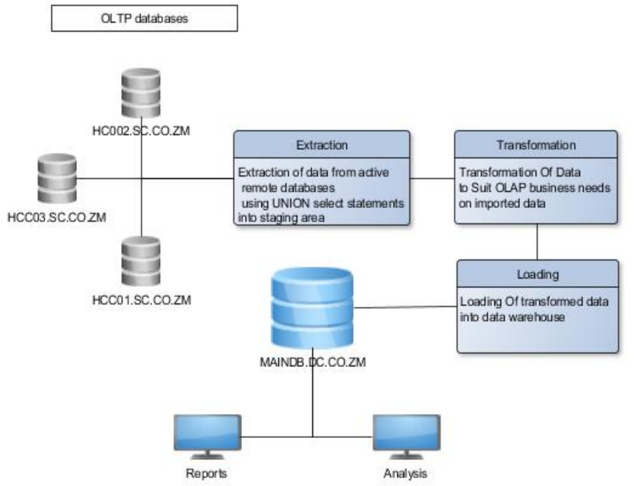

Figure 3 Extraction, transformation and loading diagram

3.2 Application Design and Development

The distributed application described in this paper has adopted the façade design pattern as the pattern of choice and based on the pattern are several interfaces that have been developed and relevant functionality exposed which also helps in maintenance of security in the system. Due to the fact that this is a heterogeneous distributed decision support system, a way of masking the differences in technologies is achieved by usage of web services. A number of business, data access object and façade classes which are shown on deployment diagram below are in the implementation. The web services are then developed to ensure there is coordination in system by masking a heterogeneity that might exist. The deployment diagram shows all the components of the system and how they interact with each other. The data access object has many methods some of which have been shown in the table below.

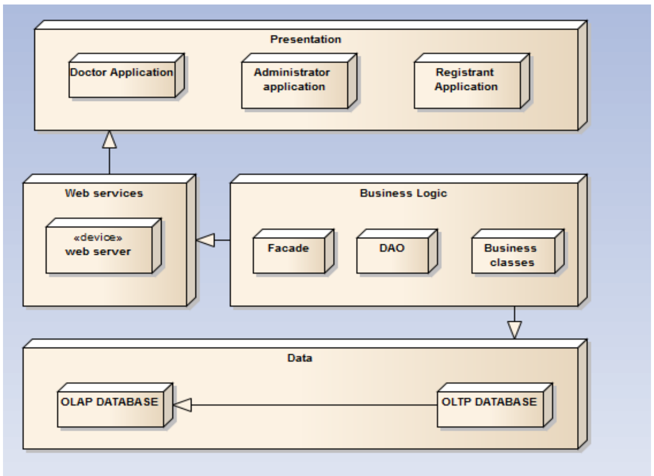

Figure 4 : Deployment Diagram

Detailed Method Description

ID	Name	Return type	Arguments
1	Authenticate	User	String username,String password
2	retrieveWare House	WareHouse[]	No Arguments
3	dropStagingArea	Void	No arguments
4	loadToWare House	Void	No arguments
5	OLTPTransformData	Void	No arguments
6	OLTPExtraction	Void	No arguments
7	getHealthCentre	HealthCentre	String healthCentreID
8	updateHealthCentre	Void	HealthCentre HC
9	deactivateHealthCentre	Void	HealthCentre HC
10	activateHealthCentre	Void	HealthCentre HC
11	newHealthCentre	Void	HealthCentre HC
12	getListOfHealthCentres	HealthCentre[]	No arguments
13	getDistricts	District[]	No argument
14	getDistrict	District	Int districtID

Table 3: list of AdminDao methods

The methods in the table above are critical to the functionality of system and are described below.

- The authentication method accepts a username and password as arguments and checks whether the current user exists or not. If the user exists it returns the user.
- The retrieveWarehouse method retrieves all the data from the warehousedata materialized view and returns an array of the warehouse class
- The method dropStaging area drops the temporal staging area created by the OLTPExtraction method.
- The loadToWarehouse method gets all the data cleaned by the OLTPtransformData method and loads it into the data house
- The OLTPTransformData Method cleans the data from the OLTPExtraction method into the temporal staging area
- The OLTPDataExtraction method goes through all active healthCentres and retrieves data from all the tables into the temporal staging area
- The getHealthCentre method simply gets a particular healthcentre by its ID
- Method updateHealthCentre updates the content of a healthCentre on demand by the admin
- Method deactivateHealthCentre changes the status of a healthcentre to deactivated and drops the database link between nit and that database.
- Method activateHealthcentre changes the status of the healthcentre to activated and create the database link between the databases.
- Method newHealthCentre creates a new healthcentre and assigns it a status of not activated.
- Method getListOfHealthCentres retrieves all the healthcentres in the database and assigns instances of the healthcentre class.

- Method getDistrict retirves a district by its ID and returns an instance of that district
- Method getDistricts returns all the districts in the database and assigns them to individual instances and loads them into an array.

AdminFacade

The AdminFacade creates abstraction for the AdminDao. It is created in a separate package and uses the same method name, return type and parameters as the AdminDao.

Web Services

The development of the web services is achieved using java and the glassfish server version 3+. The URL for the WSDL of the web service is http://localhost:30199/sharecarews/adminWS?WSDL the web service class implements all the AdminFacade class methods.

3.3 Report Generation

The generation of reports has been achieved in the prototype.Here are examples of some of the reports generated by the program.

SHARECARE

Health Centre Report

Friday 10 May 2013

ID	NAME	TYPE	STATUS	DISTRICT
Central				
HCC010	Kabwe Mine	Hospital	Not Activated	Kabwe
Lusaka				
HCC09	Kafue General	Hospital	Not Activated	Kafue
HCC07	Chilenje	Clinic	De-activated	Lusaka
HCC02	Kabwata	Hospital	Activated	Lusaka
HCC08	Libala	Clinic	Not Activated	Lusaka
HCC01	University Teaching	Hospital	Activated	Lusaka

Figure 5: Health centre summary

PatientCondition Summary

This report looks at how many people suffer from a particular illness

SHARECARE

Monthly Condition Summary Report

Friday 10 May 2013

HEALTHCENTRE	CATEGORY	CONDITION	NUMBEROFPATIENT
APRIL			
University Teaching Hospital	General	cachexia	2
University Teaching Hospital	General	loss of appetite	1
University Teaching Hospital	General	malaise	1
University Teaching Hospital	General	muscle weakness	1
University Teaching Hospital	General	weight gain	1
University Teaching Hospital	General	weight loss	1
MAY			
Kabwata Hospital	General	dry mouth	1
Kabwata Hospital	General	loss of appetite	1
Kabwata Hospital	General	weight loss	1

Figure: 6 patient condition summary report Patient registration Summary Number of patients that have registered at a health centre

Patient Registration Statistics

Friday 10 May 2013

REGISTEREDPATIENTS	GENDER	AGE
University Teaching Hospital		
february		
6	Male	57
Kabwata Hospital		
february		
6	Male	38
1	Male	58

figure 7: patient registration statistics

Patient list report

List of patients gotten remotely through the database links

Friday 10 May 2013

FNAME	LNAME	DOB	GENDER	HEIGHT	WEIGHT	DATEREGIS
KabwataHospital						
Bernedatte	Sampa	08-feb-1978	female	1.55	78.5	2/9/13 12:00 AM
University TeachingHospital						
Chileshe	Mwansa	08-feb-1971	male	1.56	71.5	2/2/13 12:00 AM
Gilbert	Sibajene	08-feb-1970	male	1.55	70.5	2/4/13 12:00 AM
Lombe	Banda	08-feb-1973	female	1.78	73.5	2/4/13 12:00 AM

figure 8: remote patient list

IV. Conclusion

There are several Clinical Data warehouse systems developed to utilize the potential knowledge in clinical data. However, most of the previous work focused on the medical information management[12]. The prototype of Decision Support System described in this paper uses data warehouse as data repository while distributed application designed using the façade design pattern. Some level of analysis has been achieved,ETL tools developed and functioning very well. However there is room for more improvement on the developed system on most modules. ETL processes need to be automatic after set intervals and there is need to implement a module for creating ad hoc queries.

References:

[1] Maldonado, J.A. E.U.I., Tech. Univ. of Valencia, Robles, M.Cano, Engineering in Medicine and Biology Society, 2001.

[2] W. H. Inmon, (2002) *Building The Data Warehouse* , Wiley Computer Publishing

[3] Paulraj Ponniah, (2001) *Data Warehousing Fundamentals: A Comprehensive Guide for IT Professionals*, Wiley Computer Publishing

[4] Ralph Kimball, Joe Casertam. (2004) *The Data Warehouse ETL Toolkit*.

[5] George Coulouris, Jean Dollimore, Tim Kindberg, Gordon Blair (2012), *Distributed Systems*, Pearson.

[6] Ajay D. Kshemkalyani, Mukesh Singhal(2008), *Distributed Computing*, Cambridge University Press.

[7] Firat Kart, Gengxin Miao, L. E. Moser, P. M. Melliar-Smith, 2007 IEEE International Conference on Services Computing SCC,(2007)

[8] Luis Marco-Ruiz, David Moner, José A. Maldonado, Nils Kolstrup, Johan G. Bellika, Archetype-based data warehouse environment to enable the reuse of electronic health record data, International Journal of

Medical Informatics, Volume 84, Issue 9, September 2015, Pages 702-714
[9] Anjana Gosain, Amar Arora, Security Issues in Data Warehouse: A Systematic Review, Procedia Computer Science, Volume 48, 2015, Pages 149-157
[10] Silvia Massa, Stefania Testa, Data warehouse-in-practice: exploring the function of expectations in organizational outcomes, Information & Management, Volume 42, Issue 5, July 2005, Pages 709-718
[11] Juan Trujillo, Emilio Soler, Eduardo Fernández-Medina, Mario Piattini, An engineering process for developing Secure Data Warehouses, Information and Software Technology, Volume 51, Issue 6, June 2009, Pages 1033-1051
[12] Xuezhong Zhou, Shibo Chen, Baoyan Liu, Runsun Zhang, Yinghui Wang, Ping Li, Yufeng Guo, Hua Zhang, Zhuye Gao, Xiufeng Yan, Development of traditional Chinese medicine clinical data warehouse for medical knowledge discovery and decision support, Artificial Intelligence in Medicine, Volume 48, Issues 2–3, February–March 2010, Pages 139-152
[13] Kanika Talwar, Anjana Gosain, Hierarchy Classification for Data Warehouse: A Survey, Procedia Technology, Volume 6, 2012, Pages 460-468

AUTHORS

Muwanei Sinyinda is currently lecturing in the Centre for ICT Education at Mulungushi University in Kabwe, Zambia. He holds a Master's degree in Engineering and Technology from Saint Petersburg Polytechnical University, Russia. He is experience in Java, C#, PHP/ MySQL web platforms, MS SQL Server and Oracle databases and CISCO networking.

Dr Douglas Kunda is currently the Director for the Centre for ICT Education at Mulungushi University in Kabwe, Zambia. He holds a Doctorate degree in Computer Science from the University of York, UK. He is Fellow of the Computer Society of Zambia and Member of Association for Computing Machinery. He worked as the Project Manager for the Integrated Financial Management Information System (IFMIS) project for the Ministry of Finance. He has presented papers at International Conferences and published in journals. He is certified SAP ERP Solution Manager Consultant with experience in Java, C++, PHP/ MySQL web platforms, Microsoft products including SQL Server databases, Oracle and Postgresql databases, Linux, CISCO networking, PRINCE2 and moodle.

Mr Sibajene Gilbert obtained bachelors degree in Computer Science at Mulungushi University in Kabwe, Zambia. He has good experience in in Java, Microsoft products including SQL Server databases and Oracle Databases