CLOUD-BASED DEVELOPMENT OF SMART AND CONNECTED DATA IN HEALTHCARE APPLICATION

S. Naisha Sultana[1], Gandikota Ramu[2] and Prof. B. Eswara Reddy[3]

[1, 2&3]Dept of Computer Science & Engineering, JNTUA College of Engineering, Ananthapuramu, Andhra Pradesh, India

ABSTRACT

There is a need of data integration in cloud – based system, we propose an Information Integration and Informatics framework for cloud – based healthcare application. The data collected by the Electronic Health Record System need to be smart and connected, so we use informatica for the connection of data from different database. Traditional Electronic Health Record Systems are based on different technologies, languages and Electronic Health Record Standards. Electronic Health Record System stores data based on interaction between patient and provider. There are scalable cloud infrastructures, distributed and heterogeneous healthcare systems and there is a need to develop advanced healthcare application. This advance healthcare application will improve the integration of required data and helps in fast interaction between the patient and the service providers. Thus there is the development of smart and connected data in healthcare application of cloud. The proposed system is developed by using cloud platform Aneka.

KEYWORDS

Cloud Computing, Data Integration, Informatica, Aneka, Electronic Health Records

1. INTRODUCTION

Information Integration and Informatics are specific challenges faced by healthcare applications. We present a cloud-based Information Integration and Informatics framework comprised of established open source cloud based technologies that are previously deployed in other domains, for collecting and analyzing clinical data from different semantic integration, geographical locations, knowledge discovery, secure access and data cataloging for clinical data. Traditional health IT systems such as Electronic Health Record(EHR) systems use separate and often conflict of technical and semantic standards for data description and storage and are based on different languages, proprietary architectures, database technologies and different representations of clinical data. The ability to exactly and smoothly integrate data from different health IT systems is important for the development of advanced healthcare applications.

A cloud based approach for design of interoperable EHRs[1] and an experimental system has been presented. Cloud Health Information Systems Technology Architecture (CHISTAR). CHISTAR achieves different types of interoperability through the use of a generic design methodology which uses a reference model that defines usual purpose set of data structures and an archetype model that defines the clinical data logic. CHISTAR supports higher security features such as identity management services, authorization services and authentication services for making secure access to healthcare data.

In this paper we propose an Information Integration and Informatics framework. The proposed framework allows: (1) Data Integration – integrating data from scattered and different sources into a same nomenclature and establishing effective use of clinical data combined from diverse EHRs, (2) Data Access – querying and getting healthcare data stored in the cloud, (3) Data Analytics – efficient data analysis of big healthcare data collected in cloud, (4) Data Storage – healthcare data storage and lifecycle management. The importance of the proposed work obtained from the using and developing of technologies for clinical data tools, integration and techniques for explicate healthcare data effectively, which strengthen the benefits and economics of present cloud computing environments already in or entering the use in other domains. With the current increase of clinical data the problems of how to collect data from scattered and different systems and how to analyze the big scale clinical data become cavilling. Data integration approaches for collecting clinical data from scattered and different health IT systems will lead to more effective healthcare applications. Approaches for large scale clinical data analytics will lead in development of more effectual healthcare applications, improve the quality of predictions and in timely decision making. The proposed Information Integration and Informatics framework allows healthcare data accumulation and organization from different stakeholders(doctors, patients, payers, specialists, physicians etc) as in Fig.1 and disparate data sources(databases, unstructured and structured formats, etc)as in Fig.2. With Information Integration and Informatics framework healthcare data can be safely exchanged between different stakeholders to assist more accurate and timely healthcare delivery. The proposed Information Integration and Informatics framework will assist the evolution of advanced healthcare applications. We now list a few healthcare applications that can be developed using the Information Integration and Informatics framework: Informatica Administrator is a Web application that is used to administer the PowerCenter domain and PowerCenter security. We can also administer application services for the Informatica Analyst and Informatica Developer Application services for Informatica Analyst and Informatica Developer include the Analyst service the Model Repository Service and the Data Integration Service[13]. Informatica is ETL tool, which means Extract, Transfer, and Load. For demonstrating the need for an Information Integration and Informatics framework, we ponder a use case of a healthcare application for smart and connected data. Smart and Connected Data (SCD) systems study the spread and determinants of health related states or events in populations and utilize these research for diagnosis of disease under surveillance at rational level to control health problems[2]. EHR systems include individual – level diagnostic, laboratory results, demographic data and treatment. Although EHRs were design for clinical interactions between provider and patient the EHR data can be utilized for disease detection, outbreak prediction, population – level health surveillance efforts and public health mapping. Because data in EHRs is intermittently updated, there is a need for a framework that integrates data from many EHR systems for perfectly and regularly predicting outbreaks. The scale of data collected by EHR systems is large. Therefore, perfect data integration and analysis techniques are required for using EHR data for advanced healthcare application.
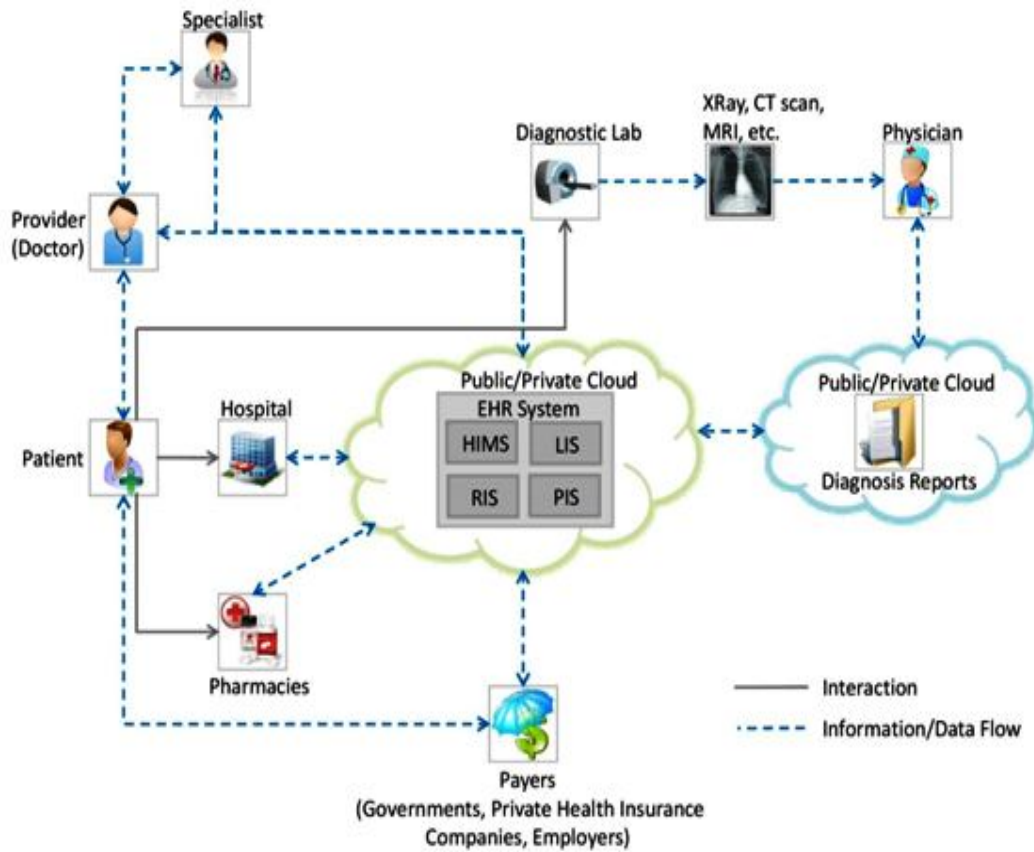
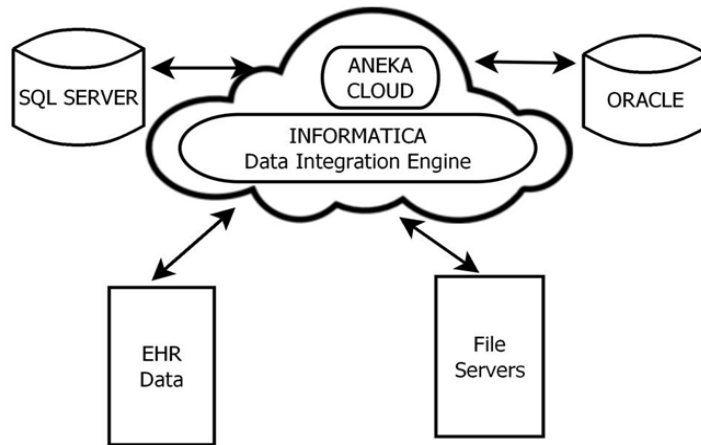Fig 1. Application of cloud computing environment

Fig 2. Data integration approach

2. RELATED WORK

VistA[3] is widely used EHR system in the United States of America. OpenEHR[4] is an EHR system which is for getting semantic interoperability. OpenEHR puts particular emphasis on

semantic interoperability better the kind of data interchange between different stakeholders in the healthcare ecosystem. OpenEHR is based on a two – level approaching which a reference model consist the first level of modeling and the formal meaning of clinical content in the form of a archetypes and templates is the second level. To have exchange of healthcare data different solutions have been made that allow combining data from different sources. Mirth Connect[5] is an open source integration engine which supports different protocols and messaging standards to connect to databases and external systems. FM Projection[6] is a set of tools which allows examing VistA File Manager data and structure by using SQL like representations.
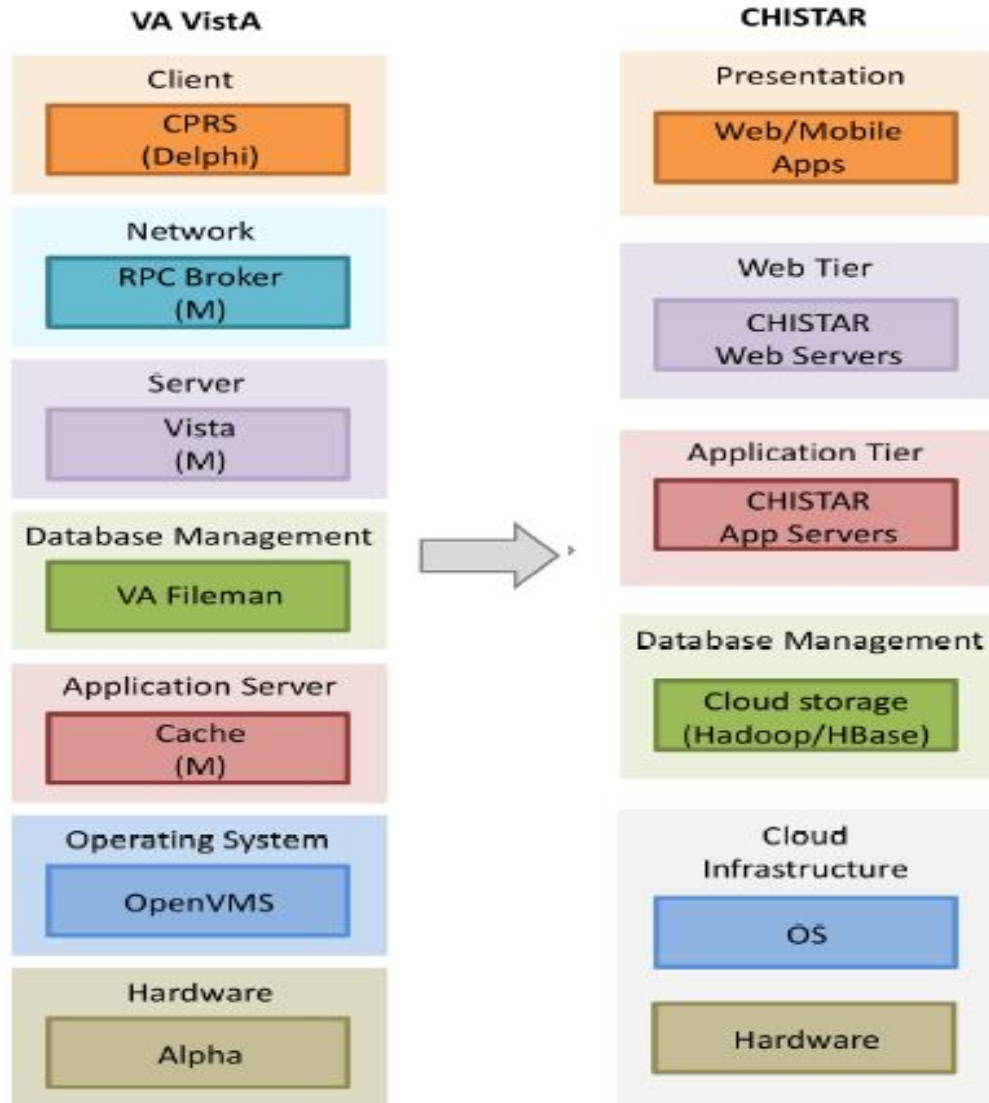
Fig 3. Technology stacks of VA VistA,CHISTAR system

A. Bagha et al[7] which proposes a data collection framework for collecting massive sensor data in a cloud. For CHISTAR a same approach for data collection which is based on a cloud – based distributed batch processing infrastructure in Fig.3. Since EHR systems handle big healthcare data, surveying the performance of such systems is very important to make sure the correctness of such systems in providing healthcare. For testing cloud – based system like EHRs there is an approach for prototyping and surveying cloud – based systems[8][9]. A similar way will be used

to determine the performance of CHISTAR. CHISTAR uses the Cloud Component Model approach for application design[10]. For the design of mobile application of cellular networks, CHISTAR adopts as in[11]. Fig4 shows the layered architecture of CHISTAR system. The infrastructure service layer comprises of the cloud instances (for application servers, load balancers, slave nodes, Hadoop masters etc.) on which CHISTAR is used. The information service layer comprises of model for data storage, data integration engine, and clinical concepts and data governance module. The application service layer provides various services such as demographic service, EHR service, terminology service and archetype service. The presentation services layer comprises of the smart and connected healthcare applications like web and mobile based.
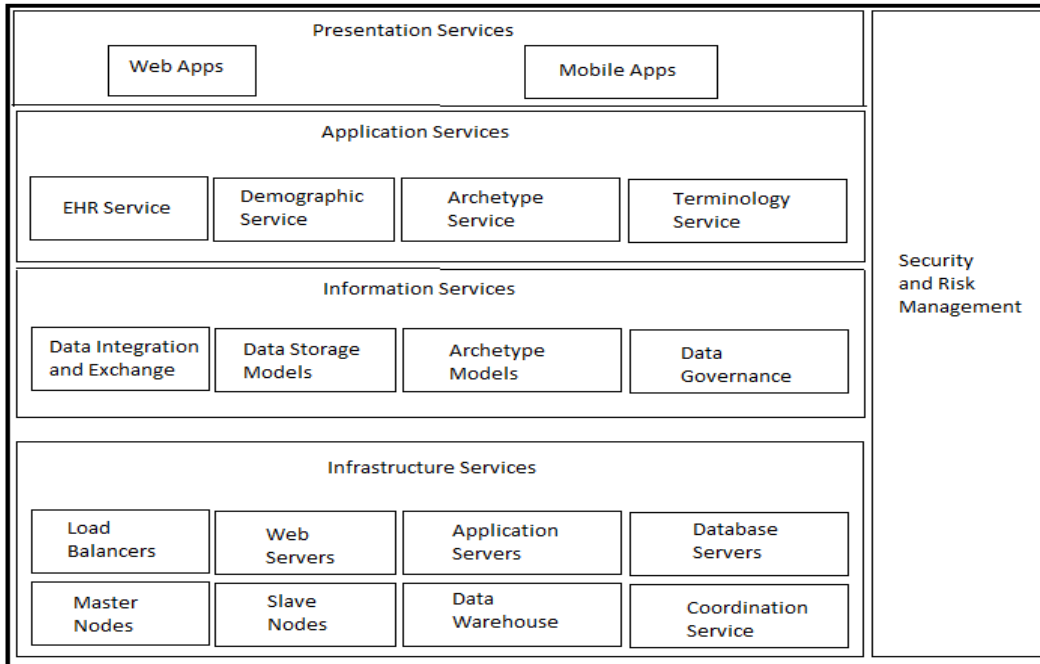

Fig 4. Architecture of CHISTAR

3. PROPOSED SCHEME

The proposed technology stack consists:

3.1. Informatics Application Builder: The Informatics Application Builder furnishes tool and APIs for constructing web and mobile healthcare application. Application developers can use the Informatics Application Builder to construct healthcare application without worrying about deployment configuration, cloud infrastructure management and data management in cloud. The Information Integration and Informatics framework and CHISTAR middleware take control of all the cloud configuration and maintenance activities.

3.2. CHISTAR Middleware: CHISTAR middleware[1] gives a platform for advancement of informatics application development.

3.3. Information Integration and Informatics Framework: Information Integration and Informatics framework makes storing, integrating and analyzing EHR data in the cloud. The Information Integration and Informatics framework consist of

(a) Data Integration Framework
(b) Data Access Framework
(c) Data Orchestration
(d) Data Analytics Framework
(e) Data Storage Framework

3.4. CHISTAR Security Service: CHISTAR gives the key requirements of HIPAA and HITECH acts that requires entities (that create, transmit, maintain, disclose and use an individual's protected health information) to make sure their customers that the availability, confidentiality and integrity of information key security features of CHISTAR contains authentication, authorization, identity management, securing data at rest, securing data in motion, key management, data integrity and auditing. The data – level integration system comprises of data storage model, archetype model, data sources and mapping as shown in Fig.2. Data sources shows the containers of data, the domain model provides a conceptual technique of the applications, data storage model gives the logical structure of data storage and mapping which maps the source data with the target data. In the data – level integration the representation of the data structure is different from the clinical information.

CLOUD – BASED Application:

The CHISTAR middleware and the proposed Information Integration and Informatics framework gives at a time development of advanced healthcare applications. For making the development of user interfaces for healthcare application CHISTAR gives platform tools for desktop platforms and mobile platforms. Application services of the CHISTAR middleware are platform independent services where the platform tools customize to handle features of required platforms. Application services and platform tools help in making healthcare application for variety of platforms and devices. The Information Integration and Informatics framework provides tools and APIs for data access, integration, analytics and storage.

4. IMPLEMENTATION DETAILS

Fig5 shows the Informatica PowerCenter Designer, which consist of

D- Designer Interface
R- Repository Manager
W- Workflow Manager
M- Mapping
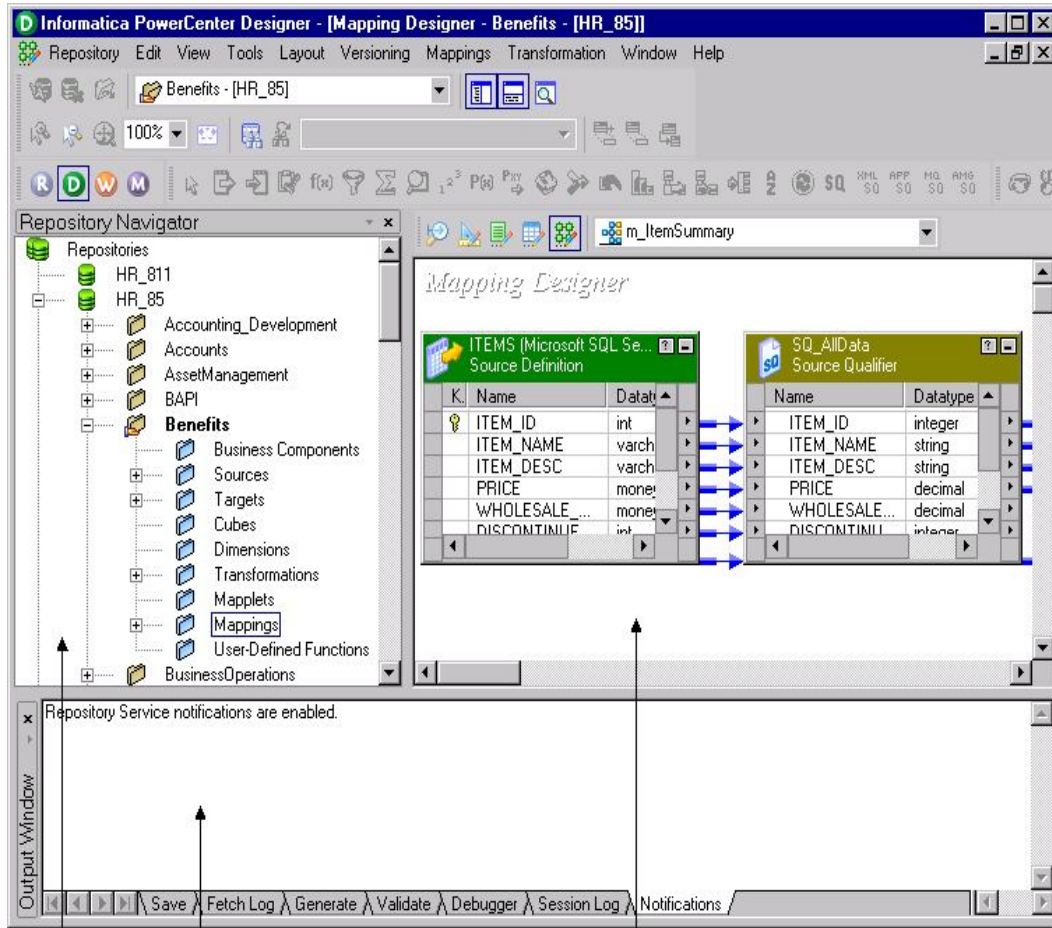
1. Navigator
2. Output
3. workspace

Fig 5. Default designer interface

In Designer Interface we connect repository and import database of source where we select ODBC datasource : remotesqlserver(SQL Server). Then we import target database. Mapping in menu and map as shown in workspace. In Workflow Manager we connect repository and select session and create workflow by name WFOracletoSQL. Creating a task and linking the task with workflow. Then Mapping is done by giving source and target as $source: Relational MicroSoft_SQL_Server$Target: Relational Oracle. By right clicking on workspace we start workflow. In Fig.8 we see the successful linking of source and target. In simple CHISTAR's multi-tier deployment configuration the naming convention are as (#L(size) = #A(size) = #H(size)), where L is the number of instances running load balancers. A is the number of instances running application servers. H is the number of instances running the database of cluster. Fig.6 shows the vertical and horizontal scaling options of CHISTAR. Comparing (1L(Large) = 2A(small) = 2H(Large)) and (1L(small) = 3A(small) = 2H(Large)) we get that by horizontal scaling (increasing the number of application server) lower response time is achieved. Fig.7 shows the average response time for the CHISTAR application for different configuration and varying number of simultaneous users. The results in Fig.7 were obtained with 1000 patient health records in the CHISTAR application with increase in number of user the mean request

arrival rate increases since CHISTAR services higher number of request per second, therefore an increase in response time is observed. The entire interface which is developed in C# is connected to cloud by using cloud platform Aneka[12]. As Aneka has workload distribution capabilities and its thread model.
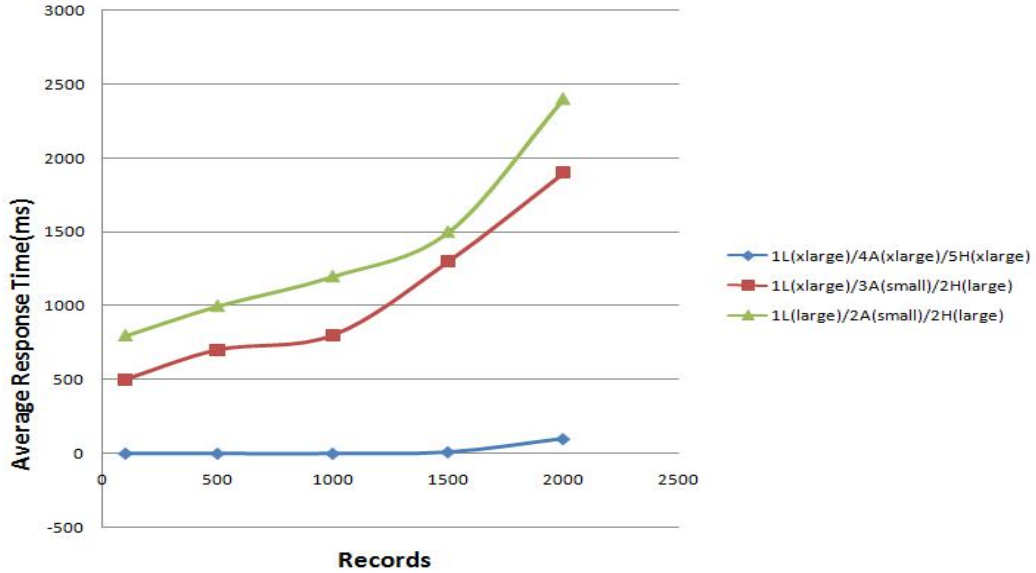

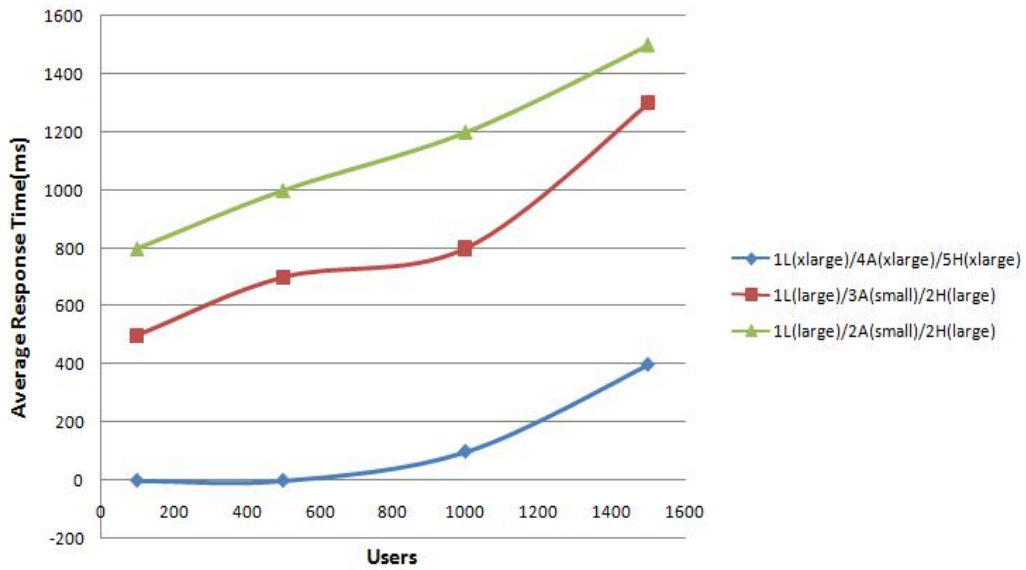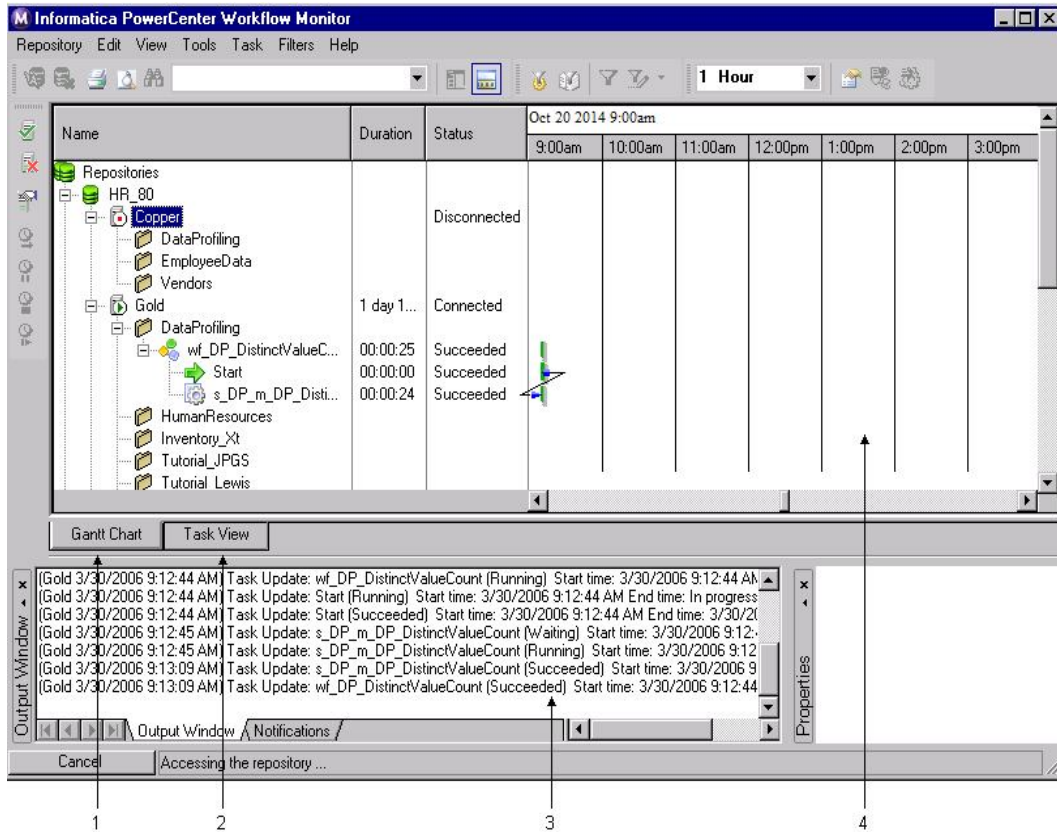Fig 6. Average response time for patient records.

Fig 7. Average response time for simultaneous users.

1. Gantt chart view
2. Task view
3. Output window
4. Time window

Fig 8. Informatica Workflow Monitor.

5. PERFORMANCE EVALUATION

5.1. Rapid Application Development: Traditional health IT system are related on client – server architectures and store all data in house which requires a server, software and hardware which are installed locally. Cloud computing makes data to be stored on external servers that can be accessed with the help of web. The proposed stack contains CHISTAR middleware and Information Integration and Informatics framework which facilitate rapid development of cloud – based healthcare application.

5.2. Integrated Health Data: The Data Integration framework makes integrating specific data from different data sources such as Oracle, SQL Server, file servers and EHR standards in a cloud – based storage.

5.3. Accessibility: Cloud – based applications built along with Information Integration and Informatics framework have high approach and understanding over client – server applications because users can securely log in to the system from everywhere if they have Internet connectivity. Information Integration and Informatics framework given a compatible representation, elucidation and access of the integrated healthcare data. Improved security

application of CHISTAR middleware such as role access control allows specialists, doctors, and physicians to access more strongly in a secure environment and provides good continuity of care.

5.4. Scalability: Cloud – based applications which are built with Information Integration and Informatics framework have good scalability compared with client – server EHRs. The computing facility can be scaled up on demand as large data is combined and new users are added. Applications which are built with Information Integration and Informatics framework can lead both horizontal (scaling out) as well as vertical scaling (scaling up) options. Information Integration and Informatics framework uses HBase in storing data. Applications which has HBase storage can be scaled linearly and at a time by adding of new nodes.

5.5. Reduced Costs: Cloud – based applications which are built with Information Integration and Informatics framework have reduced services and operation costs as related to traditional client – server applications. Client – Server   applications needs a team of IT experts to configure, install run, test, secure and update software and hardware. In cloud – based applications all of that is done in the cloud by the cloud provides.

6. CONCLUSION

This paper describes the design of an Information Integration and Informatics framework that allows storing, integrating and analyzing healthcare data in the cloud. The Information Integration and Informatics framework allow the development of advanced healthcare application with data integrated in difference database. Application developers can quickly develop healthcare applications by not thinking about the data management in cloud and cloud infrastructure management deployment configuration which are taken care by the Information Integration and Informatics framework. We use Aneka Cloud Platform for the development. The effectiveness of the framework is demonstrated with Health Mapper Applications and Patient Health Records.

ACKNOWLEDGEMENT

REFERENCES

[1]   A. Bhaga, V. Madisetti, A Cloud-based Approach for Interoperable Electronic Health Records (EHRs), IEEE Journal of Biomedical and Health Informatics, vol. 17, iss. 5, pp. 894 – 906, 2013.
[2]   C. Hajat, An Introduction to Epidemiology, Methods Mol Biol. 2011; 713:27-39.
[3]   VistA Monograph [Online]. Available: www.va.gov/vista. monograph, 2012.
[4]   OpenEHR. (2012). [Online]. Available: https://www.openehr.org
[5]   Mirth Connect. (2012). [Online]. Available: http://www.mirthcorp.com/products/mirth-connect
[6]   Medsphere FM Projection. (2012). [Online]. Available:http://medsphere.org/community/project/fm-projection.
[7]   A. Bahga and V.K. Madisetti, "Analyzing massive machine maintenance data in a computing cloud," IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 10, pp. 1831-1843, Oct. 2012.
[8]   A. Bahga and V.K. Madisetti, "Synthetic workload generation for cloud computing applications," J.Softw.eng.Appl., vol. 4, no7, pp.396-410, 2011.
[9]   A. Bahga and V.K. Madisetti, "Performance evaluation approach for multi-tier cloud applications," J.Sofw.Eng.Appl., vol. 6, no.2, pp.74-83, 2013.
[10] A. Bahga and V.K. Madisetti, Rapid Prototyping of Advanced Cloud Based Services and Systems, IEEE Comput., under review, Aug. 2012.
[11] N. Radio, Y. Zhang, M. Tatipamula, and V.K. Madisetti, "Next-generation applications on cellular networks: Trends, challenges, and solutions," Proc. IEEE, vol. 100, no.4, pp. 841-85, Apr. 2012.

[12] Aneka[Online]. Available : http://www.manjrasoft.com/
[13] Informatica Powercenter[Online]. Available : http://www.informatica.com/in/ETL/

Authors

Ms. S. Naisha Sultana received the B.E Degree in Information Technology from G. Pullaiah College of Engineering and Technology affiliated to JNT University Anantapur, Anantapuramu, India in the year 2012, and pursuing M.Tech Degree in Software Engineering at Computer Science Department of JNT University Anantapur, Anantapuramu. Her research area in M.Tech deals with Cloud Computing.

Mr. Gandikota Ramu received B.Tech degree in Information Technology and M.Tech degree in Computer Science from JNT University Anathapur, Anathapuramu, India. He is currently working towards Ph.D Degree at JNT University Anantapur, Ananthapuramu, India. He is working as Assistant Mentor in Master of Science in Information Technology (MSIT) program, JNTUA (in collaboration with Consortium of Institutions Higher Learning and Carnegie Mellon University, USA).

Dr. B. Eswara Reddy Graduated in B.Tech. (CSE) from Sri Krishna Devaraya University. He received Masters Degree in M.Tech.(Software Engineering), from JNT University, Hyderabad. He received Ph.D in Computer Science & Engineering from JNT University, Hyderabad. He is working as Professor in CSE Dept. and Vice-Principal at JNTUA College of Engineering, Ananthapuramu. He has more than 50 Publications in various International Journals and Conferences.. He has received UGC Major Research Project (MRP) titled 'Cloud Computing framework for rural healthcare in Indian Scenario'. His research interests include Pattern Recognition & Image Analysis, Data Mining and Cloud Computing. He is a Fellow of IE (India). He is a life member of ISTE, ISCA and member of CSI and IEEE.