ISSN XXXX XXXX © 2017 IJESC

Research Article Volume 7 Issue No.5

A Cloud Based Distributed Medical Record System with Big Data Analytics

Poornima .R. Patil[1], Prof. Amareshwari Patil[2]
PG Student[1], Assistant Professor[2]
Department of CSE
PDA College of Engineering, Kalaburagi, Karnataka, India

Abstract:
Medical health records are also known as public health record (PHR)/electronic health record (EHR).Generally the records stored in the hospital are very complex it contains data such as personal information which is in the relational database format, several medical diagnosis test which may be blood sugar or blood pressure, observation by the doctor or review of other medical diagnostic team, various medical imaging such as x-ray etc. All this data must be stored in a single database or server such system takes immense amount of time for processing the data. The hospital may also have active data associated with the patient. An active data need to be analyzed in the real time. Processing such large volume of data is not being supported by the current technology. Several recent works in big data technology has made such medical health records application to be migrated into the cloud that to be migrated into big data framework, further such big data framework are been processed by the front end system such as perl, php and ruby etc. The objective of this work is to consider a large amount of data with variety, velocity and process the data in real time on cloud over the big data technology. The current system of accessing the big data technology does not support cloud API and integration. In this work we extent big data system for processing medical records with cloud integration and demonstrate that our proposed architecture can not only analyze the data in real time but is much more efficient
in comparison to current technology as well as cluster based computing.

Keywords: API, Big data, Cloud, Health care Management

I. INTRODUCTION

Hospital Management System is encountering off-the-outline development in data era. Increment in number of clinical solutions in-tern generate vast measure of data consistently including Electronic Medical Records, digital image and so on. The other clinical data incorporate the wearable sensor networks aggregate information on patient's heart rate, brain activity, temperature, muscle motion and so forth. This generates enormous measure of information. The data generated is confidential and ought to be overseen properly.The availability of patient's data totally and rapidly can make the capacity for new revelations and leap forward, and additionally giving in-time treatment. The general performance of the Hospital Management System can be enhanced by the better data and better results for patients. The health care management incorporates substantial measure of data which can be described as big data, which can be characterized by 5Vs as far as Volume, Velocity, Variety, Value, and Veracity. The patient data are of expansive bytes, which depicts about the volume. The data entry rate from the patients can be known as the velocity. The diverse sorts of data sets regarding the structured, semi structured and unstructured data sets, for example, clinical reports, EHRs, and radiological images and so forth allude to the assortment and veracity clarifies the honesty of the data sets as for data availability and authenticity. The gathered data are changed into important judgment which refers to the value the data of patients are the primary and main entities in healthcare big data analytic. Hence, valid raw data must be collected in an efficient manner in a medical environment. In advanced healthcare systems, the patient data are collected through wearable devices equipped with different types of sensors. If the hospital tries to locate an another branch of its in some other geographic location then migrating this big data to that selected geographic location might result in data loss. This can be overcome by migrating the big data over the cloud. Cloud computing is nothing but the integration of several grid computers along with their data center which virtually represents the single computer from outside

II. RELATED WORK

A framework for handling different Big Data which can be in structured, semistructured and unstructured format. The concept uses Hadoop ecosystem and H-Base, which is a non-relational database executed in a distributed manner for storing, fetching and retrieving of heterogeneous datasets[1] The healthcare industry needs to reap the benefits of emerging technologies such as mobile computing and cloud computing, along with the use of Health Information Technology (HIT) to help solve the ever growing operating cost problems [2].With convergence of advanced computing and numerous Big Data technological options like commercial solutions, Open Source, Cloud etc. It is now possible to attain high performance, scalability at a relatively low cost. Big data solutions often come with set of innovative data management solutions and analytical tools, when effectively implemented can transform the healthcare outcomes[3].Big data analytics in healthcare is evolving into a promising field for providing insight from very large data sets and improving outcomes while reducing costs. Its potential is great [4].Cloud computing technology can greatly improve the level of healthcare services. Here a cloud computing based remote healthcare service system was introduced[5].The proposed approach not only utilizes clinical information but also personalized information by correlation to find hidden information using big data health analytic for improvement of life-care[6].

III. PROPOSED SYSTEM

Here we propose a novel architecture where by the traditional data stored in the relational database is first migrated to a big data framework using ETL(Enterprise transfer logic) tool this is followed by implementing a data analytic service on the bigdata system and followed by extending the services through the cloud API. Once the entire medical health record analytics is imported into cloud environment our work also provides a platform for data analytics of the medical data. It also supports device independent access of the data by using API's.

A. Block Diagram of the Proposed System

The hospital management includes large amount of patients data the data might be the patient's vitals data, patients visit data, and patients clinical report etc. All this form the health record. The traditional database can't manage such large amount of data as their will be chances of data loss due to the increased load on the system so this data is migrated into the big data system through ETL(Enterprise transfer logic) tool. The data is further passed through the cloud for easy accessing of data through API's and analytics are performed on the data, which include the predictive analysis, diagnosis analysis and the statistical analysis. Apache Spark is the big data framework which helps us performing general data analytics on the distributed cluster like hadoop.It runs on top of the existing hadoop cluster and access hadoop data store. It is the alternative to the traditional map reduce model.Its features includes the speed, easy to use and can run everywhere.It provides memory computation for the increase speed and data process over map reduce.

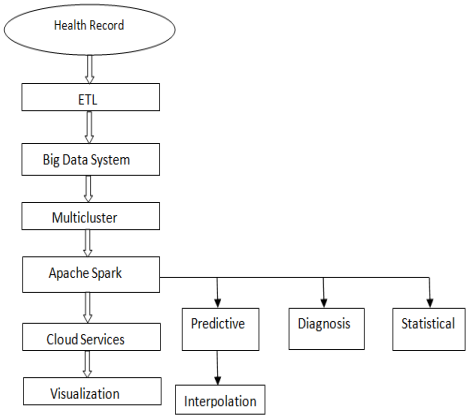

Figure.1. Block Diagram

IV. IMPLEMENTATION AND RESULTS

We have proposed an open source real time cloud based medical health record system which is HIPPA complaint based on popular OpenMRS health care system architecture. OpenMRS is the software platform and the reference application which empowers the outline of a redid medical records system The system is based on the theoretical database structure which is not subject to the real sorts of medical information required to be gathered or on specific information accumulation structures thus can be altered for various employments. Through its open source group it has developed into a medical informatics stage utilized on each mainland OpenMRS is based on the principle that information ought to be put away in a way which makes it simple to outline and break down, i.e., insignificant utilization of free content and most extreme utilization of coded information. The OpenMRS is intended to be a non specific medical record system that can bolster the care of patients, observations, encounters, notes, and other information from the healthcare system and rendering those in synopses, reports, and information sees that would enhance the viability of the general population utilizing the system. OpenMRS is a Java-based, web-based electronic medical record. We began from a straightforward (at any rate it used to be basic) information demonstrate, wrapped that into an API, and after that assembled a web-based application that uses the API.OpenMRS is a client-server application, which implies it is intended to work in a domain where numerous client PCs get to similar information on a server. Through the OpenMRS application which we have implemented using the android studio as the tool we can find the records of the patient who are registered, the start and end date and time of the particular patients visit can be known, the diagnosis of the patient can be known. Through the software platform of OpenMRS we can even perform the analytics like statistical and diagnostic analysis. The information related to the particular patient vitals or parameters based on his/her every visit can be drawn from the OpenMRS patient's record. There might be the chances of missing the vitals data due to the patient's irregular visits or the hospital management negligence, those missing data can be predicted by the predictive analytics. We are using the matlab as distribution for predictive analysis through which the analysis can be done easily and accurately. Fig shows the statistical analysis performed on the systolic blood pressure parameter .where the number of patients who are associated with this parameter or vital or concept is known and the minimum and maximum is set by default, mean and median value is calculated all this is known as the data definition.The number of occurrences of the patients for the particular vital is represented on the y-axis and the value associated with the patients on the x-axis.

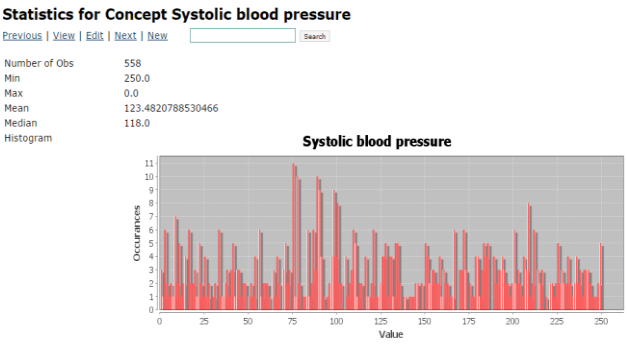

Figure.2. Histogram Graph of Systolic Blood Pressure Parameter

Fig below shows the histogram graph of systolic blood pressure without the outliers i.e the patient cannot have the 0 systolic blood pressure and can't exceed the 250 so the outliers are removed and displayed

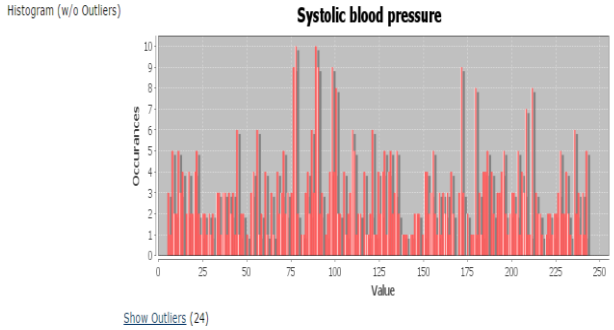

Figure.3. the Histogram of Systolic Blood Pressure Graph Without the Outliers

Fig below shows the time series graph for the systolic blood pressure parameter, which gives the information of the dates on which the patient's visited and what was the minimum or maximum systolic blood pressure reading of the patients.

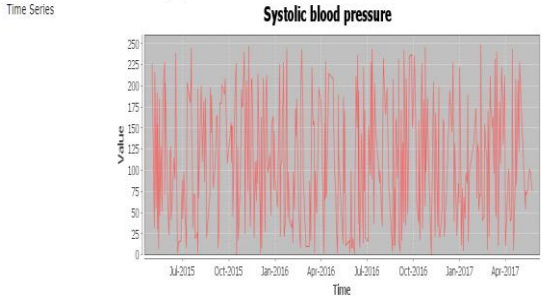

Figure.4. Time Series Graph for The Systolic Blood Pressure Parameter

Diagnostic analysis performed can be shown below. Fig below shows the diagnostic analysis for the particular diseases i.e pregnancy, miscarriage. Object concepts need to be added for performing diagnosis anlaysis this can be done by adding the observation and then save the encounter details based on which the graph will be displayed.The overall patients out of which the patients suffering with the particular diseases is shown through the dark color and No represents the rest patients who are not suffering with the particular disease.

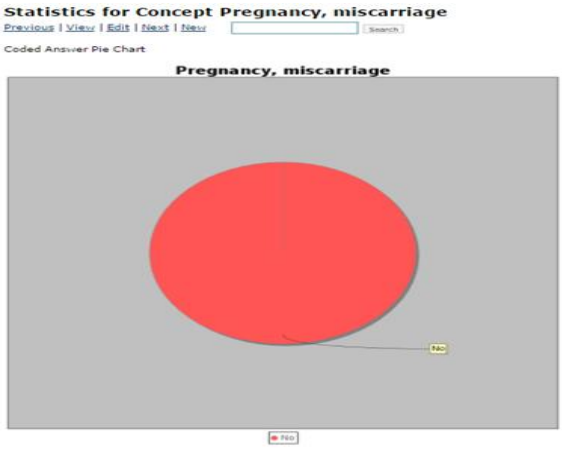

Figure.5. Diagnosis Analysis Performed On The Diseases Pregnancy, Miscarriage

Fig below shows the predictive analysis graph of the temperature parameter. As we can see the green colored line which represents the missing data or the incomplete series that is predicted using the interpolation method. Where x-axis represents the samples and y-axis represents the (' C') % for temperature parameter.

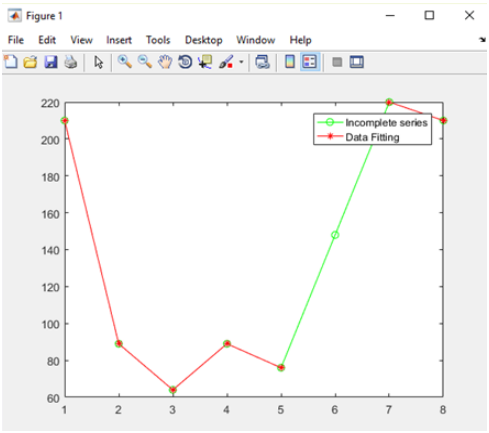

Figure.6. predictive analysis performed on the temperature parameter

V. CONCLUSION AND FUTURE WORK

The ongoing infrastructure improvement and software as a service more and more solution has been migrated into the cloud, a cloud based solution provides a distinct feature of device independent access where by the same features can be accessed across various platforms like mobile, browser etc. With the advancement of the cloud based system many medical applications are also migrated over the cloud. The advantage of such a system is that it not only gives all time access of the patient data but also provides an easy analytics service through the big data framework. Even though various past authors have proposed efficient architecture for big data based medical solution ,however an open cloud based API with data analytics and predictive analytics has never been proposed in the past(to the best of our knowledge). In this work we have proposed an open source real time cloud based medical health record system which is HIPPA complaint based on popular OpenMRS health care system architecture. Our system provides a cloud based API to access modify and upgrade patients data in the cloud at the same time a data visualization extension in the mobile further our system also provides an easy analytics services on various parameters and at the same time it offers an easy predictive analytics of the missing data. Our system has been tested across wide range of devices such as mobile, browser and the performance of the access time as found to be extremely homogenous. This system can be further improved by incorporating common predictive analytics module, disease prediction module. And can be further improved by including independent machine learning framework that can assist the user to diagnose their health system based on the vitals and improved parameters efficiently.

VI. REFERENCES

[1].Dr. Nagaratna P. Hegde, Prof. Sindhu .C.S"A Framework to Handle Data Heterogeneity Contextual to Medical Big Data" Vasavi College of Engineering, Hyderabad, India. 2015 IEEE International Conference on Computational Intelligence and Computing Research.

[2].Nina S. Godbole, John Lamb," Calculating a Hospital's IT Energy Efficiency and Determining Cost Effective Ways for Improvement"2014 IEEE.

[3].Prabha Susy Mathew, Dr. Anitha S. Pillai "Big Data Solutions in Healthcare: Problems and Perspectives" IEEE Sponsored 2nd International Conference on Innovations in Information Embedded and Communication Systems 2015.

[4].Wullianallur Raghupath and Viju Raghupathi "Big data analytics in healthcare: promise and potential" Health Information Science and Systems 2014

[5].Gao zhiqiang, He lingsong , Tian hang, Ling cong "A Cloud computing based mobile healthcare service system"proc of the 2015 IEEE 3rd International conference on smart Instrumentation, Measurement and applications (ICSIMA 2015)

[6].Wajahat Ali Khan, Muhammad Idris, Taqdir Ali, Rahman Ali, Shujaat Hussain, Maqbool Hussain, Muhammad Bilal Amin, Asad Masood Khattak, Yuan Weiwei, Muhammad Afzal, Sungyoung Lee, Byeong Ho Kang "Corelating health and wellness analytics for personalised decision making" 2015

17th International Conference on E-health Networking, Application & Services (HealthCom)

[7].Rabi Prasad Padhy, Manas Ranjan Patra, Suresh Chandra Satapathyv "Design and Implementation of a Cloud based Rural Healthcare Information System Model" Rabi Prasad Padhy et al, UNIASCIT, Vol 2 2012, 149-157.

[8].Nina S. Godbole, John Lamb "Using Data Science & Big Data Analytics to Make Healthcare Green" 2015 IEEE

[9].Survey on technologies for utilization of resource in cloud computing system(ICCRAES-16),Vol 23,issue 06,2016,Vol23 16 P068 [IJETCSE]

[10].K. Lin, F. Xia, W. Wang, D. Tian, and J. Song, ``System design for big data application in emotion-aware healthcare," IEEE Access, vol. 4, pp. 6901_6909, 2016.

[11].V. Tresp, J. M. Overhage, M. Bundschus, S. Rabizadeh, P. A. Fasching, and S. Yu, ``Going digital: A survey on digitalization and large-scale Data Analytics in healthcare," Proc. IEEE, vol. 104, no. 11, pp. 2180_2206,Nov. 2016.

[12].S. Rallapalli, R. R. Gondkar, and U. P. K. Ketavarapu, ``Impact of process-ing and analyzing healthcare big data on cloud computing environment by implementing hadoop cluster," Procedia Comput. Sci., vol. 85, pp. 16_22,May 2016.

[13].T. Huang, L. Lan, X. Fang, P. An, J. Min, and F. Wang, ``Promises and challenges of big data computing in health sciences, " Big Data Res., vol. 2,no. 1, pp. 2_11, 2015.

[14].J. Dean and S. Ghemawat, ``MapReduce: Simpli_ed data processing on large clusters," Commun. ACM, vol. 51, no. 1, pp. 107_113, Jan. 2008.

[15].J. Andreu-Perez, C. C. Y. Poon, R. D. Merri_eld, S. T. C. Wong, and G.-Z.Yang, ``Big data for health," IEEE J. Biomed. Health Inform.,vol. 19,no. 4, pp. 1193_1208, Jul. 2015.