Automated Health Care Management System Using Big Data Technology

Ravindra Ch, G Rajesh, Annapurna G, Ch Swetha, M.Ashish Reddy, G.Goutham Krishna
Guru Nanak Institutions Technical Campus, Hyderabad.

Abstract – Generally Automation plays an important role in the global economy and in daily experience. The Automated Healthcare Management System is an automated system that is used to manage patient information and its administration. In existing system challenges at large scale performing large-scale computation is difficult. To work with this volume of data requires distributing parts of the problem to multiple machines to handle in parallel. Whenever multiple machines are used in cooperation with one another, the probability of failures rises. In a single-machine environment, failure is not something that program designers explicitly worry about very often: if the machine has crashed, then there is no way for the program to recover anyway. This paper addresses the problem of data quality in electronic patient records using a computerized patient records report system with Apache HIVE and abstraction of Map reduce of big data technology. We analyzed which patient is spending more money than the others with the Map reduce. We got the data to be processed from traditional system to Hadoop via ETL's. We organized this with Oozie scheduler in Hadoop. The data what you are going to analyze is an semi-structured data. After uploading their data to cluster anyone can access them again provided they got to be in the cluster or can also use virtual machines that contain the right software to analyze them without any need for conversion.

Index Terms – AHMS, Oozie, HIVE, Mapreduce,ETLs,Hadoop.

1. INTRODUCTION

Generally Automation plays an important role in the global economy and in daily experience. Engineers strive to combine automated devices with mathematical and organizational tools to create complex systems for a rapidly expanding range of applications. The Automated Healthcare Management System is an automated system that is used to manage patient information and its administration. It is meant to provide the Administration and Staff, with information in real-time to make their work more interesting and less stressing.

2. EXISTING SYSTEM

In existing system challenges at large scale performing large-scale computation is difficult. To work with this volume of data requires distributing parts of the problem to multiple machines to handle in parallel. Whenever multiple machines are used in cooperation with one another, the probability of failures rises. In a single-machine environment, failure is not something that

program designers explicitly worry about very often: if the machine has crashed, then there is no way for the program to recover anyway.

3. LITERATURE SURVEY

The Existing Management System at Murab Hospital, Ilorin, Kwara, Nigeria.

The information flow used is a one directional system where the receptionist refers patient to doctors, doctors referring patients to the pharmacist either in or out patients and the same way out. The system that is currently being used in the hospital is entirely manual. When a patient requests drugs from the staff, all the information is recorded manually from the drug dispenser (Pharmacist). Similarly when the supplier delivers drugs all the information from the dispenser to the account on drugs is recorded manually. The following are the weaknesses of the current system at the hospital:

1. The hospital staff finds it tiresome and time consuming when computing patient data, drug supplier and staff Payment receipts and voucher cards this leads to delay in medical reports.

2. The hospital Administration currently uses health record files for storing patients and drug supplier's information. This system of information storage is susceptible to security problems such as illegal modification and update of records.

3. The Staff usually waste a lot of time in retrieving data.

4. The paper work reduces the efficiency of the System.

DESCRIPTION:

Hospitals can also be regarded as organizations based on high technology and information intensive processes.

According to Lawrence and Dyer, such organizations are not hierarchically structured bureaucracies, but are often based on democratic control mechanisms with institutionalized stakeholder influence in decision processes. A survey under 2752 European hospital managers indicates that technology can substantially influence hospital activities and services. It is also expected that health care budgets and funding will depend significantly on sophisticated patient and diagnosis classifications.

The use of IT in diagnostic and treatment processes will add to the development of networks of clinical, hospital and health care processes (Smith and Gert van der Pijl). Healthcare management is a growing profession with increasing opportunities in both direct and non-direct care settings.

As defined by Buchbinder and Thompson, direct care settings are those organizations that provide care directly to a patient, resident or client who seeks services from the organization. Non-direct care settings are not directly involved in providing care to persons needing health services, but rather support the care of individuals through products and services made available to direct care settings. The construction of medical information is important to improve the hospital medical care capability, the management decision-making level of health and the hospital operational efficiency.

Nowadays, comprehensive hospital information services and management platform have been established, centering on electronic medical records and clinical pathway. The establishment and use of these information systems played an important role in improving the degree of patient satisfaction, enhancing hospital efficiency and healthcare quality, protecting the safety of healthcare, and reducing healthcare costs.

4. PROPOSED SYSTEM

The proposed system addresses the problem of data quality in electronic patient records using a computerized patient records report system as an example. Physicians extracted five parameters from a traditional free text report and encoded these parameters thus producing a computer processable report. The proposed system is divided into Receptionist's module, Doctor's module and Pharmacist's module.

5. MAPREDUCE ALGORITHM

Map Reduce works by breaking the processing into two phases: the map phase and the reduce phase. Each phase has key-value pairs as input and output, the types of which may be chosen by the programmer. The programmer also specifies two functions: the map function and the reduce function.
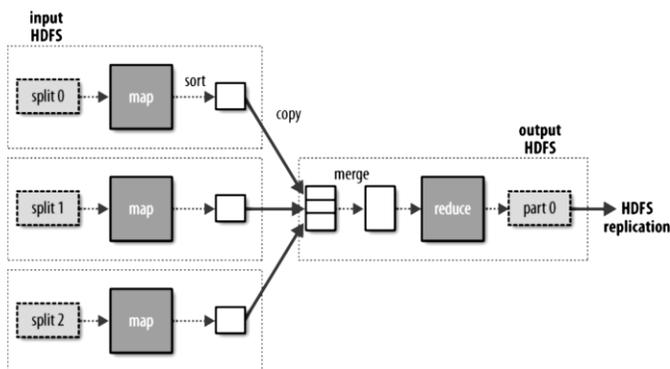

Fig 1. Map and Reduce
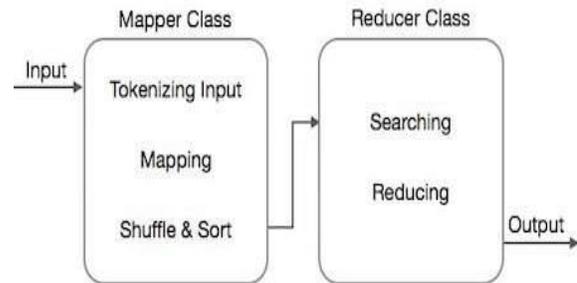
Fig 2. MapReduce Diagram

The MapReduce algorithm contains two important tasks, namely Map and Reduce.

- The map task is done by means of Mapper Class

- The reduce task is done by means of Reducer Class.

Mapper class takes the input, tokenizes it, maps and sorts it. The output of Mapper class is used as input by Reducer class, which in turn searches matching pairs and reduces them.
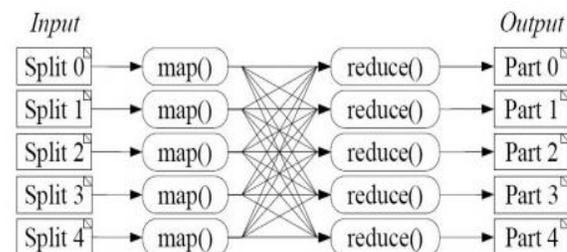

Fig 3. MapReduce Block Diagram

MapReduce implements various mathematical algorithms to divide a task into small parts and assign them to multiple systems. In technical terms, MapReduce algorithm helps in sending the Map & Reduce tasks to appropriate servers in a cluster.

These mathematical algorithms may include Sorting, Searching, Indexing and TF-IDF.

6. IMPLEMENTATION

Receptionist's module, Doctor's module and Pharmacist's module.

Map Reduce Algorithm is a programming model for data processing. The model is simple, yet not too simple to express useful programs in Hadoop can run Map Reduce programs written in various languages; in this chapter, we shall look at the same program expressed in Java, Ruby, Python, and C++. Most important, Map Reduce programs are inherently parallel, thus putting very large-scale data analysis into the hands of anyone with enough machines at their disposal. Map Reduce comes into its own for large datasets, so let's start by looking at one.

Patient Module:

CREATE DATA ACCOUNT -A Patient ID is assigned for new patients

VERIFY USER - F*or an* existing patient, "Patient ID'' is verified to check for the validity of the account

PRG (Payment Receipt Generator) **-** this assigns a receipt from the voucher generated by the doctor.

Search Engine: To search for the patients data (both payment records and also account authenticity)

Inbox: an advanced feature which will manage messages received from the medical superintendent.

Doctor's module (Lab Module Manager):

CREATE DATA ACCOUNT (For new patients, a "Unique No" is assigned for medical records on behalf of the patient.

VERIFY USER: Verifying "Registration No" is assigned to the patient if coming as existing patient in the hospital. Just as to check for the validity of the account)

HRE (Health Record Entry) a link or a function where patient's records will be inputted either inpatients or outpatients

(LMM) Lab Module Manager takes care of common laboratory test performed in the hospital

DPE (Drug Record Entry) a link or a function where doctor can input patients" drugs prescribed *WBE* (Ward& Bed Entry) A link which takes care of how beds are assigned and also change of rooms occur

PVG (Payment Voucher Generator) this will assign a voucher No so as for information of Number to flow well in the hospital

DMG (Discharge Manager) this assigns only for Inpatients and has expected date of leaving and also the date left will also be inputted into the system.

Search Engine: first field: search "With Patient Id" and choose a subdirectory for records to search.

**Inbox**: an advanced feature which will manage messages received from the medical superintendent.
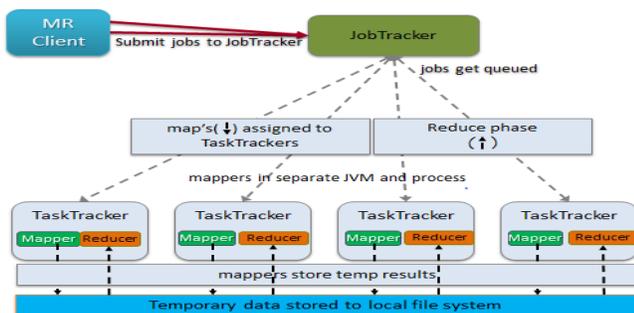
Fig 4. MapReducer Job Process
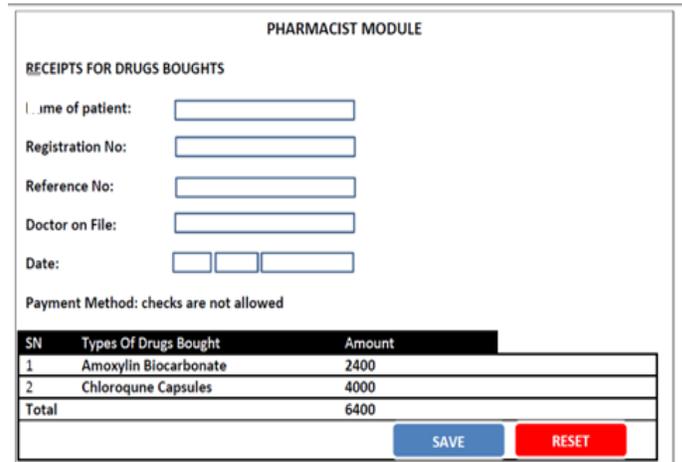
Pharmacist's Module (Receipts for Drugs bought)

Enter New Stock (ENS): This is a link which is strictly meant for entering new stocks.

CREATOR (RECEIPT CREATOR): A tabular form of entering drugs bought from the pharmacist and also assigning amount for it.
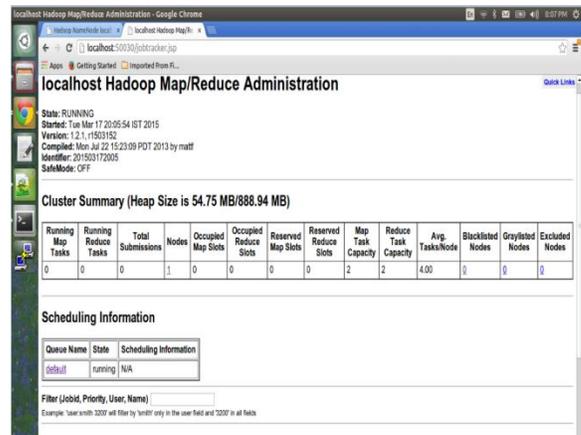
Sales record Inventor (S.R.I): This is a rendering functions of all sales made in a day which will be a search box that will be queried with date and also accounts for the amount made in a day. Also searching for Stock history will be also be a subcategories under the search record inventor.

Inbox: An advanced feature which will manage messages received from the medical superintendent.

Map Reduce Works

Run a Map Reduce job with a single method call: submit () on a Job object (note that you can also call waitForCompletion (), which will submit the job if it hasn't been submitted already, then wait for it to finish).1 this method call conceals a great

deal of processing behind the scenes. This section uncovers the steps Hadoop takes to run a job.
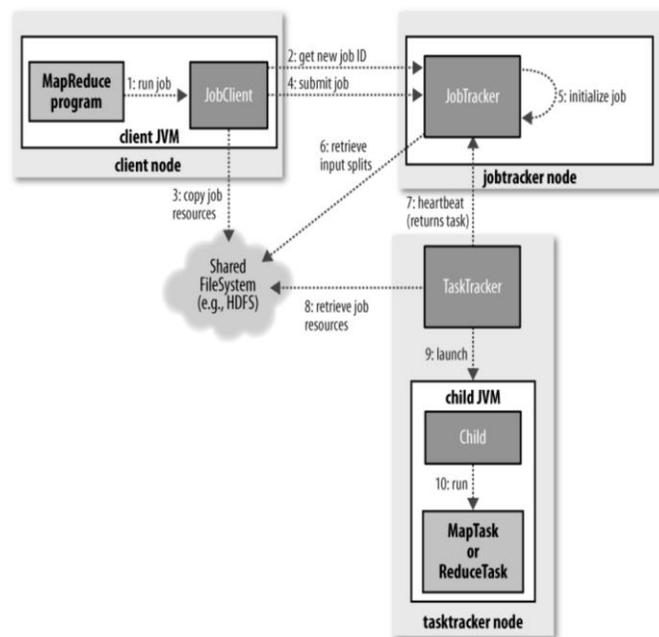

Figure 2.3: Map/Reduce Works

7. SYSTEM ARCHITECTURE

The systems architect establishes the basic structure of the system, defining the essential core design features and elements that provide the framework for all that follows, and are the hardest to change later. The systems architect provides the architects view of the users' vision for what the system needs to be and do, and the paths along which it must be able to evolve, and strives to maintain the integrity of that vision as it evolves during detailed design and implementation.

8. CONCLUSION

In this paper, we presented Automated Healthcare Management System is a project developed with Apache Hive, an abstraction of Map reduce. The data what you are going to analyze is a Semi-structured data. Computerized HMS has been developed. The system solved the problems associated with the existing manual system. Security is also enhanced since access to the system requires authentication. However, the system does not alert the pharmacy of the expiry date of drugs. Also, departments such as security and assets are not included in the design. Therefore, developing an HMS that can alert the pharmacist of the expiry date of drugs at a given time and handle all departments in the hospital will be an attractive research in future.

9. FUTURE ENHANCEMENTS

With the rise of Apache Hadoop, a next-generation enterprise data architecture is emerging that connects the systems powering business transactions and business intelligence. Hadoop is uniquely capable of storing, aggregating, and refining multi-structured data sources into formats that fuel new business insights. Apache Hadoop is fast becoming the defacto platform for processing Big Data.

Hadoop started from a relatively humble beginning as a point solution for small search systems. Its growth into an important technology to the broader enterprise community dates back to Yahoo's 2006 decision to evolve Hadoop into a system for solving it's internet scale big data problems. Eric will discuss the current state of Hadoop and what is coming from a development standpoint as Hadoop evolves to meet more workloads.

REFERENCES

[1] Bittencourt, L.F. and Madeira, E.R.M. "A Performance-Oriented Adaptive Scheduler for Dependent Tasks on Grids," Concurrency and Computation: Practice and Experience.

[2] Caron, E. Chis, A. Desprez, F. and Su, A. "Design of Plug-in Schedulers for a GRIDRPC Environment," Future Generation Computer Systems, vol. 24, no. 1, pp. 46-57.

[3] Dinda, P.A. and O'Hallaron, D.R. "Host Load Prediction Using Linear Models," Cluster Computing, vol. 3, no. 4, pp. 265-280.

[4] Dinda, P.A. "Design, Implementation, and Performance of an Extensible Toolkit for Resource Prediction in Distributed Systems," IEEE Trans. Parallel and Distributed Systems, vol. 17, no. 2,b pp. 160-173.

[5] Eddy Caron , Andreea Chis , Frederic Desprez , Alan Su (November 2011)"Plug-in Scheduler Design for a Distributed Grid Environment".

[6] Liang Hu, Xi-Long Che, (2012) "Online System for Grid Resource Monitoring and Machine Learning-Based Prediction" IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 23.

[7] Massie, M.L. Chun, B.N. and Culler, D.E. "The Ganglia Distributed Monitoring System: Design, Implementation, and Experience," Parallel Computing, vol. 30, no. 7, pp. 817-840.

[8] Peter Dinda, A. and David R. O'Hallaron (July 2012) " AN Extensible Toolkit for Resource Prediction in Distributed Systems" School of Computer Science Carnegie Mellon University Pittsburgh, PA ,15213.

[9] Sam Verboven, Peter Hellinckx, Frans Arickx and Jan Broeckhove (2011) "Runtime Prediction based Grid Scheduling of Parameter Sweep Jobs" University of Antwerp Antwerp, Belgium.

[10] Sodan, A.C. Gupta, G. Han, L. Liu, L. and Lafreniere, B. "Time and Space Adaptation for Computational Grids with the ATOPGrid Middleware," Future Generation Computer Systems, vol. 24, no. 6, pp. 561-581.

[11] Swany, D.M. and Wolski, R. "Multivariate Resource Performance Forecasting in the Network Weather Service," Proc. ACM/IEEE Conf. Supercomputing, pp. 1-10.

[12] Vetter, J.S. and Reed, D.A. "Real-Time Performance Monitoring, Adaptive Control, and Interactive Steering of Computational Grids," Int'l J. High Performance Computing Applications, vol. 14. no. 4, pp. 357-366.

[13] Waheed et al., "An Infrastructure for Monitoring and Management in Computational Grids," Proc. Fifth Int'l Workshop Languages, Compilers and Run-Time Systems for Scalable Computers, vol. 1915, pp. 235-245.

[14] Wolf, F. and Mohr, B. "Hardware-Counter Based Automatic Performance Analysis of Parallel Programs," Proc. Conf. Parallel Computing (ParCo '03), pp. 753-760.

[15] Hospital information systems definition given by US based". Consultant.com. Retrieved 2012-04-15.

[16] "Policy and Procedure Management Systems for Hospitals (2012)". PolicyStat LLC. 2012-07-18. Retrieved 2012-07-18.