WIRELESS TRAFFICDENSITYCONTROL USING SENSOR

Kirubasankar, Dillibabu, B.V.SANTHOSH KRISHNAN

Department of Electronics and communication Engineering Department, Velammal Institute of Technology Chennai (TN), INDIA

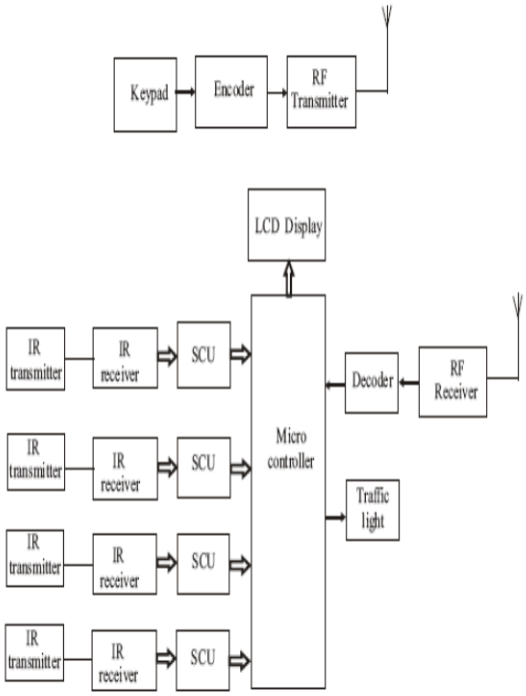
Abstract-Due to the heavy traffic in the road many people are suffering. Even though there is traffic signal available at various places it is not easy to control the crowd. In order to avoid traffic problem heavily on the road, we designed and developed a system so called as traffic density controller with wireless control. In the block diagram we will be having the components like IR sensor (4) on each side, LCD display, a Microcontroller, RF transmitter and receiver, key pad, 12 LED driver circuits and 12 LED's. The IR sensors, both the transmitter and the receiver are placed on either side of the road. The IR sensor senses how many vehicles are crossing on the corresponding side. All the sensors send the corresponding output signal to microcontroller. The LCD display is used to indicate the side through which the vehicle has to be passed.

Keywords- IR sensor (4) on each side, LCD display, a Microcontroller, RF transmitter and receiver, key pad, 12 LED driver circuits and 12 LED's

I. INTRODUCTION:

Due to heavy population in the countries such as China, India vehicle density is more compared to the other countries. Because of high density traffic we have to shift over to a remedy that by using the wireless remote access traffic control through RF transmitter and receiver which is sensed by IR transmitter and receiver. The microcontroller compares which road side is the highest density (i e, more vehicles) and it delivers the corresponding signal to LED driver circuits. The LED driver circuit enables the LED to display for particular time depends on traffic density. When this type of system is followed, we are sure that we can control the traffic to certain extent to certain extent. From this system, we can free up the traffic in the VIP's traveling road by wireless. There is a key pad connected remotely to control that. For each key present in that the corresponding road will be free up and other roads are blocked. The traffic density controller system will be very useful for nowadays and moreover it's very easy to implement this system. It will be more helpful to the public as well as the traffic department.

Block Diagram:



REFERENCES

[1]. MILL MAN J and HAWKIES C. "INTEGRATED

[2]. ELECTRONICS" MCGRAW HILL, 1972

[3]. ROY CHOUDHURY D. SHAIL JAIN, "LINEAR INTEGRATED CIRCUIT", New Age International Publishers, New Delhi, 2000"THE 8051 Microcontroller AND EMBEDDED SYSTEM" by Mohammad Ali Mazidi.

WEBSITES:

1. http://www.atmel.com/

2. http://www.microchip.com/

3. http:// www.8052.com.com/

A Study on Big Data and its Traffic Management – Separation Based on Varieties of Data

Naveen Kumar, S.Vijayabaskaran and S.Sankari

Department of Information Technology, Sri Sai Ram Engineering College, Chennai-600 044.

***Abstract* - Big data is a massive volume of data sets which are difficult to process by traditional techniques within elapsed time. Big data is associated with its properties of five V's. Big data is large amount of data that requires new technologies to abstract values for capturing and analysis process. Big data-due to properties put forward many challenges. Big data-new upcoming technology in the market and also that it can bring many benefits to many organizations also that many challenges are associated in bringing and adopting to this technology are brought into light. This paper says about the importance of big data and how it is being managed using indexing and key generation. Due to emerging technology, the traditional system adapt to changes in the concept of science into big science. The tool that is being implemented in big data is HADOOP with the concept of map reduce and the problems that are faced with hadoop is also discussed. This paper concludes with the good big data practices that are to be followed. The proposed system provides the 4 major components as switch, key-generator, and subnet switch and map workers that uses the concept of indexing and map reducing to manage the Big Data in secured way.**

***Keywords*: indexing, key generation, Hadoop, map workers, map reducing.**

I. INTRODUCTION:

Big Data is a massive volume of datasets difficult to store and process. Hadoop is a distributed Frame work for processing the Datasets and it acts as the solution for the Big Data. In Big Data, Data's are stored in peta bytes. It is very difficult to handle such large amount of data because the volume is increasing rapidly in comparison to the computing resources. Big data specifies the size of data and is not considering (or) putting attention to other properties. There are three methods to represent the data being available (i.e.) structured data, Semi-Structured data and unstructured data.

Big data defined with the following properties associated with it:

1.1 Variety

Data produced may not look to single category as it also includes the semi-structured data that are from resource like web pages, web log files, social media web sites, e-mail, documents and sensors devices. All this data is consists of raw, structured, semi structured and even unstructured data so in order to handle these different categories of data emerges this new technology of big data.

1.2 .Volume

The word big data, itself defines the large volume. Now-a-days data's are present in peta bytes and in future it may be increased to Zetta bytes. The Big data system is mainly used to handle the data produced everyday by the social websites which are in the order of terabytes.

1.3. Velocity

Velocity of big data defines the speed of data coming from various sources. The speed not only specifies the incoming data but also specifies to the data that flows out. Big data concept analyzes the data that are constantly in motion.

1.4. Variability

As the word variable specifies to the inconsistency of data flows. Weightage of data refers social media when the data loads appear to be in peak with that of certain events being occurred.

1.5. Value

The user run queries to obtain the results according to their requirements. The technical issues in handling data are storage and processing. The main challenges in handing big data are

- Design the systems to handle the large amount of data efficiently and effectively.
- Filtering the most important data from collected huge amount of data.

II. CHALLENGES AND ISSUES ON BIG DATA

2.1. Privacy and Security

It is the most important issue with big data is sensitive, conceptual, technical and legal. Problem occurs in financial transaction. When we are combining the personal information of a person with large data sets leads to interference and if we used in order to add value to the business of organization. This leads to creating insights to their lives. This big data used by law enforcement will increase the chunks of certain tagged people.

2.2.Storage and Processing Issues

The storage available in traditional systems is not enough for storing such large amount of data. Storage can be done at cloud is the one solution but uploading it to cloud is more complex and it also takes large amount of time and it is not secure. Big data requires all data collected and then linking it in a way to extract information. Transportation of data from storage point to processing point is done in two ways:

- Processing at storage point and sending only results.
- Transport only the data which is needed and important.

A. *Technical Challenges*

1. Fault Tolerance

Usually big data is intended that when failure occurs, the damage should be done within acceptable. Everytime each process should be started from beginning. The main task is to reduce the probability of failure to “acceptable” level. This leads to increase of cost.

ii) Scalability:

This issue of big data leads us to cloud computing because in order to handle such large data. Data transfer is expensive and also brings various challenges. This deals with running the program and executing the data so that we can reach our goal. It also needs to deal with system failure in efficient way which normally happens in system with large clusters.

iii) Handling unstructured data

Unstructured data is not in an organized way so that we cannot handle easily.

Example: meeting, social media interactions, handling PDF files, fax transfers etc.. It is costly. Data can't be converted to structured data.

III. TOOLS AND TECHNIQUES

3.1 Hadoop

Hadoop is open source software. It has many small sub projects. There are two types being involved, file system – for Hadoop and programming paradigm – for map reduce.

Though the storage capacities of disks are improved massively but reading the data from it has not improved much. This reading is taking much larger time, to eliminate this reading from multiple disks once and using hundreds of disks seems waste. If there are one hundred datasets, each of which is one TB providing shared access to them is also a solution. This method increases chances of failure while using multiple disks.

This can be managed by replication of data. The process of creating copies requires reading from multiple disks. All this problems can be handled by hadoop. Combining problems is can be handled by map reduce programming paradigm. There are different components of Hadoop is shown in Figure1.

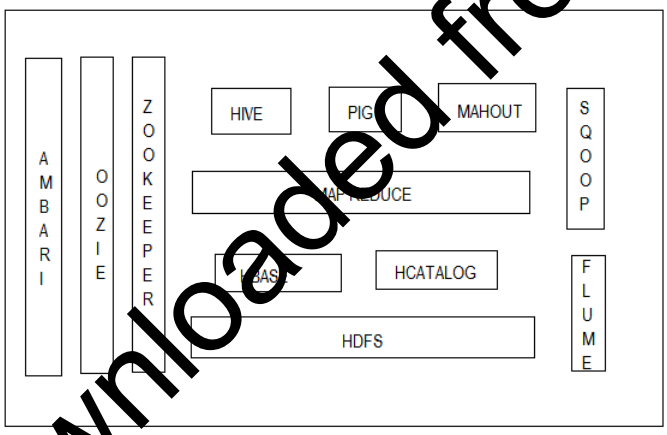


Figure1 Components of Hadoop

IV. MANAGING BIG DATA:

There are number of solutions to handle big data but the efficient and easier method is by using hadoop. It is a processing infrastructure which consists of hadoop kernel. HDFS and Map Reduce.

Hadoop Components

4.1. Hadoop Distributed File System (HDFS)

HDFS is used to store the huge amount of data, using noSql (MongoDB,Cassendra) Databases. It manages the computational resource effectively, less expensive and less time consumption. This is designed for storing very large files with streaming access patterns, running on clusters on commodity hardware. Its block size is larger than normal system i.e., 64MB (default). The large size is to reduce more disks. Clusters in hadoop has two types of nodes, they are name node which is the master and the number of Data nodes which acts as workers.

i) Name node

The name node manages name space and also manages file system and metadata of all files.

ii) Data node

Data node consists of blocks as per the instructions of name node and it returns the list to name node. Without name node it is not possible to access the file, so maintaining it is very important. The drawbacks in this are low latency, multiple writers, and arbitrary file modifications.

4.2. Map Reduce

Map reduce is the programming paradigm. This is used for processing the data from the database (HDFS). Here the operation of distributing the code or data to be processed to the different servers and the processed data is sent back to the central node. This does two tasks of Map task and Reduce task.

i) Map Task:

In the map task, input is from distributed system. This generates a sequence of key value pairs from input and this is done with the code written for map function. These values are collected by master and assign to the reduce keys after sorting. The reduce task combine all the values associated with a key.

Worker handles map task (map worker) and reduce task (reduce worker). Master assigns the workers with a task and reduces tasks; this is done by user program. After completing this task function, the workers inform to masters then it re-assigns with the new task. The failure of computer node is detected by master as it periodically pings the workers.

Drawback in this is the one that all the tasks assigned to the node is restarted even it is completed. The status of each of these map tasks is set to idle by master.

V. EXISTING SYSTEM:

5.1. Grid Computing:

Grid computing is one of the distributed computing. This is introduced in early 1990's along with the SFTI (Search for Extra Terrestrial Intelligence). This is created in order to save, distribute and analyze. Grid center is represented by number of servers that are inter-connected by high speed network. Each server plays one (or) many roles.

Advantages of this are mainly deals with high storage and processing power.

Grid Center:

There are three elements in grid center being used; they are computing Elements (CE), Storage Elements (SE) and Worker nodes (WN).

Computing Elements is used to manage resource of grid and manages jobs.

Storage Elements is used in storage and data transfer services.

Worker Node are servers that offers the processing power.

User Interface is also present. In order to access Grid at higher level, user must be certified. Virtual organization is the set of peoples defines rules for accessing the grid. It has work load management, which tracks of available Computing Elements for user jobs. User access User Interface through secure shell (SSH) and he receives priority certificate. The user then sends the job written in Job Description Language (JDL) and PYC to WMS. This WMS checks availability of Computing Elements and assigns Worker Nodes to process it. Worker Node completes job and send to WMS and state of job to Computing elements.

Hadoop with HPC and Grid Computing:

In grid computing data's are separated among clusters and they are connected through a common distributed file system. Here the locating of data among the file system clusters is difficult. Traffic separating and sending large data in a small network is difficult. In Hadoop, the Map reduces component makes use of data locality property, where it collaborates the data with the compute node itself and so access is fast. Grid basically uses Message Passing Interface (MPI), though it provides greater control to user, but the user needs to control the flow of data.

Hadoop's map reduce operates only at higher level where the data flow is implicit and programmers just think of the keys and value pairs. This map reduce operates on shared nothing architecture i.e., tasks independent. Implementation of map reduces itself give information about Failed tasks and reschedules it. This checking and pointing MPI needs to be programmed at first.

i) Volunteer Computing Technique:

In this Volunteer Computing technique the work is broken into chunks called work units which are sent over the computers all over world to be analyzed. After analysis it is checked for accuracy, by sending into three different machines. If at least two matches, then data is safe.

But in case of Map Reduce, it is designed to run jobs last minutes or hours or trusted hardware running in single data center with high aggregate between connections.

VI. PROPOSED SYSTEM:

The analysis is the major task in big data, because the big data is generally from large institutions, business organization, social media etc., According to statics which was taken at 2012 the amount of big data is mostly from America which is about 34% and in India only 4% of world. Amount of big data is 2.8 Zeta bytes in 2012.

If this process continues, in 2020 the amount of big data may be 50 times larger than today and India contributes 6% of big data. With this big data only 36% of all the data will be useful and currently we have analyzed only 0.5% till 2012. With the use of big data we can gain much knowledge and wisdom.

In our system we are going to analyze the big data in an efficient way be separating relevant data initially, then we are analyzing it.

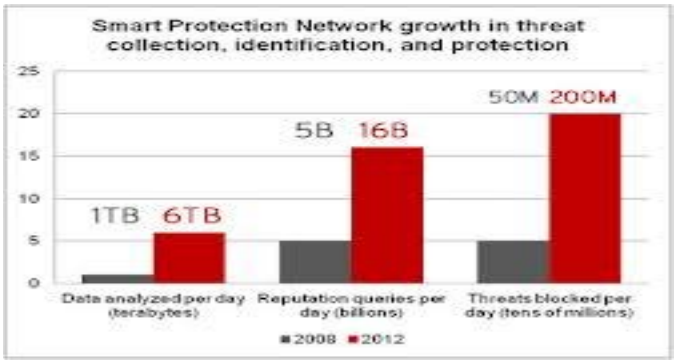


Figure2. Statistics on analyzing the big data for the year 2008-2012

Our system consist of 4 major components

- Main- Switch
- Key- generator
- Sub-net Switch (master)
- Map-workers

Functions:

i) *Main switch*

Main switch takes input as data of any kind and it separates the relevant data into different kinds. Varieties of data's are

1. Real-time
2. Social media
3. Business Organization
4. Other data

This main switch is used to separate relevant data. This separation is based on centre law by looking at a data and separates it to different kinds. We take any one relevant data such as a Real-time data is chosen and then we use the hadoop technique which has a property called map reduce.

This map reduce concept has 2 main functions

a. *Indexing*: Indexing is just like a reference to a data after the key is generated. This indexing is just works like a book index reference.
b. *Key generation*: This has an instrument which generates key from the input data and send it to a Subnet-Switch.

ii) *Key Generator*

The key generator takes input as data and it delivers keys to subnet switch, after storing the key value n index. The key Generation for the user data is by Kerberos for authentication and encryption of data for securing the data in Financial Transaction.

iii) *Subnet Switch*

In Hadoop, Oozie component acts as a job scheduler to the map worker. This subnet switch works like a master where it does not do any work other than watching the other workers works properly or not. This just assigns works to map workers. A large works is sub divided and is given to map workers. It just checks whether all workers did their job or not. In case any crash in map worker the jobs will be reassigned to it from ports even if the job is nearly completed.

iv) *Map workers*

Map workers just work with the data key which was assigned by the subnet switch. After completion of work it sends the result of job to subnet switch and gets other jobs that are to be processed.

Our concept mainly focuses on bottom up approach because when going for bottom the data amount will be less so we can analyze it easily. Top down approach fails in big data because analyze on large data is not possible. The proposed system is as given in Figure 3.

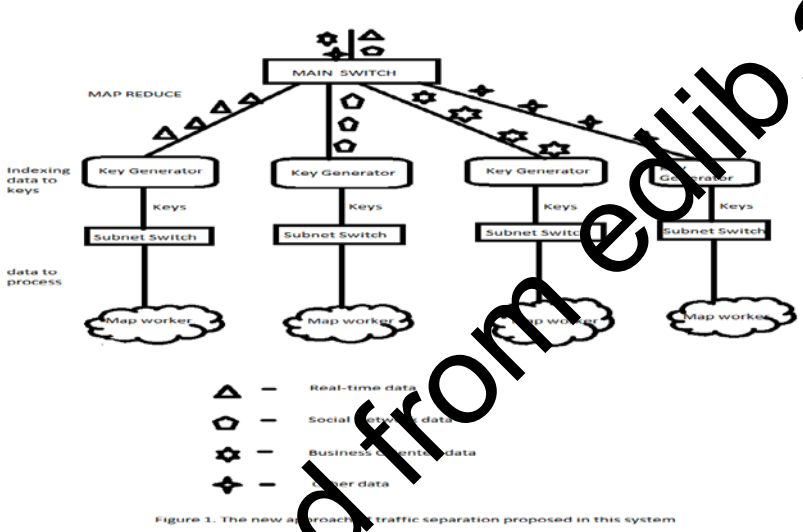


Figure 1. The new approach of traffic separation proposed in this system

Figure3. The new approach of traffic separation proposed in this system

VII. BIG DATA GOOD PRACTICES

- Creating dimensions of all the data being stored.
- All data's should have a durable keys
- Structured and unstructured data are analyzed together.
- Importance of big data to consumers is growing.
- Data quality needs to be better.
- The scalability has certain limits of storing the data.
- Investment in quality and metadata reduce processing time.

CONCLUSION

In this paper, a framework for the analysis of large amount of data (Big Data) with the use of map reduce concept is developed. This developed framework brings out a better analysis of data over the previous methods. This method uses the map reduce concept for evaluation of related data in an effective way which provides indexing for the easy reference and also providing better analysis using map worker. This system ensures that the data is analyzed in a best way to use it, because the previous analysis method does not give analysis more than 0.5%. Our concept not just only makes indexing, it also generate keys for the given input data and this is just indexed to subnet switch and then to map worker. Hope the proposed system would achieve the best analysis of the big data. This paper uses Hadoop technologies concept for helping in addressing a best analysis.

REFERENCES

[1] Stephen Kaisler, Frank Armour, J. Alberto Espinosa, William Money, "Big Data: Issues and Challenges Moving Forward", *IEEE, 46th Hawaii International Conference on System Sciences,* 2013.

[2] Sachchidanand Singh, Nirmala Singh, "Big Data Analytics", *IEEE, International Conference on Communication, Information & Computing Technology (ICCICT),* Oct. 19-20, 2012.

[3] Martin Courtney, "The Larging-up of Big Data", *IEEE, Engineering & Technology*, September 2012.

[4] Matthew Smith, Christian Szongott, Benjamin Henne, Gabriele von Voigt, "Big Data Privacy Issues in Public Social Media", *IEEE, 6th International Conference on Digital Ecosystems Technologies (DEST),* 18-20 June 2012.

[5] T. Benson, A. Akella, and A. Maltz, "Network Traffic Characteristics of Data Centers in the Wild", IMC'10, Novem1-3, 2010, Melbourne, Australia.

[6] S. Kandula, S. Sengupta, A. Greenberg, P. Patel and R. Chaiken, " The nature of Data Center Traffic: Measurement &Analysis", IMC'09 November 4-6, 2009, Chicago, Illinois, USA

[7] T, Benson, A. Anand, A. Akella and M. Zhang, "Understanding Data Center Traffic Characteristics", WREN '09, August 2009, Barcelona, Spain.