Research Article

Literati Journals
Empowering Researchers

A Survey on Analytical Architecture of Real-Time Big Data for Remote Sensing Applications

Raghavendra,[1]* Ashwinkumar U M

Abstract: The remote senses generate very large amount of real time data from the Satellite or from the Aircraft with the help of the sensors. Now a day there is a great demand added to the real time big data for remote sensing applications, these data has to be processed and extract the useful information can lead to computational challenges. From these above mentioned factor need to design an architecture that can supports both offline as well as the real time data. In this paper we will discuss the proposed architecture for the remote sensing application. The three main units comprises the proposed architecture the three units are First, Remote sensing data acquisition unit (RSDU) takes the data from the satellite and sends to the Base Station, where processing starts in this unit. Second, Data processing unit (DPU) is the main role in the architecture, the real time data will process efficiently by filtering, load balancing and parallel processing and Third, Data Analysis and Decision unit (DADU) this unit is responsible for the storing the results and generates the decision based on the results of the data processing unit. The proposed architecture can store raw data for the analysis of the offline data when required.

Asian Journal of Engineering and Technology Innovation
Volume 4, Issue 7
Published on: 7/05/2016

Cite this article as: Raghavendra and Ashwinkumar U M. A Survey on Analytical Architecture of Real-Time Big Data for Remote Sensing Applications. Asian Journal of Enginnering Technology and Innovation, Vol 4(7): 120-123, 2016.

INTRODUCTION

As of late, a lot of enthusiasm for Big Data has risen, for the most part determined from across the board number of research issues emphatically identified with genuine applications and systems. Day by day the data is increasing very large volume from social media, videos, emails, online transitions, logs, Scientific data, mobile phones, Remote sensors and other applications. These data store in the database and grow rapidly with a massive amount becomes complicated to store, process, manage and analyze.

The advanced technology in the big data give a way to the remote data, which can be collection, managing, analyzing and processing. Recently designed remote sensors that are used for the earth observatory streams the data continuously and generates large amount of data. Many of the work have been done in the different fields of remote sensing data from the satellite, such as gradient based edge detection [4], change detection [5] and etc. This paper is concentrated on the high speed continuous real time streaming data or large amount of offline data i.e. Big data, this leads to a new challenge. Such consequences for scientific understanding of transformation of the remote sensed data is critical task [6], [3].

[1]Reva Institute of Technology and Management, Rukmini Knowledge Park, Kattigenahalli, Yelahanka, Near Border Security Bustop, Bengaluru, Karnataka-560064, India.
E-mail: raghavrao.mar@gmail.com
*Corresponding author

Data is collected from the remote sensors; these remote sensors generates a very large volume raw data this is also called as data acquisition. The collected data has no meaning in it, the sensor simply collects all the information. So the data need to be processed and filtered to extract the useful information from it. The main challenge in this is the data accuracy, the information that are generated by the remote sensors are not in the correct format for analysis. Now the data need to be extracted to pull the useful or meaningful data and converted into to the structured format for best analysis. Sometimes the data might be not clear or it may be erroneous too.

To address the above needs, the architecture is introduced, for the remote sensing big data. This architecture has the capacity to analyze both type of data, offline data as well as real time data. First, the data has to be remotely processed in the readable format of the machine then the useful data is sent to the base station of the earth for the further data processing. The earth base station processes 2 types of data one is offline data and the other is real time streaming data. The offline data are sent to the offline data storage device incorporation of the later usage of data. Where in the real time data, the data is directly processed to filtering and the load balancing server. Filtering extracts the meaningful or useful data from the big data and the load balancing will balance processing by distributing the real time data equally to the server. These filtering and the load balancing server will also improve the system efficiency.

Next, the data is directly sent to the data aggregation unit for comparison by analyzing and the decision server. The proposed method is implemented by the Hadoop framework using the map reduce programming by the data of remote sensing.

RELATED WORK

This section provides the detail summary of the previous work done in the remote sensing real time big data.

The digital world generating the high amount of the data continuously, current technology and the tools to store and analyze the large amount of data not an easy task, since it is not able to extract the needed data sets. So there is a need of an architecture that can analyze both the offline data as well as real time data sets. There is an influential benefit in the business enterprise by obtaining the required information from the Big data than sample data sets. Some of the areas that are described below where big data can play very important role.

Understanding the earth atmosphere or environment requires large volume of information or data gathered from different sources, such as air and water quality monitoring sensors, amount of oxygen, co2 and the other gases present in the air, remote access satellite for the observing the characteristics of the earth and so on. In the healthcare scenarios, there is large amount of the data about the medications, patients, medical history and other details gathered by the medical practitioner. The above mentioned data is very complex in nature, there is a chances of missing the important data.

Day by day the data becoming very large by social networking, online streaming, system logs, mails and remote data, it will be very difficult to compute massive amount of data. Main problem is how to store the large amount of data i.e. big data and what data is to keep and what data is to be discarded, extracting the useful data from the big data is the challenging task [2].

Most of the data is generated by the streaming data. In data stream model, the data will arrive at a very high speed and the algorithm have to process them. This data stream causes several challenges in design of the data mining algorithms. First, algorithm has to make use of less number of resources. Second, it can deal with data that can change over time. Resources are managed in an efficient and low cost way, by the green computing [7]. Green computing is the process or study to use the computing resources in an efficient way. Here, the problem is not only the scaling issue but also error handing, lack of structure, heterogeneity, privacy, visualization and timeliness.

The challenge is to design a high performance computing systems that can be able integrate resources from different location. Even though the cloud computing systems shown high level performance for RS applications, there are challenges still remaining regarding energy and the time consumption. The big challenge emerges when collecting and the managing Remote Sensing (RS) big data. The RS data are collected form spacecraft, airplanes, satellite and other sensing devices. Remote sensing data growing explosively, we have entered in the period of very high resolution, observation of the earth. Remote sensing data also considered as a “Big Data”. With the advance sensors we can take even high spatial resolution images, spectral resolution and also temporal resolution. The advancement in the technology of the computers and the remote sensing devices increases a massive growth remote sensing data [9]. The earth observatory data that is streamed from the spacecraft approximately around 1.7GB, this data is collected by single satellite and increased many terabytes per day. The global records of observatory data of the earth would exceed to one Exabyte, according to the OGC statistics.

Various standard format data sets of remote sensing are stored in structured files, the formats including ASCII, HDF, netCDF and so on. Different organization have different standard format of the data sets, different format of data has its own format libraries and operation interfaces.

Huge amount of data need to compute in an efficient way and only the useful information need to be extracted from the big data. So there was a need of the architecture for filtering the data, load balancing, aggregating and the decision analysis.

PROPOSED WORK

Even the big data has the variety technologies as in cloud computing. Recent arise of the big data architecture in scientific applications, many efforts applied toward the analytical architecture of the big data already found in the previous written works or literature. The proposed architecture, analyze the remote sensing big data in an efficient manner [1].

The architecture for the remote sensing big data in fig. 1, which has n numbers of satellites that used to take the earth images by the sensors or from the conventional cameras. They have divide the architecture of the remote sensing big data into 3 parts, 1) RSDU (Remote Sensing Big Data Acquisition Unit); 2) DPU (Data Processing Unit); 3) DADU (Data Analysis and Decision Unit). The working of these 3 units are described below.

RSDU (Remote Sensing Big Data Acquisition Unit)

Remote sensing encourages the growth of observatory system of the earth as cost efficient parallel data acquisition system to fulfil certain computational requirement. For efficiently analyzing big data there is a need of the parallel processing to process the big data in an efficient way. For this reason, the proposed method i.e. RSDU (Remote Sensing Big Data

Acquisition Unit) is introduced in the architecture of remote sensing big data, that collects the data from different satellite from the globe. There is a possibility that raw data received can be distorted by various atmospheric gases and the dust particle. We assume that the distorted or the erroneous data, satellite can correct. However, the remote sensing satellite uses the algorithm Doppler or SPECAN to make the raw data into the image format. The data is sent for further processing to the earth base station by direct communication link.

In this two types of data processing takes place one is offline data processing and the another one is the real time data processing. In offline data processing, data will be transferred for storage, to the data center by the earth base station. This data is used for the future analysis. In the real time data, the data is directly sent to the FLBS (Filtration and Load Balancing Server).

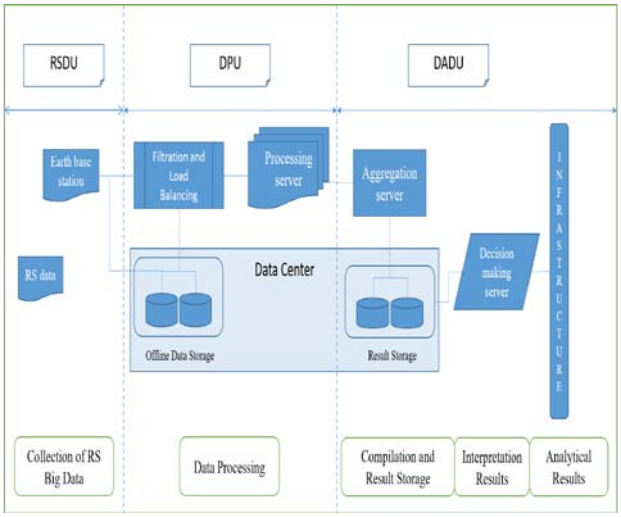

Figure 1: Analytical Architecture of remote sensing Big data

DPU (Data Processing Unit)

In DPU that is Data Processing Unit, it has two responsibilities, such as first, data need to be filtered by the filtration process. Second, balance the processing power by the load balancing server. Filtration recognizes or identifies the useful information, remaining data discarded of blocked. Hence, it improves the results of performance of the system. The load balancing server give the facility to divide the filtered data into parts and each part will be processed by the processing server. This load balancing and the filtration algorithm changes from analysis to analysis; example, if there is a need for only temperature data and the sea wave, then the needed data is filtered out and it is divided into parts.

Every processing server has its algorithm, to process the incoming segments of data from the filtration and the load balancing server. The processing servers performs some measurements, statistical calculations and makes other logical or mathematical operations to create the intermediate results from every segments of data. Since each processing servers executes the tasks in parallel and independently, hence the proposed system dramatically boosts the performance. The results obtained by every processing server are sent to further processing to the aggregation server for organization, compilation and storing.

DADU (Data Analysis and Decision Unit)

Data Analysis and Decision Unit has three major servers, such as compilation and the aggregation server, server to storage results and server to make decision. After the filtering process the data is ready for the compilation, in the data processing unit (DPU) the processing server sends part of the filtered results to the compilation and the aggregation server, since the results are not well organized and compiled form.

There is a need to organize the data or the results in proper form for further processing and storing. The proposed architecture supports different algorithm organize, compile and storm the results.

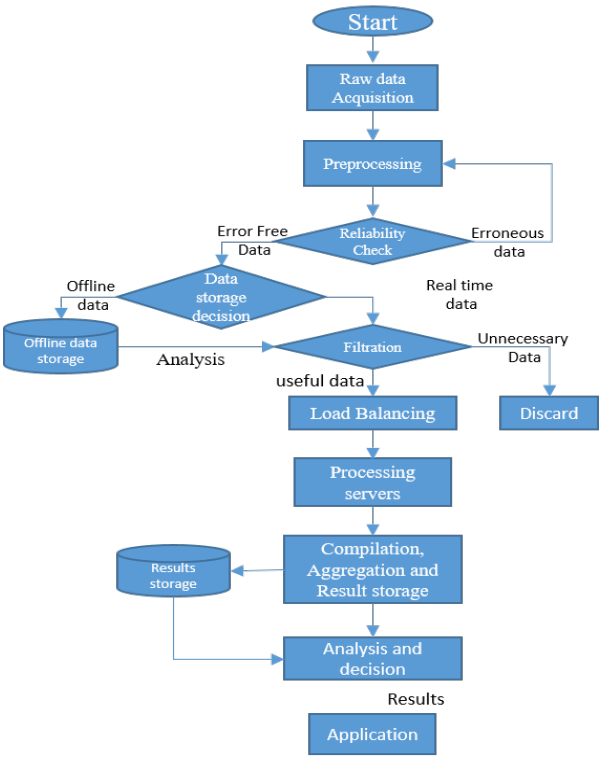

Figure 2: Flowchart for remote sensing big data architecture.

Aggregation server stores the results into the results storage this helps any other server to use it process at any time. DM (Decision making) server for making the decisions. The decision making server has decision algorithm, to make the various decisions. So any applications make use of these decisions to make their development at real time. The application can be any general purpose software, other social

networks or any business software that need decision making. The Figure 2 shows the flowchart for the proposed architecture.

CONCLUSTION AND FUTURE DIRECTIONS

The remote senses generate very large amount of real time data from the Satellite or from the Aircraft with the help of the sensors. Now a day there is a great demand added to the real time big data for remote sensing applications, these data has to be processed and extract the useful information can lead to computational challenges.

In this paper, we discussed the proposed analytical architecture of real time big data for remote sensing applications. The proposed architecture is designed in such a way that; it can analyze both the offline as well as the real time data in an efficient way. The proposed architecture for the remote sensing application. The three main units comprises the proposed architecture the three units are First, Remote sensing data acquisition unit (RSDU) takes the data from the satellite and sends to the Base Station, where processing starts in this unit. Second, Data processing unit (DPU) is the main role in the architecture, the real time data will process efficiently by filtering, load balancing and parallel processing and Third, Data Analysis and Decision unit (DADU) this unit is responsible for the storing the results and generates the decision based on the results of the data processing unit. In this paper, the remote sensing data analyzed by each proposed unit to make better decision making.

For future work, this proposed architecture can use to compute for more complex data for decision making at real time of earth observatory, such as fire detection, tsunami prediction, earthquake prediction, etc. The architecture need to make compatible for all applications for big data analysis.

REFERENCES

1. Muhammad Mazhar Ullah Rathore, Anand Paul, Bo-Wei Chen, Bormin Huang, and Wen Ji, "Real-Time Big Data Analytical Architecture for Remote Sensing Application," *IEEE journal of selected topics in applied earth observations and remote sensing*, 2015.
2. D. Agrawal, S. Das, and A. E. Abbadi, "Big Data and cloud computing: Current state and future opportunities," in *Proc. Int. Conf. Extending Database Technol. (EDBT)*, 2011, pp. 530–533.
3. A. Cuzzocrea, D. Saccà, and J. D. Ullman, "Big Data: A research agenda," in *Proc. Int. Database Eng. Appl. Symp. (IDEAS'13)*, Barcelona, Spain,Oct. 09–11, 2013.
4. A. Paul, J. Wu, J.-F. Yang, and J. Jeong, "Gradient-based edge detection for motion estimation in H.264/AVC," *IET Image Process.*, vol. 5, no. 4, pp. 323–327, Jun. 2011.
5. J. Shi, J. Wu, A. Paul, L. Jiao, and M. Gong, "Change detection in synthetic aperture radar image based on fuzzy active contour models and genetic algorithms," *Math. Prob. Eng.*, vol. 2014, 15 pp., Apr. 2014.
6. S. Kalluri, Z. Zhang, J. JaJa, S. Liang, and J. Townshend, "Characterizing land surface anisotropy from AVHRR data at a global scale using high performance computing," *Int. J. Remote Sens.*, vol. 22, pp. 2171–2191,2001.
7. R. A. Dugane and A. B. Raut, "A survey on Big Data in real-time," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 2, no. 4, pp. 794– 797, Apr. 2014.
8. E. Christophe, J. Michel, and J. Inglada, "Remote sensing processing: From multicore to GPU," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 3, pp. 643– 652, Aug. 2011.
9. Yan Ma, Haiping wu, Lizhe wang, Bormin Huang, Rajiv Ranjan, Albert Zonmaya, wei jie, "Remote sensing big data computing: Challenges and opportunities," *Article in press., October 2014.*