International Journal of Computer Science and Engineering Communications
Vol.4, Issue.2 (2016): Page.1314-1318.
www.scientistlink.com/ijcsec

Monitoring Network Traffic Based on Mapreduce in Simplified Data Processing Using Intrusion Detection System

Sukanya .D[1], Parvathi .S[2], Kumaresan .A[3],Soniya.N[4]
[1,4]PG Scholar, [2]Assistant Professor, [3]HOD,
Department of Computer Science and Engineering,
Department of Computer Science, SKP Engineering College, India,
dsukanyabe@gmail.com[1], 5680sparvathi@gmail.com[2],kummaresan@gmail.com[3], springsoni25@gmail.com[4]

Abstract- *Big data is having more key roles as infrastructure information e-pay and e-business social media in Internet booming due to the user conveniences and user benefits. Now a day Internet security problems is most of challenges for some security session. The privacy and security should be Reinforce via secured cloud computing and hadoop server. It Refer a comprehend security solution for attack detection and attack prevention. In this article research distributed clustering scheme and propose a cluster- based routing protocol for Delay Tolerant hadoop Networks (DTHNs). Due to the lack of continuously communicate among hadoop network nodes and possibility of error to estimate the nodal contact probability convergent and stability is become a major challenge in distributed clustering network. The basic idea is distributed group of hadoop nodes with quasi mobility pattern into a clusters, It can be interchanged data without overhead and load balancing then it also achieve scalable and efficient routing via distributed hadoop server .To detect and prevent the data from attacker. An exponentially weighted average moving time (EWAM) schema is worked for online update with nodal contact probabilities, which means prove the converge to true contact probability and it carried out to evaluate the efficiency and effectiveness of cluster based routing protocol.*

Keywords— *Big data, cloud computing, delay Tolerant hadoop Networks, distributed clustering, privacy and security*

1. INTRODUCTION

Big data play a more processing in business model and applications then it improve the affordable and faster communication with multiple users. Big data indicate to large and complex data set that are challenged to search, data creation, data privacy and security ,store analyse, and visualize .The main characteristics of big data is volume, velocity and Varity of data. volume it is a scalability of information to add multiple data or connect multiple users at the same time For example facebook .velocity of data is a speed of the data transaction , variety of data is format or types of data for example text ,audio, image, time sequences numerical .Generally speaking, big data is decreasing cost of storage then flexibility and cost effectiveness of data centres and developed then new framework such as hadoop.

2. RELATED WORK

2.1Monitoring Network Traffic

Most of the network contain a traffic analysis in the process of inter cluster mobility pattern communication. [1]Data summarization is a key distributing mining. Network host can be increasing the connectivity of the network it makes a major problem for network task manager to understood the nature of traffic flow and it carried in whole network.[2] Denial of service [DOS] this type of attack are aggrieved and managed intrusive behaviour of online server.

DoS attacks is mostly degrade the availability of the immolation it can be a router, host otherwise entire network. That impose drastic estimation task of the victimise by denial of service attack can detected system it use (MCA) analysis is multivariate correlation analysis based Dos attack detection system work the rule of anomaly based on the detection is recognize attack.

2.2 General Scheme for Intrusion Detection Systems

Normally virtual networks contain vulnerabilities leave loopholes for cyber intruders to exploit Cloud environment services and pretend threats to the privacy and security of big data. This intrusion detection system address the security issues, a variety of security schemes are introduced over the past years. It includes data encryption, user authentication, firewalls, access control, Data Leak Prevention System (DLPS). This complex network computing environment however

one can tuff find a single scheme that fits all cases. IDS having important role of security schemes .Their main aim is provide a security layer of self defence in against malicious user from cloud environment by sensing the known and unknown user and alerting the user from attacks. Impossible to prevent the all cyber attacks .This intrusion detection system contain more and more essential secured schemes.

3. INTRUSION DETECTION SYSTEMS FOR SECURE CLOUD ENVIRONMENT

Intrusion detection system is classified network based intrusion detection system and host based intrusion detection system. Host based IDS is efficient to handle the insider attack and user to root attack it try to gain root offer privileges to host and virtual machine. Network based IDS is monitoring network traffic flow in flooding attack. The firewalls can be block unknown network traffic packets according to the predefined perceptive they are inflexible to detect sophisticated intrusion try which means flooding attacks and network insider attack [3].IDS system is classified into misuse based on intrusion detection systems, it provide high detection with accuracy but vulnerable to all network zero rate intrusion and Anomaly based on intrusion detection system with suitable detection mechanism can be applied. it is a show trustable performance in the detection with zero day intrusion [4].The current network contain a multiple entry point with cloud environment [5] .This type of topologies is goal is improve the availability and accessibility of the network .which means leaves the security vulnerability for attacker to exploited used in most advanced infliction techniques .Distributed the network attack traffic volume to the different level of entry points. This type of network traffic behaviour at each level of entry does not present a derivation from the normal network. Most of the existing intrusion can be occurred by simultaneously and collaborative on the throughput of a network intruder can be start automated attacks it is targeting all type of vulnerabilities within the network simultaneously [6]. The network host fellows' intrusion detection system with collaborative IDS is managed either hierarchical manner otherwise decentralised manner of the large network. This type of clusters host machine can communicate directly with each other then also use centre coordinator with flexible mode of organisation can be applied [7][8].The hierarchical collaborative intrusion detection system is enable performance analysis on the aggregative information [9].This type of system avoid the unwanted network flow[10].

4. DEFICIENT OF CURRENT INTRUSION DETECTION SYSTEM

The collaborative intrusion detection system is seemed not trustable in detect the intrusion. However the existing technology is not flexible for large type of network and dataset with different parameter. This network traffic information is limited in additional computational of network is their data summarization is proportional to the amount of network traffic flow experience [11]-[13].The main disadvantage is of this approach can be found efficiency and accuracy in data summarization. In machine Accuracy without the knowledge of network host that work information from other host which means man in the middle attack. In efficiency terms is required additional computation is more denser traffic in data summarization. In IDS software on a one network device analysis and managed the network information of the device and is not built in security properties are confidentiality, authentication and integrity control[14][15] .The existing collaborative intrusion detection system is cannot control the traffic flow in largest data sets with parameter and also increase the time and space complexity. The performance is slower for multiple transactions occurs at the same time .service delay is high for existing system.

5. ALGORITHM

This system used RSA algorithm for public key infrastructure .It also used for signature, encryption, decryption and key exchange for user specification. Now a day most of the systems are used to key exchange and signature then it also improves the performance of execution of computational network.

Pseudo code

```
E=exponential value
N=modulo value M
M=message to encrypt
C=chipper text value
C=1.....set as default value
While E
    if  E is odd
        C=C*M
        C=C%N
```

E=E/2
*M=M*M*
M=M%/n
While loop

6. MAPREDUCE FRAMEWORK FOR TRAFFIC FLOW ANALYSIS

Hadoop can provide the efficient technology for big data it deals with distribute the file for different types of environment. Apache foundation is developed by hadoop with efficient and scalable performance for user specification. Hadoop employee the computation paradigm is call mapreduce which is classified into map and reduce terms of phase. In terms of map phase job computation is divide into multiple segments and is distributed with different cluster nodes. The individual segment results are estimate with the finally mapping phase it collaborate and reduced to the one output in the reduce session phase is generate the result with in the short time of period.

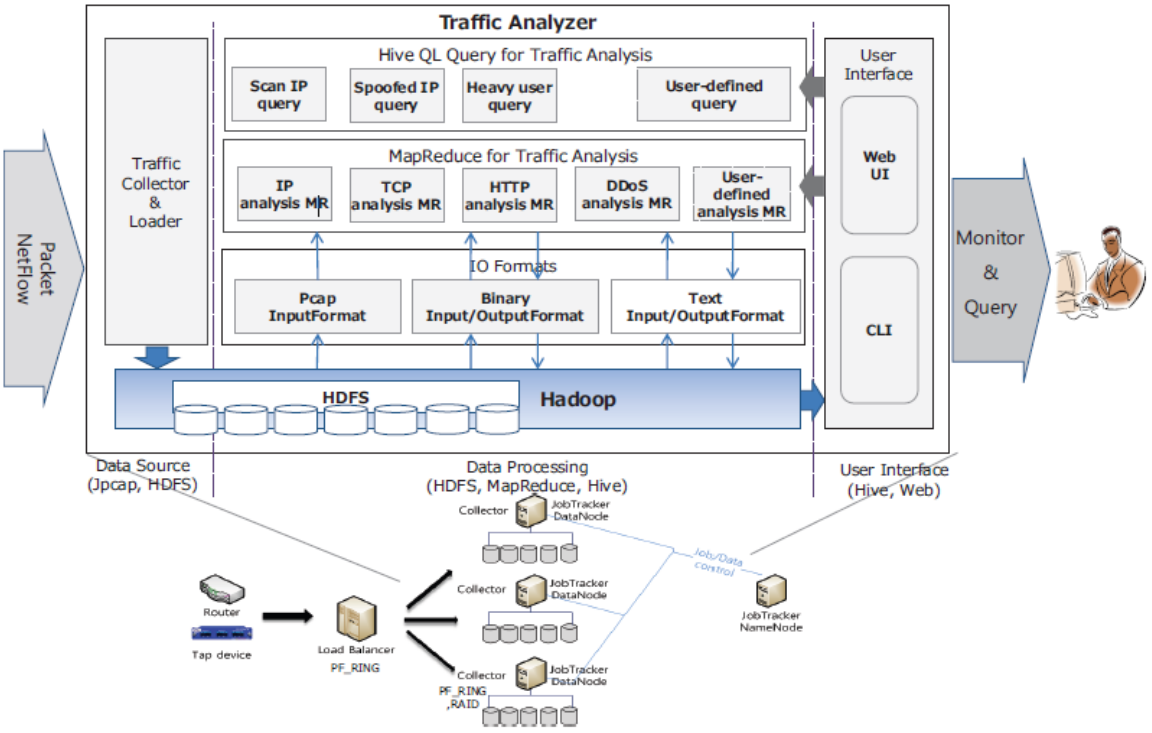

Figure1.Mapreduce Framework for Network Traffic Flow

7. PROPOSED SYSTEM

This system proposed cluster based routing protocol with help of Delay Tolerant hadoop networks (DTHNs) using Hadoop distributed file system. Mobility pattern can interchange the large datasets with load balancing and reduce overhead of network traffic flow. This HDFS system ensure to achieve the scalable and efficient transactions between the different clusters .To detect and prevent data from attacker and it also monitoring the network traffic flow with hadoop and mapreduce framework. Hadoop distributed file system is analysis the network traffic flow with low computation complexity of job. An exponentially weighted average moving time (EWAM) schema is worked for online update with nodal contact probabilities, which means prove the converge to true contact probability and it carried out to evaluate the efficiency and effectiveness of cluster based routing protocol.

8. SYSTEM ARCHITECTURE

Initially user requested admin controller the admin controller check what the database are available in ready queue with user specified data is available means it provide the authentication key. if it is authorized user it will dispatch the data from the database. Now the data generator can generate the key for requested user. It checks an performance for cluster

routing host and it measure the target service delay if delay is occurred in transaction then performance monitor rework the service delay with help of feedback controller. feedback controller check which problem should arrive the network after that query sent to feedback regulator if it is no problem arrive in network it achieve the exponentially weighted moving average time with better performance.

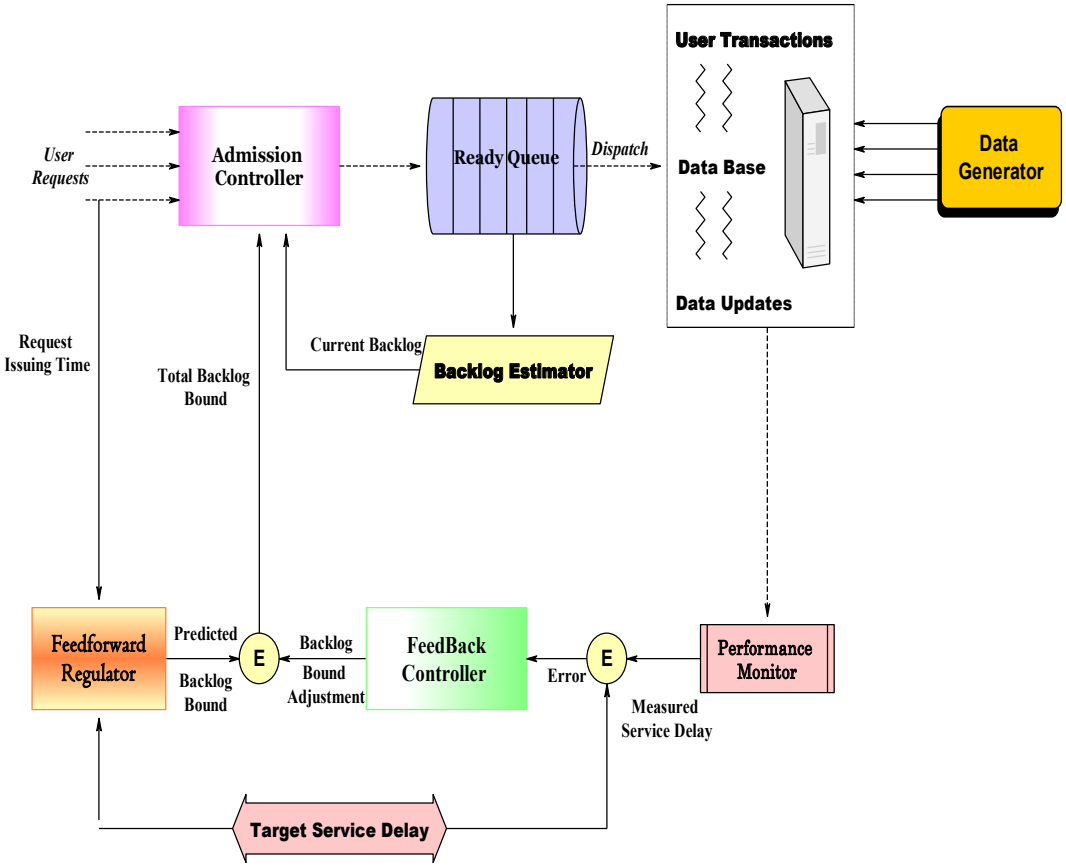

Figure2. System Architecture

9. EXPERIMENTAL RESULTS

For the experiment, we are used to hadoop distributed file system .In this system are mainly estimate the performance flow of network and it also achieve the scalable and efficient routing in DTHN.hadoop can provide the efficient technology for big data it deal with distributed environment. Hadoop system is developed apache foundation with efficient and scalable performance for user specification.

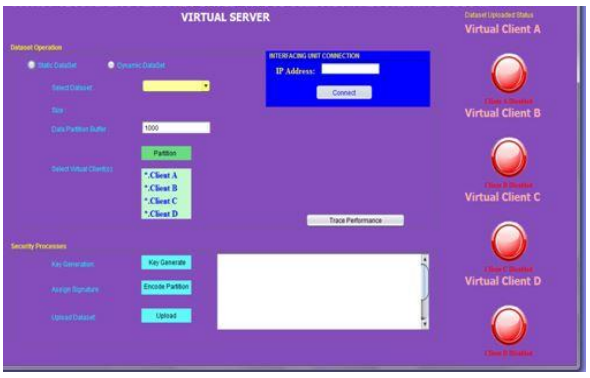

Figure3. Authentication with Interface Connection

CONCLUSION

In this paper, presented monitoring network traffic based on mapreduce framework with hadoop distributed files system. This propose scheme produce the cluster based routing protocol for delay tolerant hadoop network. This system is provide privacy and security data transmission among the large dataset without data losing .Hadoop distributed file system enable the interfacing unit for secured based cluster routing and also interface unit are used to overcome the attacks are man in the middle attack, Denial of service attack, and Sybil attack. Finally these systems achieve the efficient network traffic flow control and share the large dataset without loss.

REFERENCES

[1] Hoplaros.D, Tari, Z., and Khalil, I.:" *Data summarization for network traffic monitoring"*, Journal of Network and Computer Applications, *37,* pp. 194-205(2014).

[2] Zhiyuan, T., Jamdagni, A., Xiangjian, H., Nanda, P., and Ren Ping, L.: "*A System for Denial-of-Service Attack Detection Based on Multivariate Correlation Analysis"*, Parallel and Distributed Systems, IEEE Transactions on,, *25*, (*2),* pp. 447-456,(2014).

[3] Patel A., Taghavi, M., Bakhtiyari, K., and Celestino JúNior, J.: "*An intrusion detection and prevention system in cloud compu-ting: A systematic review*", Journal of Network and Computer Applications, , *36, (1),* pp. 25-4(2013).

[4] Dhage, S.N., and Meshram, B.: "*Intrusion detection system in cloud computing environment*", International Journal of Cloud Computing, *1, (2),* pp. 261-282(2012).

[5] Dean, J., and Ghemawat, S.: "*MapReduce: simplified data processing on large clusters*", Communications of the ACM, *51, (1),* pp. 107-113(2008).

[6] Ram, S.: "*Secure cloud computing based on mutual intrusion detection system"*, International journal of computer application, *, 2, (1),* pp. 57–67(2012).

[7] K. Fall. :"*Delay-tolerant network architecture for challenged Internets*, "in Proc. ACM SIGCOMM, pp. *27.*–34, (2003).

[8] Y. Wang and H. Wu ."*Delay fault tolerant mobile sensor network (DFTMSN)*: *a new paradigm for pervasive information gathering,*." IEEE Trans. Mobile Computing, vol. *6*, no. 9, pp. 1021.–1034, (2007).

[9] Y. Wang, H. Wu, F. Lin, and N.-F. Tzeng,"*Cross-layer protocol design and optimization for delay/fault tolerant mobile sensor networks*,." IEEE J. Sel. Areas Commun, vol. *26*, no. 5, pp. 809.–819, (A preliminary version was presented at IEEE ICDCS.'07,(2008).

[10] C. Liu and J. Wu, ."*Scalable routing in delay tolerant networks*,." in Proc. ACM MobiHOC,(2007).

[11] K. Shvachko, HairongKuang, Sanjay Radia, Robert Chansler, *"The Hadoop Distributed File System",In Proceedings of the IEEE 26thSymposium on Mass Storage Systems and Technologies,* (2010).

[12] K.Vijaya Kumar and G.Nanda Kumar, P.Sudha, A.Kumaresan," *Geographical approximate string search for retrieving errorious data in spatial database"*,(2014).

[13] Yuanjun Cai ,Min Luo, "*Flow Identification and Characteristics Mining from Internet Traffic using Hadoop*" in 978-1-4799-4383-8/14/ at IEEE (2014).

[14] J. Leguay, T. Friedman, and V. Conan, ."*DTN routing in a mobility pattern space*,." in Proc. WDTN.'05: 2005 ACM SIGCOMM Workshops on Delay-tolerant Networking, pp. 276.–283, (2005).

[15]T. Benson, A. Akella, and D. A. Maltz, "*Network traffic characteristics of data centres in the wild*," in Proceedings of the 10th ACM SIGCOMM conference on Internet measurement. ACM,, pp. 267–280,(2010).