Quintessential Cluster Tags For Extend Network

G.Anitha[1], Dr.A.Muthu Kumaravel[*2]

Assistant Professor, Dept. of Master of Computer Application, Jerusalem College of Engineering, Chennai, Tamil Nadu, India [1]

Professor& HOD, Dept. of Master of Computer Application, Bharath University, Chennai, Tamil Nadu, India [2]

[*] Corresponding Author

ABSTRACT: With the increasing popularity of the spread computing on cluster the novel algorithm for efficient cluster method ,super parallel and mining with continues process multiple computing to build booster division. To enhance change data update of paramount and to deploy singly inherently combination data. By exploiting their own architecture the algorithms gain significant speedup. Due to the recent overwhelming growth rate of large-scale data, the development of faster processes algorithms with optimal performance as becomes a need of the time.

I.INTRODUCTION

Several trends are opening up in gather information that contains millions of training samples with thousands of feature. Internet-based development and use of computer technology. The ever cheaper and more powerful processors, together with the software and computing architecture, are transforming data centers into pools of computing service on a huge scale. The increasing network bandwidth and reliable yet flexible network connections make it even possible that users can now subscribe high quality services from data and software that reside solely on remote data centers. To efficiently handle such large-scale data, faster processing and optimization algorithms has become critical in these applications. investigated the problem of data security and mapping the shared data in network nodes, which is essentially a distributed storage system. The algorithms can induce boosted models whose generalization performance is close to the respective baseline classifier. Moreover, the algorithms do not require individual computing nodes to communicate with each other, to share their data or to share the knowledge derived from their data and hence, they are robust in preserving privacy of computation as well. After analyzing the requirements of the task to be performed, the next step is to analyze the problem and understand its context. The first activity in the phase is studying the existing system and other is to understand the requirements and domain of the new system. Both the activities are equally important, but the first activity serves as a basis of giving the functional specifications and then successful design of the proposed system. Understanding the properties and requirements of a new system is more difficult and requires creative thinking and understanding of existing running system is also difficult, improper understanding of present system can lead diversion from solution.

II. PROBLEM DEFINITION AND DESCRIPTION

We present several events in greater detail to specify the competition of measuring and managing wide-area performance [1]. We propose metrics that quantify the latency contributions across sets of servers and routers. Based on the design, we implement the fingers tool for diagnosing large latency increases for CDN. After detecting large increases in latency, our classification must first determine whether client requests shifted to different front-end server for the labeling details in forcing the wide area data exchange where taken in across in the medium [2].

III. REVIEW OF LITERATURE

A)*Cluster Intimation*

At the beginning of the process, each element is in a cluster of its own [3]. The clusters are then sequentially combined into larger clusters until all elements end up being in the same cluster. At each step, the two clusters separated by the shortest distance are combined. The definition of 'shortest distance' is what differentiates between the different agglomerative clustering methods. In complete-linkage clustering, the link between two clusters contains all element pairs, and the distance between clusters equals the distance between those two elements (one in each cluster) that are farthest away from each other [4]. The shortest of these links that remains at any step causes the fusion of the two clusters whose elements are involved. The method is also known as **farthest neighbour clustering**. The result of the clustering can be visualized as dendogram which shows the sequence of cluster fusion and the distance at which each fusion took place [5].

Mathematically, the complete linkage function — distance $D(X,Y)$ between clusters $D(X,Y)$ and $D(X,Y)$ — is described by the following expression : $D(X,Y)$ where

- $D(X,Y)$ is the distance between elements $D(X,Y)$ and $D(X,Y)$;
- $D(X,Y)$ and $D(X,Y)$ are two sets of elements (clusters)

Complete linkage clustering avoids a drawback of the alternative method - the so-called *chaining phenomenon*, where clusters formed via single linkage clustering may be forced together due to single elements being close to each other, even though many of the elements in each cluster may be very distant to each other. Complete linkage tends to find compact clusters of approximately equal diameters [6-8].

B) Network Overlay in cluser

Each node points to another node called its successor. The last node points to the first node and forming the network overlay. For now we will not discuss how to maintain this overlay, rather concentrate on store and lookup assuming we have the overlay [9]. Nevertheless, it is very simple to simulate the overlay, you can use a simple Circular Linked List, where each node points to a successor node forming the circular list.

IV. SYSTEM DESIGN

Software design sits at the technical kernel of the software engineering process and is applied regardless of the development paradigm and area of application [10]. Design is the first step in the development phase for any engineered product or system. The designer's goal is to produce a model or representation of an entity that will later be built.

The importance can be stated with a single word "Quality". Design is the place where quality is fostered in software development [13]. Design provides us with representations of software that can assess for quality. Design is the only way that we can accurately translate a employee's view into a finished software product or system.

Input design is the process of converting user-originated inputs to a computer-based format. Input design is one of the most expensive phases of the operation of computerized system and is often the major problem of a system. In the project, the input design is made in various web forms with various methods [14].

Output design generally refers to the results and information that are generated by the system for many end-users, output is the main reason for developing the system and the basis on which they evaluate the usefulness of the application. In the project, the admin details, user details, are the web forms in which the output is available. The database design is a must for any application developed especially more for the data store projects. Proper handling of the table is a must. In the project, admin table is designed to be unique in accepting the username and password should be greater than zero [15].

V. ARCHITECTURE OF SYSTEM

CDN server storage system. Domain name space system collects all the data and it will store in the load balancer 1 and load balancer 2.It will directly send to application 1 and application 2 to process the request send by the user. The MasterDB collect those requests and store in the CDN storage and also frequently accessed data's, if any repeated of program needed it can be done in SlaveDB which is a proxy. These entire process takes place will be in a backup of CDN Storage[11-12]
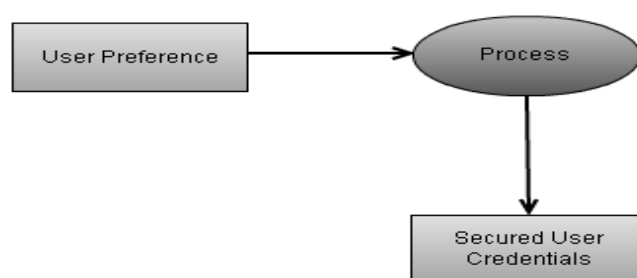

Fig4.1user preferences with credentials

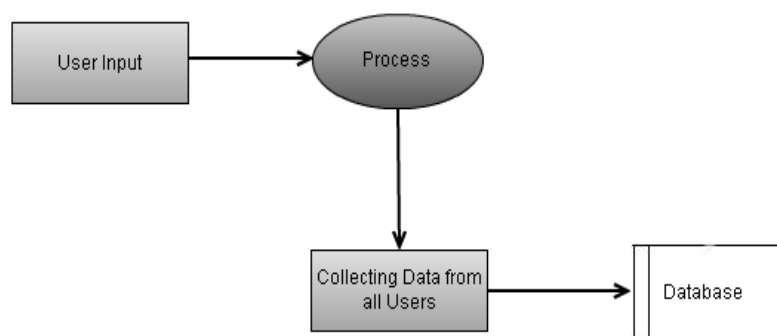The prefrence user communicating with the secured user credentials to retiving te data

Fig 4.2 User with data collected views

V. PROCESS DESCRIPTION

A Networked computer system that harnesses the resources of several servers to complete tasks and store data is shaping up to be the next big trend in the computing industry. Since the debut of the personal computer, we've become used to storing information either on an external storage device like a compact disc or on a computer's hard drive [16-17].

We are also conditioned to buy new machines or upgrade old ones whenever applications require more processing power than our current computers can provide. With CDN, the responsibility of storage and processing power falls to the network, not the individual computer owner as possible.

The most obvious of these advantages is that the applications aren't tied to a specific computer. There's no need to download and install software on a particular machine. Any computer connected to the Internet can access our network [18].

Users don't have to worry about which version of a document is the most current it will always be saved in our CDN. Users Can Upload File on server with high speed. It will be stored many server [19].

A)Cluster Tags Format

Cluster TAG is a distributed programming system which is capable of processing large data sets. Map function and the Reduce function are the two main functions involved in Cluster Labeling Framework[2].

User can also set the inner for the search performed in the nodes. The Map tasks are processed in parallel by the nodes in the cluster without sharing data with any other nodes.

Classifiers from different iterations will ensure algorithm's convergence and the definition of base classifier has been changed to a ternary classifier to form a mapped network. The framework maintains a vector of probability estimates (p) for each data point which is initialized and updated during each iteration [20].

B)File Download on CDN

We propose a decision tree for separating the causes of latency changes from their effects, and identify the data sets needed for each step in the analysis. We analyze the measurement data to identify suitable thresholds to identify large latency changes and to distinguish one possible cause from another.

Our tool Fingers can analyze latency increases and traffic shifts over sets of servers and routers. Once path performance is known, CDNs can optimize inter domain path selection based on performance, load, and cost.

We get the file with high download speed compare to other networks. Another advantage is that multiple users can make download to the same files at the same time. This is called online collaboration, and it could streamline teamwork over the web.

C)Find On CDN

User can search for documents or files using some of the standard CDN query parameters. Categories are used to restrict the type of resource (word processor, spreadsheet, presentation, collection, etc) returned. The full-text query string is used to search the content of all the documents.

User can search for resources matching an exact title or portion of a title by using the title-exact and title query parameters, respectively. User can search the content of documents by using the query parameter on the feed. Efficient learning framework of nested cascades of boosted classifiers model induced much faster and with a good generalization capability with special emphasis on the training procedures.
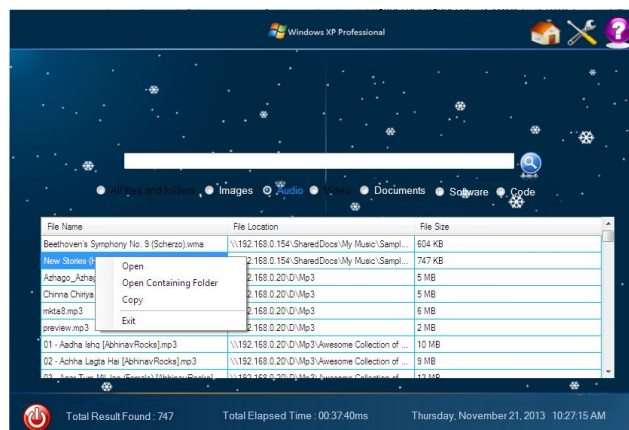

Fig 5.1 Documents and Software

It contains various modules all files and folders, images, audio, video, documents, software and code

D)Modification System Mapping

An algorithmic stream coordinate with its efficiency in the generalization performance makes it one of the most powerful ensemble methods. It has the interface between the networks. The node structures are designed to work in mesh environments where each node is a store house of vast amount of informatics and dates.

The implemented method gives more focused dependent iterative setting that makes boosting inherently serial independent iterations. Once authenticated, firewall enforces access policies such as what services are allowed to be accessed by the network users

E)Master Role

Admin can manage all the files on the server. He can delete all files. It can view all the Information about Download and Upload which was made by User. Admin can search for resources matching an exact title or portion of a title by using the title-exact and title query parameters, respectively. It can search for documents or files using some of the standard CDN query parameters.

F) Vector Factor
The standard SUPPORT VECTOR algorithm ensemble learning method each iteration a strong classifier from a pool of weak hypotheses. In the Final iteration on the classifiers, a weighted linear combination of base classifiers are analyzed where each of them are entered in the array list.

The classifiers with lower error rate are entered first in the array list which indicates higher weight. MINING METHODS is a powerful boosting method that uses regression functions instead of classifiers and these functions output real values in the same form as prediction.

The base classifiers should have a higher rate than a random classifier. Simple decision trees with only one non-leaf node often perform well for SUPPORT VECTOR.

TCP (Transmission Control Protocol) and UDP (User Datagram Protocol) are two of the protocols that make up the TCP/IP protocol suite which is used universally to communicate on the Internet. Each of these has ports 0 through 65535.

The systems are inter-connected with the neighboring system in LAN network. The IP address and the DNS servers are configured in order to collect the details need to map the files in network sharing.

Once the server Connection is started the Server Service will be accessible to the clients connected in the LAN Network.

A performance boosting algorithms used in collecting the information about the system connected in the network. Once the systems are connected in the network link the share file information are collected from each node.

An algorithmic stream coordinate with its efficiency in the generalization performance makes it one of the most powerful ensemble methods. It has the interface between the networks.

The node structures are designed to work in mesh environments where each node is a store house of vast amount of informatics and dates.

The implemented method gives more focused dependent iterative setting that makes boosting inherently serial independent iterations. Once authenticated, firewall enforces access policies such as what services are allowed to be accessed by the network users.

The standard SUPPORT VECTOR algorithm ensemble learning method each iterations a strong classifier from a pool of weak hypotheses. In the Final iteration on the classifiers, a weighted linear combination of base classifiers are analyzed where each of them are entered in the array list.

G)MINING TERMS

The classifiers with lower error rate are entered first in the array list which indicates higher weight. MINING METHODS is a powerful boosting method that uses regression functions instead of classifiers and these functions output real values in the same form as prediction.

The base classifiers should have a higher rate than a random classifier. Simple decision trees with only one non-leaf node often perform well for SUPPORT VECTOR.Cluster Labeling is a distributed programming system which is capable of processing large data sets. Map function and the Reduce function are the two main functions involved in Cluster Labeling Framework.

User can also set the depth for the search performed in the nodes. The Map tasks are processed in parallel by the nodes in the cluster without sharing data with any other nodes. Classifiers from different iterations will ensure algorithm's convergence and the definition of base classifier has been changed to a ternary classifier to form a mapped network.

The framework maintains a vector of probability estimates (p) for each data point which is initialized and updated during each iterations. Settings and Secured Distributed Computations

The cost of finding the non-leaf node is evaluated and the data samples are sorted in each attribute. Sorting all the attributes is done once before starting the first iteration
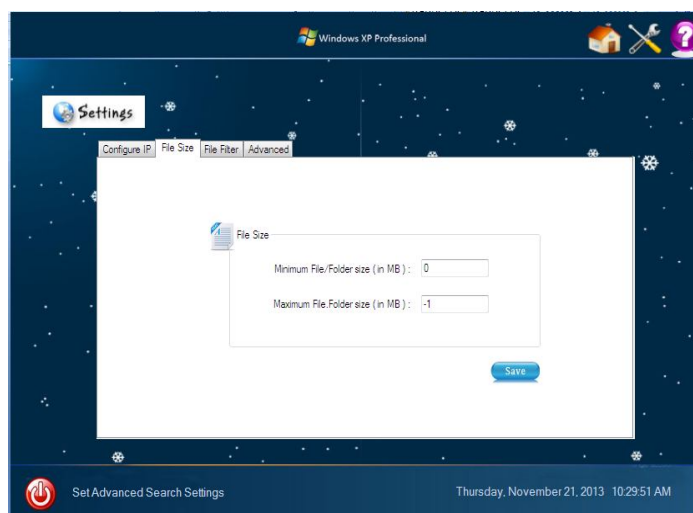

Fig 5.2 File size view report

After receiving the classifiers or functions from all the Map functions, the Reduce function merges them at the same sorted level and averages the classifier weights to derive the weights of the merged classifiers. The iterations are executed in each paralleled node and the complete algorithm builds a weak classifier on its own data structure.
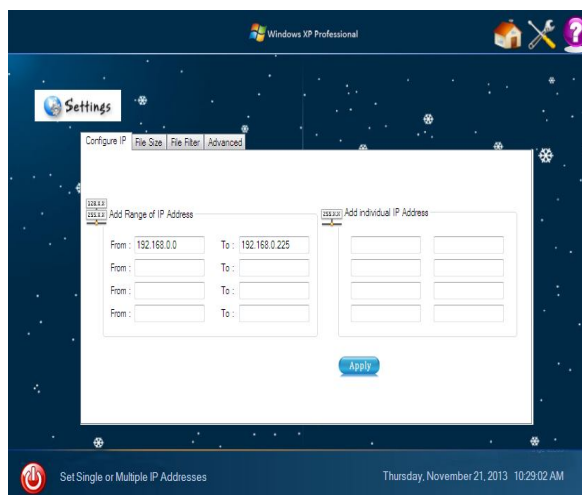
Fig 5.3 IP Range Format

For the computation to be considered completely secure, each nodes after the completion of the task except for what can be inferred from their own input. After receiving the merged classifier, the user program sends it to all of the Map nodes. Each of the Map nodes calculates the error rate of the received merged classifier on its own data and transmits it to the Reducer

VI. CONCLUSIONS

Thus the problem of data security in distributed network sharing, which is important a distributed storage system. The searching process is boosted with efficient algorithm. This proves the efficiency and protects the ability of the clients. The direction in which we have moved our project is to avoid or prevent the data losses and also maintaining the updates in the server for all events which have took place in the client side. The proposed algorithms is competitive to the original versions and is even better in some cases privacy preserving based computation, but its programmatic construct is suitable for a parallelized implementation. It employs two or more computing workers to construct the boosting classifiers each of the workers has access to only a specific subset of training data.

VII. FUTURE ENHANCEMENT

The direction enhanced is implementing this type of application in a widely connected environment. The scale up performance of our algorithms shows that they can efficiently utilize additional resources when the problem size is scaled up. In the future, we plan to explore other data partitioning strategies (beyond random stratification) that can improve the classification performance even further.

REFERENCES

1. Manning, Christopher D., Prabhakar Raghavan, and Hinrich Schutze. *Introduction to Information Retrieval*. Cambridge: Cambridge UP, 2008. *Cluster Labeling*. Stanford Natural Language Processing Group. Web. 25 Nov. 2009.
2. Consummate Cluster Labeling for Spread Network IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 6, August 2014
3. Jayalakshmi T., Krishnamoorthy P., Ramesh Kumar G., Sivamani P., "Optimization of culture conditions for keratinase production in Streptomyces sp. JRS19 for chick feather wastes degradation", Journal of Chemical and Pharmaceutical Research, ISSN : 0975 – 7384, 3(4) (2011) PP.498-503.
4. Indranil Palit, Chandan K. Reddy, "Scalable and ParallelBoosting with MapReduce", IEEE Transactions onKnowledge and Data Engineering, vol. 24, no. 10, pp. 1904-1916, Oct. 2012, doi:10.1109/TKDE.2011.208.
5. Geetha Ramani R, Jacob SG, "Prediction of P53 Mutants(Multiple Sites) Transcriptional Activity Based onStructural (2D&3D) Properties", PLoS ONE 8(2): e5
6. Andrew Eisenberg and Jim Melton. *Advancements in SQL/XML*. ACM SIGMOD Record ,Vol. 33, No. 3, September 2004.

7. Jebaraj S., Iniyan S., "Renewable energy programmes in India", International Journal of Global Energy Issues, ISSN : 0954-7118, 26(4Mar) (2006) PP.232-257.

8. Andrew Eis enberg and Jim Melton. *An Early Look at XQuery API for Java™ (XQJ)*. ACM SIGMOD Record ,Vol. 33, No. 2

9. Ramez Elmasri and Shamkant B. Navathe. *Fundamentals of Database Systems*, second edition. Addison-Wesley Publishing Company.

10. Gopalakrishnan K., Prem Jeya Kumar M., Sundeep Aanand J., Udayakumar R., "Thermal properties of doped azopolyester and its application", Indian Journal of Science and Technology, ISSN : 0974-6846, 6(S6) (2013) PP. 4722-4725.

11. Donald Kossmann and Konrad Stocker. *Iterative Dynamic Programming: A new Class of Query Optimization Algorithms*. ACM Transactions on Database Systems, Vol. 25, No. 1, March 2000, Pages 43- 82.

12. Chiang Lee, Chi - Sheng Shih and Yaw - Huei Chen. *A Graph-theoritic model for optimizing queries involving methods*. The VLDB Journal — The International Journal on Very Large Data Bases, Vol. 9,Issue 4, Pages 327 -343.

13. Beula Devamalar P.M., Thulasi Bai V., Srivatsa S.K., "Design and architecture of real time web-centric tele health diabetes diagnosis expert system", International Journal of Medical Engineering and Informatics, ISSN : 1755-0661, 1(3) (2009) PP.307-317

14. K. Bielenberg and M. Zachera. Groups in socialsoftware: Utilizing tagging to integrate individualcontexts for social navigation. 2005.

15. U. Brandes, M. Gaertler, and D. Wagner. Experimentson graph clustering. In Proceedings of the 11th AnnualEuropean Symposium on Algorithms (ESA'03), volume2832 of Lecture Notes in Computer Science, pages568–579. Springer-Verlag, 2003.

16. Sharmila S., Jeyanthi Rebecca L., Saduzzaman M., "Biodegradation of domestic effluent using different solvent extracts of Murraya koenigii", Journal of Chemical and Pharmaceutical Research, ISSN : 0975 – 7384, 5(2) (2013) PP.279-282.

17. M. E. Newman and M. Girvan. Finding and evaluatingcommunity structure in networks. Physical Review E,69(2), 2004.

18. A. Pothen, H. D. Simon, and K.-P. Liou. Partitioningsparse matrices with eigenvectors of graphs. SIAM J.Matrix Anal. Appl., 11(3):430–452, 1990.

19. S. White and P. Smyth. A spectral clustering approachto finding communities in graphs. In SIAM International Conference on Data Mining, 2005.

20. Consummate Cluster Labeling for Spread Network IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 6, August 2014

21. B Karthik, TVU Kirankumar, MS Raj, E BharathKumaran,Simulation and Implementation of Speech Compression Algorithm in VLSI, Middle-East Journal of Scientific Research 20 (9), PP 1091-1092, 2013

22. M.Sundararajan & R.Pugazhanthi," Human finger print recognition based biometric security using wavelet analysis", Publication of International Journal of Artificial Intelligent and Computational Research, Vol.2. No.2. pp.97-100(July-Dec 2010).

23. .M.Sundararajan & E.Kanniga," Modeling and Characterization of DCO using Pass Transistor", proceeding of Springer – Lecturer Notes in Electrical Engineering-2011 Vol. 86, pp. 451-457(2011). ISSN 1876-1100.(Ref. Jor- Anne-II)

24. .M.Sundararajan & C.Lakshmi, "Wavelet based finger print identification for effective biometric security", Publication of Elixir Advanced Engineering Informatics-35(2011)-pp.2830-2832.

25. .M.Sundararajan, "Optical Instrument for correlative analysis of human ECG and Breathing Signal" Publications of International Journal of Biomedical Engineering and Technology- Vol. 6, No.4, pp. 350-362 (2011). ISSN 1752-6418.(Ref. Jor-Anne-I

26. M.Sundararajan, C.Lakshmi & D.Malathi, "Performance Analysis Restoration filter for satellite Images" Publications of Research Journal of Computer Systems