IFERP

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 3, Issue 7, July 2016

Integrated Fragmentation Clustering Allocation Approach for Promote Web Telemedicine Database System

[1] Miss. Tejashri S. Nimbalkar [2] Mr.Nagaraju Bogiri
Department Of Computer Engineering
K.J.College of Engineering & Management Research,
Pune, India
[1] Tejashri25.nimbalkar@gmail.com [2] nagaraju@gmail.com

Abstract- **— Although the application of telemedicine evolutionary to cover a wide area of users' needs, data centralization is not achieved yet. Also, approach to patient data from any remote location, through a secure environment, is necessary to achieve doctor's collaboration and remote data handling. So web telemedicine database systems used for access to patient data from any location. The real world healthcare challenging application makes it hard to induce the database administrative staff. Traditional approaches for promote web telemedicine database systems focus on small networks involves minimum number of sites over the cloud. These are works on limited clustering algorithms. To overcome this Limitation, we need a new approach for promoting web telemedicine database system. In this project we proposed a novel Integrated Fragmentation clustering allocation approach for increase care admissions and decrease care difficulties on Approach that manages the computing web services that are required to promote telemedicine database system performance. Our approach focused on large scale networks involving large number of sites over the cloud. To perform more intelligent data redistribution, we apply different types of clustering algorithms and introduce search based techniques. The security concerns, need for addressing over data fragments will be taken into consideration for better results.**

Keywords---: **WTDS, IFCA, Fragmentation, Clustering**

I. INTRODUCTION

Telemedicine is an easy to use technology and it is new, flirtatious, and outwardly, there tends to be a belief among health service managers that in clinicians it can simply be made available who will automatically accept and use the telemedicine systems. With the advancement of communication technology and information, Internet connects millions of hosts ecumenical has been getting more and more popular recently. Computer networks have made it possible to share through remote consultation of electronic medical records and to deliver medical expertise.

Prevailing information Systems do not satisfactorily support these characteristics of telemedicine due to they are still record-oriented in lieu of case oriented, and users cannot directly optically discern an association and an explicative picture of a case.

The software application researchers to propose several computing services' techniques to achieve more efficient and effective management of web telemedicine database systems (WTDS). Significant research progress has been made in the past few years to improve WTDS performance. Different kinds of patient information such as ECG, temperature, and heart rate need to be accessed by means of various client devices in heterogeneous communications environments.

WTDS enable high quality continuous delivery of patient's information wherever and whenever needed. Recently, many researchers have fixated on designing web medical database management systems that satisfy certain performance levels. Such performance is calculated by measuring the amount of relevant and irrelevant data accessed and the amount of transferring medical data during the transactions processing time. Improve database performance the several techniques have been proposed in order, telemedicine, control medical data proliferation and optimize medical data distribution. These techniques believed that high performance for such systems can be achieved by amending at least one of the database web management services, namely—data distribution, database fragmentation, distributed caching, database scalability and websites clustering.

Data records may be overlapped or even redundant with it, which increase the processing time, I/O transactions and so the system communications overhead. These works often investigated fragmentation, sometimes clustering problems and allocation. The transactions should be executed very expeditious in a flexible load balancing database environment. When the number of sites in a web database system increases to an immensely colossal scale, the intractable time intricacy of processing a sizable voluminous number of medical transactions and managing an immensely colossal number of communications make the design of such methods a non-picayune task.

a. Existing System

- Many researchers have been busy on creating web medical DBMS that satisfy certain performance levels.
- Several have been input in front in order to improve telemedicine database performance, optimize medical data distribution, and control medical data Expansion.
- These methods are believed that high performance for such systems can be gotten through using at least one of the database web management services, namely—database fragmentation, websites clustering ,database scalability.
- In an Existing threefold approach that manages the computing web services that are required to promote telemedicine database system performance. But its focus on small networks involves a minimum number of sites over the cloud. These are works on limited clustering algorithms.

Existing Advantages

- The conclusion requires more investigation and experiments.

Existing Disadvantages

- It cannot focus on a large number of networks.
- It cannot work on different types of clustering algorithms.

b. Proposed System

- In our project, we propose a Novel Integrated Fragmentation Clustering Allocation Approach that manages the computing web services that are required to promote telemedicine database system performance.
- In this work, we address the previous drawbacks and propose a threefold approach that manages the computing web services that are required to promote telemedicine database system performance.
- We are combining 3 Techniques of computing services named data fragmentation, site clustering and fragment allocation.
- We perform both external and internal evaluation of our integrated approach. We develop a fragmentation computing service technique by dividing our database table into small disjoint fragments.
- This method generates a less number of disjoint fragments that would be allocated to the web servers in the data distribution phase. Thus reducing the total amount of data transfer and accessed by different sites reducing cost of communication.
- We introduce a high speed clustering method that groups the web database sites into sets of clusters according to their communications cost.
- This helps in grouping the sites that are similar into a single cluster thus it minimize data allocation operations, which in turn helps to avoid allocating duplicate data.

Proposed System Advantages

- Focus on a large number of networks involves large number of sites.
- It supports different types of clustering algorithms.
- Generates a less number of disjoint fragments that would be allocated to the web servers in the data distribution phase.

II. LITERTURE REVIEW

In this we have various researches on data fragmentation.

Vertical Fragmentation and Allocation in Distributed Deductive Database Systems

There are various researches on data allocation and vertical fragmentation, but lack in distributed deductive database systems (DDDBSs). This paper has different approaches which are used for relations which are referred by the rules and allocation of rules and fragments in DDDBS. Minimize the communication cost and Maximum locality of query evaluation in a distributed system is the main advantage of the proposed system [1].

Integrating vertical and horizontal partitioning into

In relational database horizontal partitioning and vertical partitioning plays a very important role. Partitioning is important for such an operation like backup, restore easily. This paper presents a novel technique for designing a scalable solution to the integrated design of indexing and partitioning, which takes both manageability and performance into account. In this implemented techniques and calculated results on Microsoft SQL Server [2].

A Mixed Fragmentation Methodology for Initial Distributed Database Design

In this paper, mixed fragmentation is introduced which applied for the horizontal and vertical fragmentation on a relation.

It can be obtained in one of two ways first performing horizontal fragmentation come after vertical fragmentation or performing vertical fragmentation come after horizontal fragmentation. Mixed fragmentation is needed in distributed databases. In this paper proposes a technique for generating candidate vertical and horizontal fragmentation and proposed a scheme by using this fragmentation for distributed database design [3].

A Framework for Server Data Fragment Grouping to Improve Server Scalability in Intermittently Synchronized Databases

Mobile computing is expensive, so clients connect periodically. Conventionally, constantly connected server has various facilities for sharing data, upload and download updates. Servers itself compute and retransmit these updates on a client basis. Complexity is depending upon the number of clients. Now days, data centric data sharing is used behalf of the client centric system [4].

Data clustering: a review

Clustering is a classification of patterns into a cluster. In many research areas clustering problem is addressed. Clustering is very important in data analysis. In this paper author presents pattern clustering methods from statistical pattern recognition perspective, which is useful for clustering practitioners. In this paper has different clustering techniques and identify crosscutting themes and recent techniques [5].

III. PROPOSED APPROACH FRAMEWORK AND DESIGN

a. Problem definition:

Several methods have been proposed in order to ameliorate telemedicine database performance, optimize medical data distribution, and control medical data expansion. These techniques believed that high performance for such systems can be gotten through using at least one of the database web management services, namely—database fragmentation, websites clustering ,database scalability. However, the intractable time intricacy of processing an immensely colossal number of medical transactions and managing a sizable voluminous number of communications makes the design of such methods a non-picayune task. Moreover, none of the existing methods consider the threefold service together, which makes them impracticable in the field of web database systems. Additionally, using multiple medical services from different web database providers may not fit the needs for improving the telemedicine database system performance. Further, the accommodations from different web data-base providers may not be compatible or in some cases it may increase the processing time because of the constraints of on the network. Finally, there has been a lack in the tools that support the design, analysis and cost-effective deployments of web telemedicine database systems.

Designing and developing fast, efficient, and reliable incorporated techniques that can handle a huge number of medical transactions on a large number of web health care sites in near optimal polynomial time are key challenges in the area of WTDS. To improve the performance of medical distributed database systems, we incorporate data fragmentation, websites clustering, and Fragmentation allocation computing services together in a new web telemedicine database system approach. This new approach intends to decrease data communication, increase system throughput, reliability, and data availability.

b. Proposed Methodology:

We propose an approach to integrate database fragmentation, websites clustering, and data fragment allocation into one scenario to resulting in ultimate web telemedicine system throughput in terms of concurrency, reliability, and data availability. We call this scenario Integrated Fragmentation Clustering Allocation (IFCA) approach.

The data request is initiated from the telemedicine database system sites. In existing system takes a small number of sites, but our proposed work we take a large number of telemedicine database system sites. The requested data are defined as SQL queries that are executed on the database relations to generate data set records. Some of data records may be in over-lapped fashion, even duplicate which cause increase in I/O transactions processing time and so the system communications overhead.

To solve this problem, we execute the proposed fragmentation technique which generates telemedicine disjoint fragments that represent the minimum number of data records. The web telemedicine database sites are grouped into clusters by using different types of clustering algorithms in a phase prior to data allocation. On Existing work only focus on the limited clustering algorithm. But we proposed different types of clustering algorithm such as kmeans++, k prototype and so on. The purpose of this clustering is to reduce the communications cost needed for data allocation. Accordingly, the proposed allocation service technique is applied to allocate the generated disjoint fragments at the clusters that show positive benefit allocation. Then the fragments are allocated to the sites within the selected clusters. Database administrator is responsible for recovering any site failure in the WTDS.

Incorporating database fragmentation, web database sites‘ clustering, and data fragments computing services‘ allocation techniques in one scenario distinguishes our approach from other approaches. The telemedicine IFCA approach is designed to support web database provider with computing services that can be implemented over multiple servers, where the data storage, communication and processing transactions are fully controlled, costs of communication are symmetric, and the patients‘ information privacy and security are met.

c. Algorithm Used In Existing System

1. DATA FRAGMENTATION:

1) Set fragmentation Size K=F.size ();

2) for int i=1;i<f.size();

3) for j=I;j<f.size();

4) intial fragments Fi,Fj
4) create new fragment Fk
Fk=Fi n Fj where, n=intersection Fk and add to list of fragment F
5)Loop until I and j reach F.size(); 6)return added fragment list F;

2. FRAGMENTS ALLOCATION ALGORITHM

Input: S=List of sites
F= list of disjoint fragments.
1)determine if F is allocated or not
2)if allocated then
Det(F,S)=1
Else
Det(F,S)=0
3)in this step we tend to compute cost of allocating and cost of not- allocating fragments components
4)on basis of values cost of allocating and cost of not-allocating fragments components calculate allocation decision value

5)return List of fragments that are allocated to each web cluster.
6)end;

d. Algorithm used in proposed system:

1. CLUSTERING ALGORITHM:
K-Means Clustering:
Input: Matrix Communication cost between sites. Output: Clustering Decision Values matrix.
Input: Matrix Communication cost between sites X.

Input: Matrix Communication cost between sites X.
Output: and V = {v1,v2,…….,vc} be the set of centriods of clusters.
Let X = {x1,x2,x3,……..,xn} be the set of Matrix Communication cost between sites and V = {v1,v2,…….,vc} be the set of centers.
1) Randomly select ‗c' cluster centers.
2) for int i=0;i<k;i++ do where K is number of Cluster

3) Calculate the distance between each sites communication costs and cluster centers.
4) Assign the site to the cluster center whose distance from the cluster center is min of all the cluster centers.
5) Recalculate the new cluster center using:

$$V_i = (1/c_i) \sum_{i=1}^{c_i} x_i$$ here Z represent summation

where, 'c_i' represents the number of sites in i^{th} cluster.

6) Recalculate the distance d between each site and new obtained cluster centers in V.

7) Run loop until updated distances converges.
8) end for loop;
9) return V;
where V is representing set of final centriods
2. K-PROTOTYPE CLUSTERING
K-prototype Clustering applies same principle as k-means as it inherits some property from it make use of Euclidian distance metric PE as distance metric for distance learning to define closeness between to objects.

Step 1: Initially start with selection of random k initial prototypes Q1, ... , Qk , one for each cluster.
Step 2: For each record Xi calculates the distances d from the DB record to the clusters prototype;
Step 3: Now finds the closest prototype Qm to the D Brecord according to the metric pE i.e Euclidean metric (distance) and allocates the record Xi to the cluster Cm with this prototype.
Step 4: Find a new prototype Q‘m for each cluster Cm such that the sum of square distances
Xi∈Cm p2 E(Xi, Q‘m) is minimum.
Step 5: If prototypes Q1, ... , Qk != Q‘1, ... , Q‘k then take new prototypes and go to step 2,
Else
Stop the procedure.

3. DATA FRAGMENTS ALLOCATION
-Once final Cluster centriods are decided we go further to data fragment allocation sub-system.
Input: T=List of transactions in database D
F= list of disjoint fragments C= list of clusters on basis of V.
1) in first step we tend to compute cost of allocating CAF and cost of not- allocating fragments components CNAF
2) on basis of values CAF and CNAF calculate allocation decision value CAD
3) return List of fragments that are allocated to each web cluster.
4) end;

e. System architecture:

These functions guarantee the minimum communications cost among websites and hence accomplish better data distribution and clustering the data compared to allocating data to all websites evenly. Combine DB fragmentation, Website clustering and fragment allocation into one system so that to output efficient web telemedicine DBMS with high throughput.
In the system architecture fig. 1 the data request is initiated from the telemedicine database system sites that is admin and it requesting for the data pre-processing and the requested data in the form of SQL query that are executed on the database relations to generate data set records after that we execute the proposed fragmentation technique which generates telemedicine disjoint fragments that represent the minimum number of data records. The web telemedicine database sites are grouped into clusters by using our clustering service technique in a phase prior to data allocation. The purpose of this clustering is to reduce the communications cost needed for data allocation the proposed allocation service technique is applied to allocate the generated disjoint fragments at the clusters that show positive benefit

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 3, Issue 7, July 2016

allocation. Then the fragments are allocated to the sites within the selected clusters. Database administrator is responsible for that.

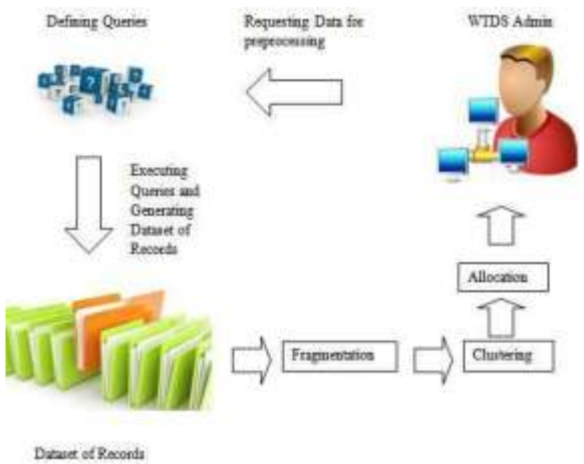

Figure 1: System architecture

IV. MATHEMATICAL MODEL OF PROPOSED SYSTEM

Let S, be a system such that,
S = {s,X,V,c,ci}
Where,
S- Proposed System
s- Initial state at Time T=0.
X - Matrix Communication cost between sites V - set of cluster centers
c - number of cluster centers
ci - number of sites in ith cluster
X- Input of System
-Training data set X={x1,x2….,xn} where X - Matrix Communication cost between sites.
Y- Output of System
- V - set of cluster centers V = {v1,v2,…….,vc}
FME- Main algorithm resulting into outcome Y, mainly focus on success defends for the solution. UP Growth Algorithm is proposed.

DD- Deterministic Data, it helps identifying the load-store function or assignment function. e.g. i= return i. Such function contributes in space complexity.
NDD- Non Deterministic Data of the system to be solved. These being computing function or CPU Time.
F friend- Set of random variables. 0,1
MEM shared- Memory required to process all these operations, memory will allocated to every running process.
CPU Core CNT- More the number of count doubles the speed and performance.
Φ – Null value if any.

V. RESULTS AND EXPERIMENTAL STUDIES

In this section we present the Module description, how it works, practical results and environment.

a. Modules

1. Database Fragmentation

2. Websites Clustering

3. Data Fragments Allocation

1. Database Fragmentation:

- In this module, the data request is initiated from the telemedicine database system sites. The requested data are defined as SQL queries that are executed on the database relations to generate data set records.

- Some of data records may be in over-lapped fashion, even duplicate which cause increase in I/O transactions processing time and so the system communications overhead.

- To solve this problem, we execute the proposed fragmentation technique which generates telemedicine disjoint fragments that represent the minimum number of data records.

2. Websites Clustering:

- In this module, the web telemedicine database sites are grouped into clusters by using different types of clustering technique (k-means, k-prototype and so on) in a phase prior to data allocation.

- The purpose of this clustering is to reduce the communications cost needed for data allocation.

3. Data Fragments Allocation:

- In this module, the proposed allocation service technique is applied to allocate the generated disjoint fragments at the clusters that show positive benefit allocation.

- Then the fragments are allocated to the sites within the selected clusters. Database administrator is responsible for recovering any site failure in the WTDS.

b. Hardware And Software Used

- Hardware Configuration
  - Processor Pentium –IV
  - Speed - 1.1 GHz
  - RAM - 256 MB(min)
  - Hard Disk - 20 GB
  - Key Board - Standard Windows Keyboard
  - Monitor - SVGA

- Software Configuration
  - Operating System -Windows XP/7/8
  - Programming Language - Java
  - Tool - Netbeans.
  - -Server -Wamp Server

c. System Boundaries

a. This new approach intends to decrease data communication

b. Input dataset is restricted only for telemedicine sites

d. Experimental Setup

Input:

c. Queries on patients training data matrix

Output:

d. Get the requested data in the preprocess form

e. Results of Practical Work

In this section we provide the results on basis of Construction of Proposed system. There are number of Parameter on which we can consider for evaluation of the performance of system so that to compare with existing work. But initially we have considered the two factors for performance analysis i.e total number of instances that took part in clustering procedure and number of data points been correctly classified according to user query.

The results in Fig. 2 show that our clustering service generates more clusters for smaller number of sites; hence it induces less communication costs within the clusters. On the other hand, the other techniques generate less clusters for large number of sites, thus, they induce more communication costs within the clusters. Fig. 2 shows that the clustering trend increases with the increase in the number of network sites and ours. In contrast, the number of clusters generated by the clustering approach is less due to their clustering approximation function that uses natural logarithmic function. This in turn results in maximizing the number of sites in each cluster which increases the communications cost. Here,

X - Represents the number of generated cluster Y – Represents the number of websites

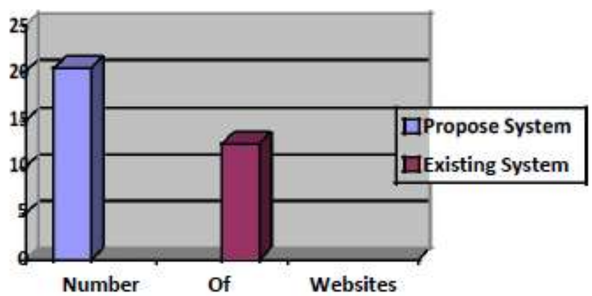

Figure 2: Clustering performance comparison

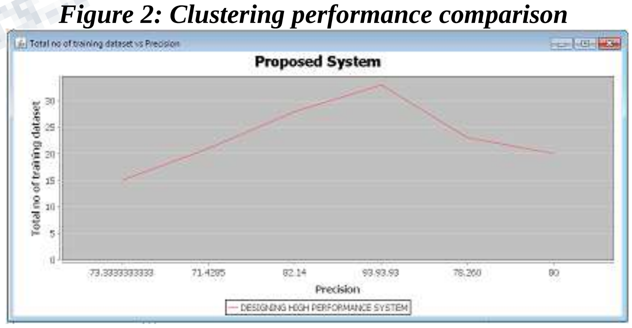

Figure 3: Performance of System

VI. CONLUSION

In this project, we proposed a novel Integrated Fragmentation clustering allocation approach for promoting web telemedicine database system performance. Our proposed approach using threefold approaches, namely, database fragmentation, network sites clustering and fragment allocation. We took a large number of telemedicine database system sites for input. We proposed an estimation model to compute communications cost which helps in finding cost-effective data allocation solutions. The results show that our proposed approach significantly improves service requirement satisfaction in web systems. This conclusion requires more investigation and experiments.

Acknowledgment:

The authors would like to thanks to all the faculty members and Department of Computer Engineering KJCOE, Engineering & Research, Pune, India for the guidance and cooperation.

REFERENCES

[1]Planning High Performance Web-Based Computing Services to Promote Telemedicine Database Management System"IEEE TRANSACTIONS ON SERVICES COMPUTING,JANUARY 2015

[2]S. Lim and Y. Ng, ―Vertical Fragmentation and Allocation in Distributed Deductive Database Systems,‖ *J. Information Systems, vol. 22, no. 1, pp. 1-24, 1997.*

[3]S. Agrawal, V. Narasayya, and B. Yang, ―Integrating Vertical and Horizontal Partitioning into Automated Physical Database Design,‖ *Proc. ACM SIGMOD Int'l Conf. Management of Data, pp. 359-370, 2004.*

[4]S. Navathe, K. Karlapalem, and R. Minyoung, ―A Mixed Fragmen-tation Methodology for Initial Distributed Database Design,‖ *J. Computer and Software Eng., vol. 3, no. 4, pp. 395-425, 1995.*

[5]W. Yee, M. Donahoo, and S. Navathe, ―A Framework for Server Data Fragment Grouping to Improve Server Scalability in Inter-mittently Synchronized Databases,‖ *Proc. ACM Conf. Information and Knowledge Management (CIKM), 2000.*

[6]A. Jain, M. Murty, and P. Flynn, ―Data Clustering: A Review,‖ *ACM Computing Surveys, vol. 31, no. 3, pp. 264-323, 1999.*

[7]Y. Huang and J. Chen, ―Fragment Allocation in Distributed Data-base Design,‖ *J. Information Science and Eng., vol. 17, pp. 491-506, 2001.*

[8]M. Halkidi, Y. Batistakis, and M. Vazirgiannis, ―Clustering Algo-rithms and Validity Measures,‖ *Proc. 13th Int'l Conf. Scientific and Statistical Database Management (SSDBM), 2001.*

[9]H. KhanS and L. Hoque, ―A New Technique for Database Frag-mentation in Distributed Systems,‖ *Int'l J. Computer Applications, vol. 5, no. 9, pp. 20-24, 2010.*

[10]G. Mao, M. Gao, and W. Yao, ―An Algorithm for Clustering XML Data Stream Using Sliding Window,‖ *Proc. the Third Int'l Conf. Advances in Databases, Knowledge, and Data Applications, pp. 96-101, 2011.*

[11]G. Decandia, D. Hastorun, M. Jampani, G. Kakulapati, A. Lakshman, A. Pilchin, S. Sivasubramanian, P. Vosshall, and W. Vogels, ―Dynamo: Amazon's Highly Available Key-Value Store,‖ *Proc. ACM Symp. Operating Systems Principles, pp. 205-220, 2007.*