ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 24, April 2016

B TREE BASED PARALLEL HASHING FOR EFFICIENT MEDICAL IMAGE FILE ACCESSING

BOOVITHA D
PG Scholar
Computer Science & Engineering
Erode Sengunthar Engineering College
Thudupathi,Erode.

KAVITHA C
Assistant Professor(SLG-I)
Computer Science & Engineering
Erode Sengunthar Engineering College
Thudupathi,Erode.

ABSTRACT

Cloud Computing is provided by parallel and distributed technology which enables computing, storage and software services to users.Electronic Medical Record (EMR) is a computerized medical record produced by a care-giving organizationi.e., hospital and doctor's surgery.Patient privacy becomes most important issue inWorld Health Organization (WHO) and in developed countries. Inter-hospital medical information is currently shared using paper-based operations. Motivate the researchers to provide an immediate exchange of electronic medical records for avoiding the duplicate prescriptions or procedures. The existing work presented amedical image file accessing system (MIFAS) built on Hadoop platform to handle exchanging, storing, and sharing of medical images. Presented a strategy for inspecting medical image involves co-allocation mechanism in a cloud environment. Hadoop platform and co-allocation mechanism developcloud environment for MIFAS. HoweverMIFAS produce reasonable redundancy in medical resources. It achieves only an average medical image file accessingperformance in cloud environment. The method is unable to provide medical image exchangingbetween patients and their caregivers. To overcome these specific drawbacks, the proposed work presented B-Tree modeled Parallel Hashing technique to improve medical image file accessing and reduce redundant cloud resource filesin cloud environment. B-Tree indexing is designed to store the medical image file along with its cloud data

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 24, April 2016

node location and provide sequential access.Parallel hashing algorithm works on the medical image files stored in cloud servers and update files in the cloud server storage. The performance measure of the proposed B Tree Based Parallel Hashing for Efficient Medical Image File Accessing technique is conducted on parameters like File redundancy rate, Speed of medical image file accessing, File Location Accuracy on Cloud data nodes.

Key terms: Hadoop platform, B-Tree, Cloud, Medical image

1.INTRODUCTION

Large scale cluster based cloud technology is widely used in the data center and cloud computing environment. The purpose of the field is to overcome the challenges in Medical Image exchanging for storing and sharing issues in EMR (Electronic Medical Record). The importance of the EMR is Patient-centered Care, Collaborative Teams, Evidence-based Care, Redesigned Business Processes, Relevant Data Capture and Analysis and Timely Feedback and Education. Cloud Computing and Virtualization are the most popular computer science fields and is facilitated by the parallel and distributed technology to develop computing, storage and software services to users. The method of Distributed Replicated Block Device is applied to replicate an entire disk image. Medical records are essential documents to store patienthealth care data and information and it works for production of records to all medical institutions which perform clinicalteam in the business. The cloud storage virtualization technology provides many availabilityservices and helps to design a medical imaging system with a distributed file system. Both HDFS servers and traditional streaming media servers are used to support client applications to access patterns characterized using long sequential reads and writes. Medical Image File Accessing System (MIFAS) is developed by HDFS of Hadoop in cloud.

2. LITERATURE SURVEY

In this paper [1], the author proposed a novel scalable B+-tree based indexing method for efficient data processing in the Cloud. Our method can be described as follows. At first, we construct a local B+-tree index for every node where only indexes data relies on the node. Secondly, we develop the compute nodes as a ordered cover and bring out a piece of the local B+-tree nodes to the overlay for well-organized query processing .At last , adaptive algorithm was developed to chose the offered B+-tree nodes in the query patterns.

In the paper [2], the author elaborated a new data structure or a new search algorithm or an adaptation of well-known algorithms and a well-known data structure. The need of the method is to insert an artificial leading key feature to a B-tree index. Here only a particular value for the important B-tree column is present, which is the common and most popular state, the B-tree index is quite like a traditional index.

In the paper [3], the author developed a novel way of indexing moving items by using the classical B-tree without balancinga query and storage efficiency. The inspirationto use theB-tree is threefold. B-tree is commonly used incommercial database systems and is proved to be most effective in queriesand updates. It provides robustness in adjusting workloads and scalable.

In this paper [4], the author established parallel-friendly data structure which provides an effective random access of millions of elements to be developed and accessed at interactive rates. The data structure possesses enormous applications in computer graphics, monitored on usages that requiresaving a sparse set of items in a dense illustration.

In the paper [5], the author depicted a new parallel hashing technique that possess high load factor with a minimum failure rate. And also the method possesses the special merits to destroy coherence in the data and to deny access patterns for earlier presentation. It develops higher area of memory accesses and consistent execution paths in a group of threads. This is suitable to Computer Graphics, where spatial consistency is ordinary.

In this paper [6], the author explains academic investigation of the proposed method and presents the plan of hashing technique. The estimation gives many interesting result. The straightforward technique vastly performs the scattered hashing techniques and present presentation similar with approaches based on shared memory, supercomputers that use particular hardware predicates.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 24, April 2016

In the paper [7], the author proposed a method called MIFAS (Medical Image File Accessing System) to resolve the exchanging, storing and allocation on Medical Images of crossing the various hospitals problems. By using this method, it is possible to improve efficiency of sharing information between patients and caregivers. And also, the system develops the best potential patient-care decisions.

In this paper [8], the author studies about simple variant of the B-tree which is well matchedto use in a concurrent database system. Thismethod expands some of the views in the paper and applies in a model to which the concurrent data structure is stored on secondary storage. And also, a key for B-trees possess the requestof demonstrated practicality.

In the paper [9], the author elaborated an easy method of parallelizing collision searches that really expands the contact of practical attacks. In the hash functions, it starts on with two messages; the first is a message that need target to digitally sign, and the next is a message that the target is eager to sign. By collision search modified for hashing collisions, it is easy toidentifysomewhatdistorted versions of these messages like the two new messages offer the similar hash effect.

In the paper [10], the author described about two parallel machine models which were planned as alternatives to the PRAM: the BSP and the QSM. The PRAM – the most powerful academic parallel machine method – allows algorithm designers to focus on the natural parallelism of the algorithms with no consideration to take machine information inaccount.

In this paper[11], the author presented a new variant of LSH known as Parallel LSH developed to be enormously efficient and able to scale out on multiple nodes and multiple cores that helps high throughput streaming of new data. This method possesses many novel thoughts like cache-conscious hash table layout by 2-level merge algorithm for hash table construction.

In this paper [12], the author briefed about 2-stage index structure which depends on neighborhood indexing and ideal hashing methods. This arrangementarranges a filtering

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 24, April 2016

phase by the neighborhood regions about the seeds in regular time and preventsprobable random memory accesses and branch divergences. And also, it fits mainly well on parallel SIMD processors due to the requirement of intensive homogeneous computational operations.

In this paper [13], the author explained about B-tree data structure to demonstrate the link between B-tree indexing method and computer forensics. B-tree is a rapid information indexing technique that develops indexes to a multi-level set of nodes in which every node consists of indexed information. This method is mostly used in databases and file systems for retrieving therecords saved in a file whileinformation is too large to fit in major memory.

In this paper [14], the author proposed sequential algorithm to plan the algorithm by an abstract model of computation called the random-access machine. Here the machine contains single processor linked to a memory system. Every vital CPU operation consists of arithmetic operations, logical operations and memory accesses in one step. The aim of this method is to improve the algorithm with reserved time and memory requirements.

In this paper [15], the author presented two security methods to assure a safe distribution of medical images in the cloud computing environment using the mean of trust administration by the official parities of the data. It also allows the privacy distribution of the Electronic Patients' Records string data in between the parties to secure the common medical image from the distortion.

3. B TREE BASED PARALLEL HASHING FOR EFFICIENT MEDICAL IMAGE FILE ACCESSING

B-Tree modeled Parallel Hashing technique enhances the medical image file accessing to reduce redundant cloud resource filesin cloud environment. B-Tree indexing is designed to store the medical image file along with its cloud data node location and provide sequential accessable to insertion and/or delete medical image file in the cloud data nodes. It contains binary search tree structure optimized for read and write large blocks of data.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 24, April 2016

Parallel hashing is done on the indexed B-Tree to utilize 128 hash key modelsand to reduce access time as wellreduction in redundancy of medical image file storage. It reduces hash collision. Operation of do write and delete with optimal cloud resource usage and minimal access time are implemented. Parallel hashing algorithm works on the medical image files stored in cloud serversupdate files in the cloud server storage. The method efficiently utilizes the cloud server nodes to reduce redundancy.

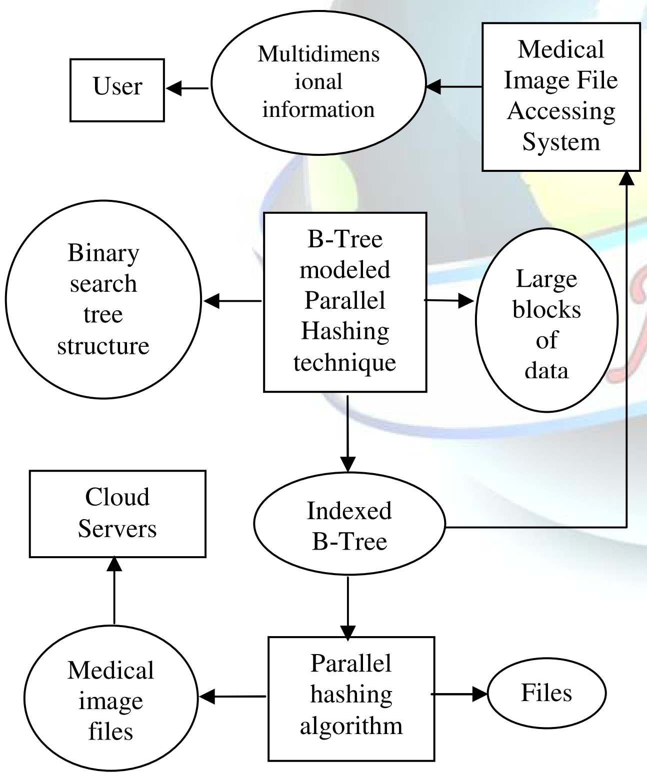

Figure 3.1 Architecture Diagram of B Tree Based Parallel Hashing For Efficient Medical Image File Accessing

It consists of the following modules. They are

a) Medical Image File Accessing System (MIFAS)

b) B-Tree Indexing on MIFAS Cloud Data

c) Parallel Hashing on Indexed Medical Image Files

d) Performance Measure

a)Medical Image File Accessing System (MIFAS)

MIFAS user interface presents multidimensional information to users. Users view the MIFAS portal after passing the validation. Need to pass authentication interface prior to the logging in MIFAS. Medical images obtain from different medical imaging instruments aremagnetic resonance (MR), ultrasound (US), positron emission tomography (PET). Examination category gives list of Medical Examination variety.

Each Hadoop node evaluate real-time state from all members achieves the best approach of transmission data from information service. Middleware helps to collect the needed information fromserver utilization rate, network efficiency, bandwidth between server and server.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 24, April 2016

b) B-Tree Indexing on MIFAS Cloud Data

B-Tree indexing is done on medical image file stored in cloud environment. B-Tree is formed by having binary structure with location of the medical image file on the cloud data node and content key of the medical image file. Indexing of B-Tree is done with location and content key of medical images.

Binary search is initiated on the indexed B-Tree for exact identification of medical image file location and faster accessing from distributed cloud data nodes. Stored medical images are mapped to points in a multidimensional space. Binary search is done sequentially with ability to insert and/or delete medical image file in the cloud data nodesoptimized for read and write large blocks of data.

c) Parallel Hashing on Indexed Medical Image Files

Parallel Hashing applied on indexed medical image file is much faster than sequential searching in cloud data nodes and scaled up for large medical image files. Parallel hash based image file accessing system detect redundant copy of image file in cloud data node. Process storage cloud server maintains a hash tablecontains two fields i.e., a hash signature and its real address. It calculates hash signature for each record requesting for backup by using parallel hash algorithm.

d)Performance Evaluation

Performance evaluation of proposed work are done with following parameters Medical image file size, File redundancy rate, Speed of medical image file accessing and File Location Accuracy on Cloud data nodes.

4. PERFORMANCE METRICS

In this section, we evaluate the performance of B Tree Based Parallel Hashing for Efficient Medical Image File Accessing through SHA-256 hash signature.One of the major contributions of this work isimproves the medical image file accessing and reduce redundant cloud resource files in cloud environment. The performance metrics of the parameters is Medical image file size, File redundancy rate, Speed of medical image file accessing and File Location Accuracy on Cloud data nodes.

The performance metrics are is

- File redundancy rate
- Speed of medical image file accessing

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 24, April 2016

- File Location Accuracy on Cloud data nodes
- File redundancy rate
-

Redundancy file data refers to the variation between the estimated compacted data length of n messages. A determination of redundancy involving two variables is the combined information or a normalized variation. An evaluation of redundancy with numerous variables is represented by the total correlation."n" represents the redundancy of raw data. The rate of a source of information is the normal entropy per symbol. For memory low sources, this is only the entropy of every symbol.

Table: 4.1 Medical image file size (MB) Vs File redundancy rate (%)

Medical image file size(MB)	File redundancy rate (%)	
	Existing MIFAS	Proposed B-Tree
20	26	29
40	30	33
60	32	35
80	36	39
100	42	45

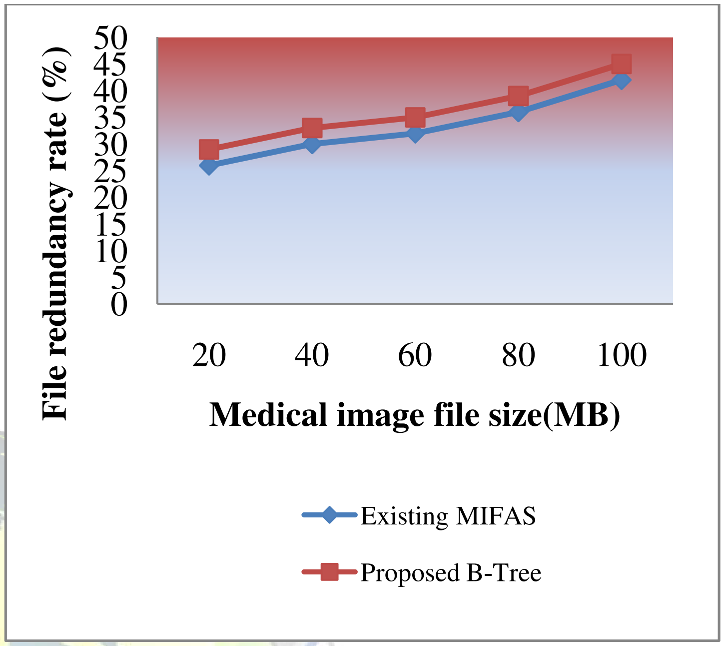

Fig: 4.1. Medical image file size (MB) Vs File redundancy rate(%)

Figure 4.1 demonstrates theFile redundancy rate. X axis represents theMedical image file size(MB)whereas Y axis denotes theFile redundancy rate (%)using both theMedical image file accessing system (MIFAS) technique and the proposedB Tree Based Parallel Hashing for Efficient Medical Image File Accessing technique. Whenthe medical image file size is increased,file redundancy rateget increases consequently. File redundancy rateare illustrated using the existing MIFAS and proposed B-Tree Technique. Figure 4.1.shows the better performance of Proposed B-

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 24, April 2016

Tree method than existing MIFAS and proposed B-Tree.B Tree Based Parallel Hashing for Efficient Medical Image File Accessing method achieves 5to 10% high performance when compared with existing system.

b) Speed of medical image fileaccessing

The average speed is the product of time taken to obtain the public/private key to the number of mobile nodes in the network. The Speed in medical image file accessing isthe rate at which someone or something moves or operates or is able to move or operate.

Table: 4.2 Medical image file size (MB) Vs Speed of medical image file accessing (rps)

Medical image file size(MB)	Speed of medical image file accessing (rps)	
	Existing MIFAS	Proposed B-Tree
20	19	22
40	23	25
60	29	31
80	35	39
100	42	45

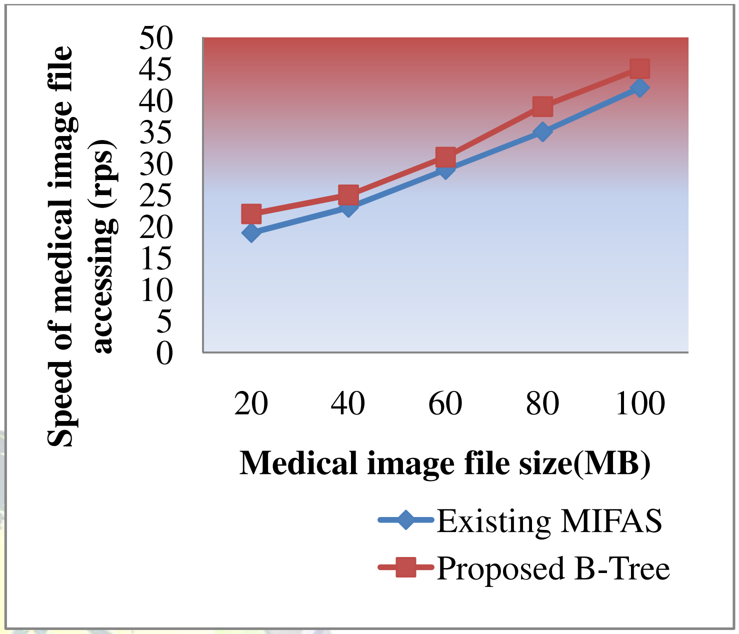

Fig: 4.2. Medical image file size (MB) Vs Speed of medical image file accessing (rps)

Figure 4.2 demonstrates theSpeed of medical image file accessing. X axis represents theMedical image file size(MB) whereas Y axis denotes the Speed of medical image file accessing (rps) using both theMedical image file accessing system (MIFAS) technique and the proposedB Tree Based Parallel Hashing for Efficient Medical Image File Accessing technique. Whenthe medical image file size is increased, Speed of medical image file accessing get increases consequently. Speed of medical image file accessingis illustrated using the existing MIFAS and proposed B-Tree Technique. Figure 4.2.shows the better performance of Proposed B-Tree method than existing MIFAS

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 24, April 2016

and proposed B-Tree.B Tree Based Parallel Hashing for Efficient Medical Image File Accessing method achieves 4 to 9% high performance when compared with existing system.

c) File Location Accuracy on Cloud data nodes

In this section, the FileLocation Accuracy on Cloud data nodes with respect to the training and test data are evaluated with respect to two state-of-the-art methods.Locationaccuracy refers to the ability of the framework to correctly preserve the privacy of the label of new or previously unseen data.

Therefore, the File Location Accuracyis the ratio of size of testing set privacy correctly preserved to the size of the testing set considered for conducting experiments.Location Accuracy is measured in terms of percentage (%). Higher the accuracy more efficient the method is said to be.

Table: 4.3 Medical image file size (MB) Vs File Location Accuracy on Cloud data Nodes (%)

Medical image file size(MB)	File Location Accuracy on Cloud data nodes (%)	
	Existing MIFAS	Proposed B-Tree
20	60	63
40	64	67
60	69	71
80	75	78
100	79	82

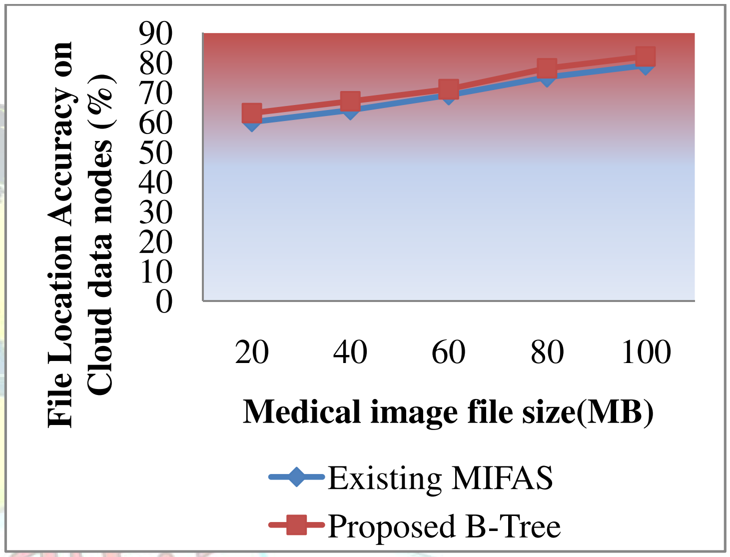

Fig: 4.3. Medical image file size (MB) Vs File Location Accuracy on Cloud data nodes (%)

Figure 4.3 demonstrates theFile Location Accuracy on Cloud data nodes (%). X axis represents theMedical image file size (MB) whereas Y axis denotesFile Location Accuracy on Cloud data nodes (%) using both theMedical image file accessing system (MIFAS) technique and the proposedB Tree Based Parallel Hashing for Efficient Medical Image File Accessing technique. Whenthe medical image file size is

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 24, April 2016

increased, File Location Accuracy on Cloud data nodes get increases consequently. File Location Accuracy on Cloud data nodes is illustrated using the existing MIFAS and proposed B-Tree Technique. Figure 4.3 shows the better performance of Proposed B-Tree method than existing MIFAS and proposed B-Tree.B Tree Based Parallel Hashing for Efficient Medical Image File Accessing method achieves 2 to 5% high performance when compared with existing system.

5. CONCLUSION

This paper proposed a B-Tree modeled Parallel Hashing technique to improve medical image file accessing and helps to reduce redundant cloud resource filesin cloud environment. In this paper a parallel architecture for inline data de-duplication is developed by SHA-256 algorithm to achieve less collision probability. The merit of SHA- 256 algorithm is that it gives 128 bit hash by which the collision probability is lower than prior de-duplication techniques. Inline parallel architecture is implemented for data de duplication toanalyze write and delete worksin various size of data blocks. At last, the positive results are produced for saving operations as well as to delete operations. In this architecture, In-line parallel architecture by server and client storage de-duplication of data is developed. Finally, Modification is completedby various hash algorithms with different block size concept. Execution of this architecture is only possible on Linux kernel.

6. REFERENCES

[1]Sai Wu, Dawei Jiang, Beng Chin Ooi and Kun Lung Wu, "Efficient B-tree Based Indexing for Cloud Data Processing", Proceedings of the VLDB Endowment, Volume 3, Issue 1, 2010, 12 Pages

[2] Goetz Graefe, "Sorting and Indexing With Partitioned B-Trees", Proceedings of the 2003 CIDR Conference

[3]Christian S. Jensen, Dan Lin and Beng Chin Ooi, "Query and Update Efficient B-Tree Based Indexing of Moving Objects", Proceedings of the 30th VLDB Conference, Toronto, Canada, 2004.

[4] Dan A. Alcantara, Andrei Sharf, FatemehAbbasinejad, ShubhabrataSengupta and

Michael Mitzenmacher, “Real-Time Parallel Hashing on the GPU”.

[5] Ismael Garcıa, Sylvain Lefebvre, Samuel Hornus and AnassLasram, “Coherent Parallel Hashing”, ACM Transactions on Graphics, Association for Computing Machinery, 2011, Proceedings of ACM SIGGRAPH Asia 2011, 30 (6).

[6] Long Cheng, Spyros Kotoulas, Tomas E Ward and GeorgiosTheodoropoulos, “Design and Evaluation of Parallel Hashing over Large-scale Data”.

[7] Chao-Tung Yang, Lung-Teng Chen, Wei-Li Chou and Kuan-Chieh Wang, “Implementation of a Medical Image File Accessing System on Cloud Computing”, 2010 13th IEEE International Conference on Computational Science and Engineering, 11-13 Dec. 2010, Pages 321 – 326.

[8] Philip L. Lehman and S. Bing Yao, “Efficient Locking for Concurrent Operations on B-Trees”, ACM Transactions on Database Systems, Vol. 6, No. 4, December 1981, Pages 650 – 670.

[9]Paul C. van Oorschotand Michael J. Wiener, “Parallel Collision Search with Application to Hash Functions and Discrete Logarithms”, 1994 August 17, 9 Pages.

[10]Hyunyoung Lee, “Parallel Hashing Algorithms on BSP and QSM Models”, 2004.

[11] Narayanan Sundaram, AizanaTurmukhametova, NadathurSatish, Todd Mostak, PiotrIndyk, Samuel Madden and PradeepDubey, “Streaming Similarity Search over one Billion Tweets using Parallel Locality Sensitive Hashing”, Proceedings of the VLDB Endowment, Volume 6, No. 14, 2013.

[12] Tuan Tu Tran, Mathieu Giraud and Jean-StephaneVarre, “Perfect Hashing Structures for Parallel Similarity Searches”, 2015 IEEE International Parallel and Distributed Processing Symposium Workshops, 25 - 29 May 2015, Pages 332 – 341.

[13] Petra Koruga and Miroslav Baca, “Analysis of B-tree data structure and its usage in computer forensics”, 6 Pages.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 24, April 2016

[14] Guy E. Blelloch and Bruce M. Maggs, "Parallel Algorithms", Pages 1 – 64.

[15]Fatma E.-Z. A. Elgamal, Noha A. Hikal and F.E.Z. Abou-Chadi, "Secure Medical Images Sharing over Cloud Computing environment", (IJACSA) International Journal of Advanced Computer Science and Applications,Vol. 4, No. 5, 2013, Pages 130 – 137.

[16]AjalaFunmilola A, OpasolaTomilola R, FalohunAdeleye S and FenwaOlusayo D, "Database Management System for a Digitized Medical Image", Journal of Biology, Agriculture and Healthcare, Vol 2, No.3, 2012, Pages 76 – 84.

[17]Luís A. Bastião Silva, "Sharing medical imaging over the cloud services", 12 Pages.

[18]Amir Akhavan, AzmanSamsudin and AfshinAkhshani, "A novel parallel hash function based on 3D chaotic map", Springer, EURASIP Journal on Advances in Signal Processing, 2013, Pages 1 – 12.

[19]IftachHaitner, Yuval Ishai, EranOmri and Ronen Shaltiel, "Parallel Hashing via List Recoverability", Advances in Cryptology -- CRYPTO 2015, 01 August 2015, 19 Pages.

[20]Yantao Li, Di Xiao, Shaojiang Deng, Qi Han and Gang Zhou, "Parallel Hash function construction based on chaotic maps with changeable parameters", Springer, 16 October 2009, 8 Pages.