RESEARCH ARTICLE   OPEN

Analyze the Critical Factors Affecting the Performance Efficiency of Social Networking Databases

Arif Hasan* Dr.P.Sasikala[#]

Research Scholar (Computer Science)Makhanlal University, Bhopal

[#] *Professor ,HOD New Media Technology Makhanlal University ,Bhopal*

Abstract: Before Starting Social Networking Sites RDBMS Was Dominant Database But As Soon As Social Media Came And For Handling It's Unstructured Data Nosql Database Came Into Picture. Nosql Database Many Advantages over Traditional RDBMS. But Due To Rapidly Growing Database Its Performance Effects. There Fore In This Paper Different Critical Factors Like Memory Utilization, Consistency, Replication Are Analyzed So That Performance Of Social Networking Sites Can Be Increased

Keywords: Social Networking Sites, Nosql, Memory Utilization, Consistency, Replication Factors

----

I INTRODUCTION:-

One Of The Major Challenge In Social Networking Site's Database Is To Store, Manipulate And Analyze The Large Amount Of Data. Since According To International Data Corporation (IDC) The Volume Of Social Networking Database Will Reach Up To 32.4 Trillion By The End Of 2020[1].Relationship Exists Between The Social Networking Site's Database Complexity Structure And It's Performance. Different Types Of Nosql Databases Make Strong Framework For Social Networking Database. Different Types Of Social Networking Databases Are Used In Face-Book, Twitter, Linked-In, Whatsapp[2].These Databases Provide Scalability Features In Social Networking Databases That Is Why Their Complexity Is Comparatively Less Then Traditional RDBMS. Since These Social Networking Nosql Database Do Not Contain Any Fix Format Of Data Storage And Their Corresponding Different Keys For Retrieving A Data. But In Social Networking Database Huge Amount Of Data Is Stored That Result In To Degradation Of Efficiency Of Social Networking Database Therefore

Different Nosql Databases Are Used At Different Places According To Their Applicability So That Efficiency Of Social Networking Database Can Be Increased. In This Paper Key Features Of Four Different Social Networking Site's Nosql Databases Are Assessed And Compared On Different Parameters. And Then Analysis Is Done On The Basis Of These Comparisons.

II CLASSIFICATION OF SOCIAL MEDIA – NOSQL

Social Media Uses Different Nosql Technologies For Its Different Purposes. Myspace Uses Dynamodb[3] For Its Storage Mechanism. Facebook Uses Cassandra[4] Which Is Open Source And Column Oriented Database As A Core Database Along With Neo4j[5] Which Is Open Source Database Used A Structure Model Which Can Be Further Used For The Analysis Purpose.

Table I

Social Networking Sites	Nosql Database Subcategories Used
Facebook	Cassandra, Hbase, Ne04j
Twitter	Flockdb, Cassandra, Hbase, Ne04j
Linkedln	Voldemort, Mongodb, Hbase, Allegrograph
Flickr	Mongodb, Neo4j
Friendfeed	Hbase, Cassandra, Orientdb
Foursquare	Mongodb, Couchdb, Riak, Cassandra, Infogrid
Myspace	Mongodb, Dynamodb, Ne04j

Table I Describes Different Databases Used For Different Social Networking Sites. These Define As A Primary Parameters Used In A Efficient Query Processing Since These Databases Have Their Different Structure Model.

All The Above Social Networking Sites Nosql Databases Can Be Classified One Of The Following Categories.

(I)Wide Column Family Store[6]: Which Has Distributed Column Oriented Database

That Consist Clustering Column Attributes Ex.Cassandra ,Hadoop/Hbase.

(Ii)Document Store[7]: Which Has Centralized Document Database That Consist Congregate Associated Documents.Ex Mongodb,Couchdb.

(Iii)Key Value/Tuple Value Store[8]: Which Has Centralized Key For Tuple Value Store That

Consist No Unique Key. Ex.Dynamodb, Riak, Voldmort.

(Iv)Graphdb[9]: Which Has Distributed Nodes With Keys That Are Completely Connected.

Ex.Neo4j

III ANALYZING CRITICAL FACTORS FOR INCREASING THE PERFORMANCE EFFICIENCY OF SOCIAL NETWORKING DATABASE

(i) Memory Utilization: From The Following Chart It Is Revealed That Different Social Networking Databases Have Different Memory Utilization While Executing Query And Fetching Records.
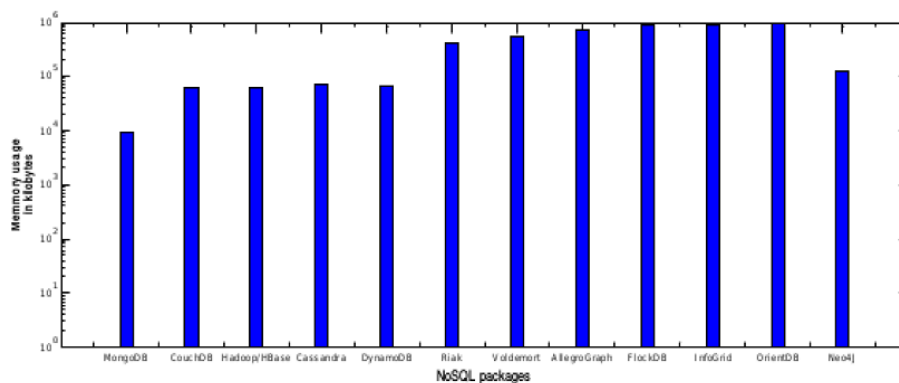
(Ii) Concurrency Control: Table II Reveals Concurrency Control Of The Different Social Networking Sites. Different Concurrency Control Defines Fault Tolerance Capacity Of Social Networking Database. Therefore By Maintaining Different Concurrency Control In Social Networking Sites We Can Maintain It's Fault Tolerance Capacity That Result Into Efficient Performance Of Social Networking Site's

Table II

Social Networking Database	Concurrency Control
Cassandra	Optimistic locking
MongoDB[10]	Locks attained separately for read and Write
CouchDB[11]	Multi-granularity
DynamoDB	Optimistic locking
Riak[12]	Optimistic locking and plug in data storage
Voldemort	ACID with no locks
Neo4J	Multi Version Concurrency Control

(Iii)Consistency In Storage/Replication: Consistency Level Defines Any Transaction Must Change Only Affected Data. Any Data Written To The Database Must Be Valid According To All Defined Rules Including All Constraints. These Provides Data Security In Social Networking Sites. In A Similar Way Replication Defines How Many Replica Of A Data Item So That If Copy Fails We Can Copy It From Other Replica.

Table III Defines Different Social Networking Databases How They Maintain Their Consistency Level And Replication And By Maintaining These Factors One Can Increase Performance Efficiency Of Social Networking Sites.

Table III

Social Networking Database	Consistency In Storage/Replication
Cassandra	Disk Level Consistency And Asynchronous Replicate Factor
Mongodb	Disk Level Consistency And Asynchronous Replicate Factor
Couchdb	Partial Framework Consistency And Asynchronous Replicate Factor
Dynamodb	Eventual Consistent And Asynchronous Replicate Factor
Riak	Consistent Synchronous Replication
Voldemort	Asynchronous Replication And Consistency At Document Level
Neo4J	Consistency High At Each Node And Replication Is Asynchronous And Persistent

IV CONCLUSION

Performance Efficiency Of Social Networking Site Is Critical While Considering Huge Amount Of Structured, Semi-Structured And Unstructured Data Since These Are Complex And Innumerous. This Steer Need For Reliable And Rapid Services From Social Networking Sites That Result In To Requirement Of Performance Efficiency Of Social Networking Sites.

In This Paper Four Types Of Nosql Database Are Analyzed And Precisely Each Database Is Taken In Account And Then Find Out Its Critical Factors And Shown How It's Performance Can Be Increased So That In Over All We Can Improve The Performance Of Social Networking Sites.

REFERENCES

[1]. I] J. Gantz And D. Reinsel, " The Digital Universe In 2020: Big Data, Bigger Digital Shadows, And

[2]. Biggest Growth In The Far East," IDC Iview: Idcanalyze The Future, Vol. 2007, Pp. 1-16, 2012

[3]. [P. Groves, B. Kay Y Ali, D. Knott, And S. Van Kuiken, " The 'Big Data' Revolution In

[4]. Healthcare," Mckinsey Quarterly, 2013.

[5]. G. Decandia, D. Hastorun, M. Jampani, G. Kakulapati, A. L Akshman,A. Pilchin, S.

[6]. Sivasubramanian, P. Vosshall, And W. Vogels, "Dynamo: Amazon's Highly Available Key -

[7]. Value Store," In SOSP, Vol. 7, Pp. 205-220, 2007.

[8]. L. George, Hbase: The Definitive Guide. O'reiuy Media, Inc. , 2011.

[9]. G. Vaish, Gelling Started With Nosql. Packt Publishing, 2013.

[10]. G. Vaish, Gelling Started With Nosql. Packt Publishing, 2013.

[11]. B. G. Tudorica And C. Bucur, "A Comparison Between Several Nosql Databases With Comments

[12]. And Notes," In Roedunet International Conference (Roedunet), 2011 1 0th, Pp. 1-5, IEEE,

[13]. 2011.

[14]. R. Cattell, "Scalable Sql And Nosql Data Stores," ACM SIGMOD Record,Vol. 39, No. 4, Pp. 12-

[15]. 27, 2011.

[16]. N. Developers, "Ne04j," Graph Nosql Database [Online], 2012

[17]. 1. Han, E. Haihong, G. L E, And 1. Du, "Survey On Nosql Database," In Pervasive Computing

[18]. And Applications (ICP CA), 20 11 6th International Conference On, Pp. 363-366, IEEE, 2011.

[19]. 1. C. Anderson, 1. L Ehnardt, And N. Slater, Couchdb: The Definitive Guide. O'Reilly, 2010.

[20]. D. Bartholomew, "Sql Vs. Nosql," Linux Journal, Vol. 2010, No. 195,P. 4, 2010.