A SURVEY ON CLOUD DBMS ENABLED DATA TRANSACTIONS AND DATA STORAGE

A. Aasha Begum
Part-Time Research Scholar
Dept of Computer Science
Madurai Kamaraj University
Madurai

Dr.K.Chitra
Assistant Professor
Dept of Computer Science
Govt. Arts College, Melur
Madurai – 625 106

Abstract

The cloud based environment enabled the distributed system along with the accessing of data at any instances. The cloud framework that makes the direct transmission along the users at any instances of time. The database management system that provides the acknowledgements and the requests along through its corresponding users to activate the way of data storage. Although the various manipulated works comes under this way of data storage made through the database transactions. For instances of data storage and database transaction provides the distributed service with the path of laaS (logging-as-a-service) and daaS (data-as-a-service). Due to the database transaction may leads the drawbacks by affecting the data storage. From this problem, to influenced the data optimization, data lifecycle management system (DLMS) and its ACID properties along with the quality of data also have the maximum number of stored data. The survey focuses on the data optimization, lifecycle of the data storage with its functional working approaches.

Key Words: *Cloud DBMS, Data Optimization, Data Lifecycle Management Systems, laaS, daaS, ACID Properties.*

I. INTRODUCTION

In recent years, the easy access of data transactions and data storage attains the cloud DBM. The cloud framework always enables the improved way of data transactions along with their respective acknowledgements and requests of read and write data transactions. DBMS organizing the distributed database and the data to be stored within the physical address. The database information accessed within the time consideration. The data transactions may leads to the execution of scheduled process with the requests and acknowledgements received. The replication and control approaches used to avails the database organization [1]. The cloud framework provides the start-up information for the long-time working environment and also reduced the resource management cost along with the proper optimized data storage. The growth of the service and its efficiency always depends upon the resiliency and scalability of data transactions [2]. The overall growth of the service and the amount of data due to the transactions makes the difficulties for Data Base Administors (DBA) in terms of affected the capacity. Due to this problem, the control and consistency of data to be optimized [3]. The storage and placement of databases evenly distributed along with the cloud framework and its efficient data transaction, which could enables the users for proper data management while some existing problem arise. The robust analytics of data, increased quality of data due to the transactions. The complexity of adequate data storage and the realistic cloud framework makes the improved data transaction in cloud DBMS [4]. The consistent and replication of data transaction the uncharacteristic interaction of low level leads to make the security problem within the cloud framework. Cloud DBMS based on the determination of storage security, distribution system,

and optimized database information systems. The read and write function. The migration of cloud framework that leads to the high amount of data to be maintained within the quality and integrity of data along with its corresponding data transaction process [5].

II. CLOUD DBMS BASED DATA TRANSACTIONS

The cloud infrastructure enables the database transaction, read and write data storage, low response time, high performance of cloud service along with the quality of data storage. The functions that makes the concurrent of data transactions within the database transaction response. This cloud DBMS ensembles the large scale data analysis with the running database transactions. The transaction of database system must maintain consistency, concurrency, scalability, and durability of improved data storage. The integrity of data makes the inconsistency of data within the given cloud framework along with the ACID Properties.

III. Data Lifecycle Management System (DLMS)

The Data Lifecycle Management System (DLMS) focuses on the data. The data storage planning and provisioning of data to be executed within the data placement of cloud framework. The exploitation and the number of vulnerability of data attain the more complex systems and the threat model of data management systems and database systems. The number of accessing and storage centers gets increased. The data processing covered along with the application specific tools through the database transactions [4]. This approach provides the management metadask, data of placement, data storage, data investigation, processing of big data, resource management to enables the database transactions. The data management is the process of organizing several works through some particular data storage, processing, and transmissions along with the database control and its security. DLMS is the process of managing the susceptible data with the less exploitation of data management, which is inference and vulnerable among the feature of data in the complex form of systems. Due to the big data transmission, the physical access to the DBMS server logging to the nodes of transmission. The transmission nodes spoofing the data to modify or change along their property of processing node. The processing node increases the number of individual works. This may access the intruder gain with the cloud framework and resources of computer technologies [6].

The increased queue problem may leads to eliminate by using the distributed systems along with the database transactions. The processing of distributed systems that eliminates the node processing and improves the efficiency. The important problem such as inefficient account featured in the resource processing. This also provides the data integrity, DB access rate, control over the data transaction nodes [7]. The social networking service provides the efficient large scale and the long term of data along with the storage and processing of massive data. The number of users may increased due to the life cycle of data and its efficient of node transmission [8]. The single instances of partitioned data avails the cloud services with the large datasets of data transmission. The highly used cloud database processing systems. The database and its working function always dynamic along with the long term of data [9]. The change in weather and climate makes the conscious life cycle timings. The effects of climate change along with the key indicators. The regional or global process defined as climate, which requires characterization of life cycles on a large scale. The time series based vegetation life cycles to be estimated through the characterization [10].

The open access of data to be described through the features of node transmission. The communication occur more frequently over the the type of database transactions to enables the number of data access. The traditional approach to electronic publication storage and access through the interface of

full text search systems is most common today; however, due to the growing volumes of electronic information and features of the electronic publication life cycle, the use of the standard services and search tools of the Internet that relate to electronic scientific information has become less effective [11]. The set of physical and digital services makes the object augmented physical life with the regular time bound. The users provides the networking capabilities, data processing, data storage. The continuous streams of contextual data and geolocalized through lifecycle SO. The distributed software system provides the surrounding of environment along with the related surrounding users [12]. The retrieval performance enabled the encrypted data with the high speed of data query processing. The external indexing file making the large volumes of data with the inclusion of archival management. The sensitive information encrypted and prevented an index file. The external indexing file leaked itself [13]. Hadoop Map Reduce approach critical to scale the multiple cloud framework in the batch processing frameworks. An inter cloud data transfer and overhead problem in synchronization due to latencies of data storage without data provisioning approach. An unified computing resources makes the dynamically provisioned cloud framework and the collection of data to be interconnected and virtualized in the compute frameworks[14]. In cloud computing environment, the big data and the database systems enables much higher. NoSQL database systems confines the big data along with the property of relational database, big data, and hybrid database within the cloud service blooms [15].

IV. Data Optimization

The data optimization algorithm is the process of reduced the queries within the allocated time frame, which solves the lowest-cost problem to enables the database. The process of natural selection based genetic algorithm that is based on search based heuristic algorithm. The current generation are selected along with the reproduction of individuals for reproduction of the next generation that produce the offspring of data [16]**.** The different data suppliers makes the security along with the insured quality of information as increased. The important is trading of secret data enables the outright essential for trading secret data such as distribution, approval of an information and exposure frequently. The data optimization may leads to the safety and the web services depends trade over the data [17]. The data optimization approach that enables the performance of large query workloads with the improved data storage and database transactions. The data base is parameterized along with the quality of node transmission and the query information with the optimized data [18]. The response time to be managed by the optimization of QoS requirement also the security strength with its changing the security mechanisms [19].

V. logging-as-a-service (laaS) and data-as-a-service (daaS)

In laaS system, the majority of data of bare mail server designed to work with the logging in to the server also the limited group of users to communicate to the SQL based database to send and receive email within the secured manner of physical data administrator. The sender's username must be logged in to delete a sender's messages along with the delete mail messages [20]. For the process of transaction modification operations, the different nodes are executed. The hoster on the initiator node can be overcome by logging the modification information within the reliability of information. The execution of SQL may leads to the initiated node along with the processing the host information. The master node is used to track the state of queue and disables the unloading of information [21]. The logging data to be taken from the distributed sources to ensembles the data storage validation. The users allows the particular persistence in memory and partition the dataset across nodes, which makes the tolerance of faults, atomicity of granular via the partitions and transaction along with the provided logging of data replication in laaS (logging-as-a-service). The cloud framework that configures the logging files across the database transactions. Logging permits the system administrators to enables the easy access into the actions [22]. In web service enables an

object-oriented web-based interface. The several sensor devices that utilizes the database server, which is sensed by the stored data or regular server managing actualized set of information.

From daaS system, daaS commended the process of data monetization concept to enables the database transactions within the given reasonable information in daaS (data-as-a-service). Data-as-a-service (DaaS) provides the large-scale databases within the cloud framework. This allows the efficacy of host and manages the data transactions. The well-known data storage and database transaction synchronized and encrypted the index of data within the daaS (data-as-a-service). The range of query count to be manipulated and responsible for outsourced database systems.

. An encrypted index of anonymized data in a DaaS service that is responsible for answering range count queries from and its database transactions. To ensure data confidentiality and integrity of outsourced databases should be considerable effort for suggest encrypting the data before entering it into the cloud. It is less effective in deterring inference of data attacks and the data to be confidentially maintained. The `secured data provides the confidentiality with the encryption of data simultaneously. The new privacy-enhancing technologies secure and provide the inference attacks to enables the query answering with real time bound [23]. The cost of transactions gives the preferences among the consumers and reducing the speeding up the database. The different types of data produce the various kind of sharing application among the users to occupies the data producers, and consumers and their assets [24]. The cloud computing makes the open source to the space within the internet discovery. The hardware provides the portability and better installation of equipment. DaaS (data-as-a-service) allows the remote storage and backup of data with the easy access of programmes. An external cloud provider gives the mass storage in an outsourcing form of local LAN. The internal users logging into the worker's workspace to makes the data transactions [25].

The transaction through online in real time data is the crucial part, which is generated by cloud framework for social access or maps the data storage. The real-time analysis and streaming makes the visualization due to big data's current infrastructure and its challenging causes the use of efficient and extraction of information within the time consideration [26]. The data storage and data management gives the function of relational database model , which is considered as a leading model. NoSQL and New SQL exploits the function of Big Data explosion provides the high volume of data with the alternate models. The improved communication technology produced the database systems with the change of architecture function.

. The technologies mostly focus on performance guarantees to can ensure the security and privacy of the information they handle. The different types of integrated security mechanism provide the different database systems in a big data applications [27]. The distributed DBMSs provides the non-relational database systems that can be enables the overlooked and the security requirements. The DBMS enables the four possible ways such as management transparent through distributed data, easy access of data with its improved performance, transaction of database like as distributed models of easy expansion. The concurrent execution of user provides the full transaction and supports the guarantees in DBMS do not provide the consistency of database. The only one exaction provides the query in the user and the correct transaction given the time bound of integrity of database [29].

VI. ACID properties:

1. The transaction execution ensures the Atomicity.
2. The correct execution maintains the Consistency in the data transaction.

3. The effect of concurrent transactions Isolation indicates correctness of data transactions execution within the shielded commit to each other.
4. The effect of committed transactions ensures the database transaction that permits the permanent the system crashes as attains the durability [28].
5. The high throughput rate makes the scalable and composed manner to equalize the transactions that has many data storage process to maintain the database transactions and its important advantages over other existing systems.

The NVM based architecture provides the storage of data in the subsystem with the reducing the number of write operations. Data Storage Subsystems provides the external memory with the subsystems as various data storage to be developed and optimized. These are

- The renovate data in the subsystem makes the current storage location
- To launch a new copy of the data element when the subsystem updated
- The journal-structured launch in the subsystem [29].

The approach of database-as-a-service (DBaaS) enabled in the cloud framework to access the databases that presents easy access management of data varied challenges. The relational databases in NoSQL, which manages the data with the decision making process. The cloud providers through the cloud databases, which is preconfigured virtual machines. The Database-as-a-service (DBaaS) model. The universe database includes the number of information present in the data transactions, with the part of the household's works and population [30]. The problem to the data along with the data centers with the migration of huge stored data and compute the better data storage location that enhanced the total number of data placement cost also the throughput of the performance [31].

SURVEY TABLE OF DATABASE TRANSACTIONS IN CLOUD DBMS FOR DATA STORAGE

S.No	Paper	Methods	Advantages	Limitations
1	N. Chauhan et al. (2019)	Replication of data	Dynamic replication	Limited storage capacity
2	W. Sul et al. (2019)	Decoupled database	Availability, reliability	Effort for good performance
3	D. Tomar, J. P et al. (2019)	NoSQL relational database based healthcare system	More flexible and easy access for cloud framework	Can't establish the relational database
4	S. Mazumdar et al. (2019)	Data storage placement methods	Provide multi-domain computing, high throughput	Optimal data storage
5	N. Semenov et al. (2019)	Cloud based architecture security	Harmonization of security	Conflicts in Data processing
6	M. Poltavtseva, et al. (2019)	Big data management and threat	Intruder model based vulnerable data	Information security

		model	storage	
7	E. Gusev,et al. (2019)	Optimizing access to memory pages	Queue problem and increased efficiency	Oracle RAC systems
8	K.-Y. Whang, et al. (2019)	Building social networking systems	ACID properties and scalability	Not enables in relational database
9	H. Leena et al. (2019)	Data optimization and portioning of data	Handling large volume of database	Specific tuple management
10	M. Johnson et al. (2019)	Bayesian dynamic linear models	Phonological event estimation	Shortcoming of processing stage
11	A. Elizarov et al. (2014)	Electronic specific journal-management	Implementing within the framework	Less effective
12	G. Fortino et al. (2016)	Cyberphysical digital libraries	Integration of database management	Critical issues in addressing the database
13	C. Ho et al. (2019)	Performance of encrypted databases	Impact of high performance	Speed of query processing
14	K.-L. Du et al. (2019)	Big data, cloud computing, IoT	Synchronization of database systems	Overhead problem
15	C. Li and J. Gu (2019)	Integration approach for hybrid databases	Flexible manner	Complexity problem
16	S. Dias et al. (2020)	Query Time Optimization	Handling the processing of data	Less database transactions
17	M. Agarwal et al. (2019)	"“Big” Data Management in Cloud Computing	Massive growth in size of data	Problem in relational database management systems
18	A. Beirami et al. (2019)	Trusted relational databases	Immutable data transactions, temporal levels	Large query workloads
19	M. S. Khatib et al. (2020)	"FGSA for optimal quality of service based transaction	Number of user requests increased	Intrusion attack, decreased database systems
20	Ghandeharizadeh, S., et al. (2019)	Nova: Diffused Database Processing Using Clouds	High speed network optimization	Workload problem
21	Gusev, E. (2019)	Optimizing Access to	Improves the efficiency and	Threshold time in packet

		Memory Pages	improved distributed systems	delivery problem, transfer function reduced
22	Firouzi, F., & Farahani, B. (2020)	Architecting IoT Cloud	Improved performance and connectivity	Limited processing and less storage capacity
23	Dagher, G et al. (2019)	"privacy-preserving data outsourcing framework	Confidentially encryption of data and more flexible	Increased data size, defective harmonic problems
24	Porter, L., et al. (2019)	planning, Land and Housing in the Digital Data Revolution/The Politics of Digital Transformations	Provides best data transactions, encryption done confidentially	Inference of data attack
25	Visconti, R. M. (2020)	The Valuation of Software and Database.	Integrated data storage	Intangible data evolution
26	Amalina, F., et al. (2019)	Blending big data analytics	Cyber security	Inefficient wide range of problems
27	Samaraweera, G. D., et al. (2019)	Security and Privacy Implications	Deploy cloud based database transactions	Less guarantees of data performance
28	Özsu, M. T et al. (2020)	NoSQL, NewSQL, and Polystores	Enhanced performance	Does not violate database consistency
29	Kuznetsov, S. (2019)	Towards a Native Architecture of In-NVM DBMS	Simplicity and efficiency of database transactions	Overhead due to the many data storage and data transactions problem
30	WM Ribeiro, M et al. (2019)	OLAP parallel query processing in clouds	Capability of high performance, database replication	Long time data processing, critical decision making process

Conclusion:

Cloud computing is use of computing resources as a service via an internet. It is a very promising technology for the future with several advantages like pay per use, availability, elasticity etc. In this paper we have discussed the various developments that have taken place in this field. There are four different models for deployment of a cloud: public, private, hybrid and community. Cloud Service providers offer their services through several service delivery models. The various cloud service models are: Software as a Service, Platform as a service, database as a service and Infrastructure as a Service .Merits and demerits of migrating applications and infrastructure of an organization must be considered before a realistic migration process is carried out in this paper.

REFERENCES:

[1]N. Chauhan and S. P. Tripathi, "QoS Aware Replica Control Strategies for Distributed Real Time Database Management System," *Wireless Personal Communications*, vol. 104, pp. 739-752, 2019.
[2]W. Sul, H. Y. Yeom, and H. Jung, "Towards Sustainable High-Performance Transaction Processing in Cloud-based DBMS," *Cluster Computing*, vol. 22, pp. 135-145, 2019.
[3]D. Tomar, J. P. Bhati, P. Tomar, and G. Kaur, "Migration of healthcare relational database to NoSQL cloud database for healthcare analytics and management," in *Healthcare Data Analytics and Management*, ed: Elsevier, 2019, pp. 59-87.
[4]S. Mazumdar, D. Seybold, K. Kritikos, and Y. Verginadis, "A survey on data storage and placement methodologies for cloud-big data ecosystem," *Journal of Big Data*, vol. 6, p. 15, 2019.
[5]N. Semenov and A. Poltavtsev, "Cloud-Based Data Architecture Security," *Automatic Control and Computer Sciences*, vol. 53, pp. 1056-1064, 2019.
[6]M. Poltavtseva, D. Zegzhda, and M. Kalinin, "Big Data Management System Security Threat Model," *Automatic Control and Computer Sciences*, vol. 53, pp. 903-913, 2019.
[7]E. Gusev, "Optimizing Access to Memory Pages in Software-Implemented Global Page Cache Systems," *Programming and Computer Software*, vol. 45, pp. 497-505, 2019.
[8]K.-Y. Whang, I. Na, T.-S. Yun, J.-A. Park, K.-H. Cho, S.-J. Kim, *et al.*, "Building social networking services systems using the relational shared-nothing parallel DBMS," *Data & Knowledge Engineering*, p. 101756, 2019.
[9]H. Leena, B. Premasudha, and P. Basavaraja, "Data optimisation and partitioning in private cloud using dynamic clusters for agricultural datasets," *International Journal of Dynamics and Control*, pp. 1-13, 2019.
[10] M. Johnson, P. C. Caragea, W. Meiring, C. Jeganathan, and P. M. Atkinson, "Bayesian Dynamic Linear Models for Estimation of Phenological Events from Remote Sensing Data," *Journal of Agricultural, Biological and Environmental Statistics*, vol. 24, pp. 1-25, 2019.
[11] A. Elizarov, D. Zuev, and E. Lipachev, "Electronic scientific journal-management systems," *Scientific and Technical Information Processing*, vol. 41, pp. 66-72, 2014.
[12] G. Fortino, A. Rovella, W. Russo, and C. Savaglio, "Towards cyberphysical digital libraries: integrating IoT smart objects into digital libraries," in *Management of Cyber Physical Objects in the Future Internet of Things*, ed: Springer, 2016, pp. 135-156.
[13] C. Ho, K. Pak, S. Pak, M. Pak, and C. Hwang, "A Study on Improving the Performance of Encrypted Database Retrieval Using External Indexing System of B+ Tree Structure," *Procedia Computer Science*, vol. 154, pp. 706-714, 2019.
[14] K.-L. Du and M. Swamy, "Big Data, Cloud Computing, and Internet of Things," in *Neural Networks and Statistical Learning*, ed: Springer, 2019, pp. 905-932.
[15] C. Li and J. Gu, "An integration approach of hybrid databases based on SQL in cloud computing environment," *Software: Practice and Experience*, vol. 49, pp. 401-422, 2019.

[16] S. Dias, S. Kolhe, R. Shinde, R. Chaudhari, and R. M. Wahul, "Query Time Optimization Using Hungarian Algorithm," in *ICCCE 2019*, ed: Springer, 2020, pp. 271-276.
[17] M. Agarwal and G. M. S. Srivastava, "“Big” Data Management in Cloud Computing Environment," in *Harmony Search and Nature Inspired Optimization Algorithms*, ed: Springer, 2019, pp. 707-716.
[18] A. Beirami, Y. Zhu, and K. Pu, "Trusted relational databases with blockchain: design and optimization," *Procedia Computer Science*, vol. 155, pp. 137-144, 2019.
[19] M. S. Khatib and M. Atique, "FGSA for optimal quality of service based transaction in real-time database systems under different workload condition," *Cluster Computing*, vol. 23, pp. 307-319, 2020.

[20]. Ghandeharizadeh, S., Huang, H., & Nguyen, H. (2019, May). "Nova: Diffused Database Processing Using Clouds of Components [Vision Paper]. In *International Conference: Beyond Databases, Architectures and Structures,"* (pp. 3-14). Springer, Cham.

[21]. Gusev, E. I. (2019). "Optimizing Access to Memory Pages in Software-Implemented Global Page Cache Systems. *Programming and Computer Software*," *45*(8), 497-505.

[22]. Firouzi, F., & Farahani, B. (2020). "Architecting IoT Cloud. In *Intelligent Internet of Things"* (pp. 173-241). Springer, Cham.

[23]. Dagher, G. G., Fung, B. C., Mohammed, N., & Clark, J. (2019). $${\textsf {SecDM}} $$ SecDM: "privacy-preserving data outsourcing framework with differential privacy. *Knowledge and Information Systems*," 1-38.

[24]. Porter, L., Fields, D., Landau-Ward, A., Rogers, D., Sadowski, J., Maalsen, S., ... & Bates, L. K. (2019). "Planning, Land and Housing in the Digital Data Revolution/The Politics of Digital Transformations of Housing/Digital Innovations, PropTech and Housing–the View from Melbourne/Digital Housing and Renters: Disrupting the Australian Rental Bond System and Tenant Advocacy/Prospects for an Intelligent Planning System/What are the Prospects for a Politically Intelligent Planning System?. *Planning Theory & Practice*," *20*(4), 575-603.

[25]. Visconti, R. M. (2020). "The Valuation of Software and Database. In *The Valuation of Digital Intangibles"* (pp. 193-217). Palgrave Macmillan, Cham.

[26]. Amalina, F., Hashem, I. A. T., Azizul, Z. H., Fong, A. T., Firdaus, A., Imran, M., & Anuar, N. B. (2019). Blending big data analytics: Review on challenges and a recent study. *IEEE Access*, *8* 3629-3645.

[27]. Samaraweera, G. D., & Chang, M. J. (2019). Security and Privacy Implications on Database Systems in Big Data Era: A Survey. *IEEE Transactions on Knowledge and Data Engineering*.

[28]. Özsu, M. T., & Valduriez, P. (2020). NoSQL, NewSQL, and Polystores. In *Principles Of Distributed Database Systems* (pp. 519-557). Springer, Cham.

[29]. Kuznetsov, S. (2019, November). Towards a Native Architecture of In-NVM DBMS. In *2019 Actual Problems of Systems and Software Engineering (APSSE)* (pp. 77-89). IEEE.

[30]. WM Ribeiro, M., AB Lima, A., & de Oliveira, D. (2019). OLAP parallel query processing in clouds with C-ParGRES. *Concurrency and Computation: Practice and Experience*, e5590.

[31]. Mazumdar, S., Seybold, D., Kritikos, K., & Verginadis, Y. (2019). A survey on data Storage and placement methodologies for cloud-big data ecosystem. *Journal of Big Data*, *6*(1), 15.