ISSN: 2395-7852

International Journal of Advanced Research in Arts, Science, Engineering & Management

Volume 10, Issue 5, September 2023

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 6.551

Optimization for Dynamic Replication in Cloud Center

BAIRYSETTI PRASAD BABU, DR. PRASADU PEDDI, DR. SUNEEL PAPPALA

Research Scholar, Dept. of Computer Science & Engineering, Shri JJT University-Rajasthan, India

Research Supervisor, Dept. of Computer Science & Engineering, Shri JJT University- Rajasthan, India

Co-Supervisor & Associate Professor, Dept. of Computer Science & Engineering, Lords Institute of Engineering and Technology, Hyderabad**,** India

ABSTRACT: The cloud Storage Providers (CSPs) present geological documents are put away with a few memory space classes with arranging costs. A noteworthy issue encountering by the cloud customers can be the methods by which to blow this additional room classes for you to serve a license request with a period, contrasting residual job that needs to be done about its materials at any rate aggregate expense. Records intermittent notwithstanding replication approaches have a fundamental activity around appropriated methodologies over the electronic encouraging. In this study arranged DRA Criteria to determine the technique which will complete a capacity zone strategy in which utilizes the specific Division and even Replication approach to putting away the data. In this procedure, the report will be isolated and these zones will be replicated and blend as demonstrated by the duplication factor before taking care of the idea upon the cloud. Usually, the parts are normally dissipated with the end goal that progressive clients in the electronic encouraging don't hold the areas of any proportionate convenience in this way while a PC is yielded no huge information is emphatically spilled on the attacker. This structure will improve cloud prosperity gauges using Category and Duplication of data on Cloud proposed for Optimal Functionality and Safety gauges reasoning. Cloud Storage Providers (CSPs) offer geographically data stores providing several storage classes with different prices.

KEYWORDS: Cloud computing, Data centers, Geology, Aggregates, Cloud Storage Providers (CSPs), Organizations, Cognition.

I. INTRODUCTION

First of all, it enables better data availability. If a system at one site goes down because of hardware issues or other problems, users can access data stored at other nodes. Furthermore, data replication allows for improved data backup. Since data is replicated to multiple sites, IT teams can easily restore deleted or corrupted data. Data replication also allows faster access to data. Since data is stored in various locations, users can retrieve data from the closest servers and benefit from reduced latency. Also, there's a much lower chance that any one server will become overwhelmed with user queries since data can be retrieved from multiple servers. Data replication also supports improved analytics, by allowing data to be continuously replicated from a production database to a data warehouse used by business intelligence teams.

There is a number of works in the literature that deals with data replication in cloud systems. These strategies were proposed for improving performances, reducing the bandwidth consumption, increasing the level of data availability and load balancing. However, these objectives appear to be conflicting. For example, replicating data ensures the availability. However, this is done on the detriment of communications between sites, which overloads the network and then, affects performances. Furthermore, most of these strategies neglect both the replication cost and the provider profit. We propose a dynamic data replication strategy that simultaneously satisfies availability and performance requirements while the Profit of the provider is taken into account (DRAPP). We focus on the replication of read only data. Hence, the proposed strategy is used for OLAP purposes.

II. LITERATURE REVIEW

Rambabu D (2023) It is critical in cloud computing to have excellent data accessibility and system performance. To improve system availability, commonly used data should be duplicated to many places, allowing users to access it from a nearby site. Deciding on a sensible number and location for replicas is a difficult problem in cloud computing. Therefore, a novel Data Replication system based on data mining techniques is being proposed in this research work. The data replication is done here by locating commonly utilized data patterns in a node's massive database. This will be accomplished using an optimization-assisted frequent pattern mining approach, with a novel hybrid algorithm performing the best threshold selection.

Jiarui Wang (2021) The computing resources are supplied by cloud computing on basis of cloud user requirements demand. Using virtualization and distributed computing, the resource allocation model is constructed to highlight the cloud services scalability. Nevertheless, a complex problem is created by the user, to manage the demand in the on-demand resource allocation model. Hence, a novel optimization approach is developed called Grey Wolf Optimization and Crow Search Algorithm (GWO-CSA) to resolve the problem in the resource allocation model. In a distributed manner, to the virtual machine, the tasks are allocated, to balance workload in the cloud.

S Manimurugan (2020) Intrusion detection system has become the fundamental part for the network security and essential for network security because of the expansion of attacks which causes many issues. This is because of the broad development of internet and access to data systems around the world. For detecting the abnormalities present in the network or system, the intrusion detection system (IDS) is used. Because of the large volume of data, the network gets expanded with false alarm rate of intrusion and detection accuracy decreased. This is one of the significant issues when the network experiences unknown attacks. The principle objective was to expand the accuracy and reduce the false alarm rate (FAR). To address the above difficulties the proposed with Crow Search Optimization algorithm with Adaptive Neuro-Fuzzy Inference System (CSO-ANFIS) is used.

Adel N. Toosi (2017) Cloud Storage Providers (CSPs) offer geographically data stores providing several storage classes with different prices. An important problem facing by cloud users is how to exploit these storage classes to serve an application with a time-varying workload on its objects at minimum cost. This cost consists of residential cost (i.e., storage, Put and Get costs) and potential migration cost (i.e., network cost). To address this problem, we first propose the optimal offline algorithm that leverages dynamic and linear programming techniques with the assumption of available exact knowledge of workload on objects.

Log-Based Incremental Replication

Some databases allow you to store transaction logs for a variety of reasons, one of which is for easy recovery in case of a disaster. However, in log-based incremental replication, your replication tool can also look at these logs, identify changes to the data source, and then reproduce the changes in the replica data destination (e.g., database). These changes could be INSERT, UPDATE, or DELETE operations on the source database.

The benefits of this data replication strategy are:

- Because log-based incremental replication only captures row-based changes to the source and updates regularly (say, once every hour), there is low latency when replicating these changes in the destination database.
- There is also reduced load on the source because it streams only changes to the tables.
- Since the source consistently stores changes, we can trust that it doesn’t miss vital business transactions.
- With this data replication strategy, you can scale up without worrying about the additional cost of processing bulkier data queries.

Snapshot Replication

Snapshot replication is the most common data replication strategy; it’s also the simplest to use. Snapshot replication involves taking a snapshot of the source and replicating the data at the time of the snapshot in the replicas.

Because it’s only a snapshot of the source, it doesn’t track changes to the source database. This also affects deletes to the source. At the time of the snapshot, the deleted data is no longer in the source. So it captures the source as is, without the deleted record.

For snapshot replication, we need two agents:

- **Snapshot Agent:** It collects the files containing the database schema and objects, stores them, and records every sync with the distribution database on the Distribution Agent.
- **Distribution Agent:** It delivers the files to the destination databases.

Snapshot replication is commonly used to sync the source and destination databases for most data replication strategies. However, you may use it on its own, scheduling it according to your custom time.

Just like the full table data replication strategy, snapshot replication may require high processing power if the source has a considerably large dataset. But it is useful if:

- The data you want to replicate is small.
- The source database doesn’t update frequently.
- There are a lot of changes in a short period, such that transactional or merge replication wouldn’t be an efficient option.

- You don't mind having your replicas being out of sync with your source for a while.

Transactional Replication

In transactional replication, you first duplicate all existing data from the publisher (source) into the subscriber (replica). Subsequently, any changes to the publisher replicate in the subscriber almost immediately and in the same order.

To perform transactional replication, you need the Distribution Agent, Log Reader Agent, and Snapshot Agent.

- **Snapshot Agent:** It works the same as the Snapshot Agent for snapshot replication. It generates all relevant snapshot files.
- **Log Reader Agent:** It observes the publisher's transaction logs and duplicates the transactions in the distribution database.
- **Distribution Agent:** It copies the snapshot files and transaction logs from the distribution database to the subscribers.
- **Distribution database:** It aids the flow of files and transactions from the publisher to the subscribers. It stores the files and transactions until they're ready to move to the subscribers.

III. RESEARCH METHODOLOGY

Through internet as a service Cloud computing access resources and cloud storage as a powerful paradigm for sharing information across the internet. Each Cloud Storage providers has different pricing policies and terms for the end users. Several ultra scale datacenters are built by cloud providers at a variety of locations in an geographical areas which includes more number of computing servers. Cloud vendors need to upgrade their infrastructure according to the changes in the environment. The one area which is challenging is to manage is the data replicas so a separate storage space for replica is maintained. For increasing availability in storage data replication concept is used. Storage providers cost their users according to the usage space and bandwidth consumption. As far as using single cloud, it is considered as vendor lock-in risk because attackers can easily steal the information. So it was shifted to multi cloud which gives service availability and data can be easily managed. Moreover, the amount of data loss can be minimized. Group sharing manner is a complicated task because anywhere group admin and all members can store and modify the data while protecting data privacy from semi-trusted cloud service provider. A basic idea provided by existing system to keep respective user's data confidential against semi-trusted cloud server is encrypting the files, before uploading the data into the cloud server unfortunately a secure data sharing for dynamic groups in public cloud is not easy task because of some challenging issues. This framework combines proxy re-encryption, one time password verification, and regular elimination of unwanted files and prevents unauthorized access. Re-encryption improves the security and privacy of the uploaded files and then the one time– passwords improve the security compared to the text-based password.

IV. RESULTS

Proof of verification

If a signature (r, s) on a message m was indeed generated by A, then s = k -1 (h (m)+dr) mod n. Rearranging gives k ≡ s-1 (e+dr) ≡ s-1e + s-1 rd ≡we +wrd ≡ u1+u2d (mod n). Thus u1G +u2Q = (u1 +u2d) G = kG and so v=r as required.

Table 1: Execution time analysis for Fully Homomorphic-ECDSA

RSA-AES			ECIES			Fully Homomorphic-ECDSA		
Key length	Encryption	Decryption	Key length	Encryption	Decryption	Key length	Encryption	Decryption
	Time (ms)			Time (ms)			Time (ms)	
1024	1.81	1.81	160	7.41	7.36	160	6.98	6.75
2048	6.42	6.53	224	13.73	13.31	224	12.37	12.13
3072	18.88	19.01	256	16.96	17.27	256	15.69	16.72
7680	240.55	239.76	384	40.77	39.75	384	38.47	38.55
15360	1827.05	1834.13	512	81.52	82.84	512	80.25	79.94

Table 1 describes the strategy of comparison of an existing method, it processes the text file of 1kb data size, and the encryption time, decryption encryption time and different parameters encryption time compared to RSA-AES over time, different levels encrypted message is a function of the main size and data size for both RSA and ECIES. The ECDSA key size is relatively small compared to the RSA key size, thus the encrypted message in ECDSA is small.

Table 2: Encryption Time Analysis Comparison With Different File Sizes

File size	Encryption time (ms)					
	DES	3DES	AES	Blowfish	RSA	Fully Homomorphic-ECDSA
25KB	495	492	512	50	512	327
50KB	500	428	618	420	924	422
1MB	723	723	650	396	1432	406
2MB	943	1267	528	389	1587	432
3MB	1245	1648	567	894	1813	512

Table 2 compares the encryption time analysis for different file sizes with different existing algorithms. The time taken to convert plaintext to ciphertext is encryption time. Encryption time depends upon key size, plaintext block size, and mode. In our experiment, we have measured encryption time in milliseconds.

Table 3: Decryption Time Analysis Comparison With Different File Sizes

File size	Decryption time (ms)					
	DES	3DES	AES	Blowfish	RSA	Fully Homomorphic-ECDSA
25KB	398	346	337	21	298	297
50KB	387	387	599	97	576	365
1MB	598	545	573	302	806	398
2MB	787	765	438	278	827	405
3MB	874	799	436	545	1635	427

This proposed Fully homomorphic-ECDSA framework provides better encryption and decryption level security for the cloud data, but the encryption and decryption time gets increased with large key size (from Table 1). The lengthier key will provide superfluous security compared with smaller key sizes and the extra larger key size may cause more complex processes and increased processing time. Moreover, this framework limits to provide secure key communication and data backup services for effective data protection in cloud computing. To solve these issues, further, we implemented this framework in Aneka cloud with added secure key communication and data backup services over the key size of 512.

Table 4: Encryption time for the proposed framework

File	File Size(kb)	Encryption Time (ms)
File-1	21	293
File-2	46	302
File-3	71	312
File-4	84	321

From Table 4, it is known that the encryption time for a proposed framework for different file size File-1, File-2, File-3, and File-4 with the size of 21, 46, 71 and 84 Kb is 293,302,312 and 321 ms.

Table 5: Decryption time of proposed framework

File	File Size(kb)	Decryption Time(ms)
File-1	21	289
File-2	46	295
File-3	71	312
File-4	84	324

From Table 5, it is known that the decryption time for a proposed framework for different file sizes File-1, File-2, File-3 and File-4 with the size of 21, 46, 71 and 84 Kb is 289,295,312 and 324ms.

Table 6 Mode Analysis of Network

Distance (loss [in dB])	0	17	40	85
Effective measurement time [s]	245	230	210	340
Average single count rate [kcps]	50	28	23	7
Total sifted key [bits]	13230	15300	8400	9200
Sifted key rate [bit/s]	460	340	230	56
Error rate	0.012	0.043	0.056	0.052
Secure key rate [bit/s]	230	203	134	43
Run time (M cycles)	406	753	2200	1502
Average latency (ms)	603	1792	17452	700

Here, Table 6 describes the mode analysis of the network in our proposed work. For 0 dB distance loss, the obtained error rate is 0.012, sifted key rate is 460 bits/s, secure key rate is 230 bits/s, the total sifted key rate is 13230 bits, and so on. Thus, from this table, it is clear that the proposed framework acts as an effective barrier against various attacks and provides secured data storage and its access by handling the security threats in the cloud service provider.
Thus, it is clear from the above analysis that the quantum cryptography provides unconditional security and the sniffing detection properties to achieve secured data transmission and access control.

V. CONCLUSION

The division of fragments and replica management in multi cloud framework highlights the performance of dynamic replication for the management and the fragments that are allocated in multi cloud in which nodes are placed by using T-coloring concept. The cloud information Sharing plan is a promising security for consistent alteration of enrolment which contains a mix of assembling signature and element telecast re-encryption systems. This framework supports different clients for sharing basic data over the individuals and every part which comprises the information elements. This enables to provide highlights like security, protection valid access control, triviality and traceability. The replication cost was considered to solve the objectives of the problem and to be tested. Some future work can be made further by improving the Data availability in each data centres, with high replication factors that can have more reliability by combining to other fault tolerant techniques. In a cloud virtualized environment the proposed system shows with the experimental setup focusing on the Multi-Master Data replication strategy.

REFERENCES

1. Adel N. Toosi (2017),"Cost Optimization for Dynamic Replication and Migration of Data in Cloud Data Centers",IEEE Transactions on Cloud Computing,ISSNno:2168-7161,Vol.(99),Pages.1-1.DOI:10.1109/TCC.2017.2 659728
2. Rambabu D (2023),"Optimization assisted frequent pattern mining for data replication in cloud: Combining sealion and grey wolf algorithm",Advances in Engineering Software,ISSNno:1873-5339,Vol.176,https://doi.org/10.1016/j.ad vengsoft.2022.103401
3. Jiarui Wang (2021),"Grey Wolf Optimization and Crow Search Algorithm for Resource Allocation Scheme in Cloud Computing",Multimedia Research, ISSNno:2582-547X,Vol.4(3),
4. S Manimurugan (2020),"Intrusion detection in networks using crow search optimization algorithm with adaptive neuro-fuzzy inference system", Microprocessors and Microsystems,ISSNno:0141-9331,Vol.79,https://doi.org/ 10.1016/j.micpro.2020.103261
5. Alireza Askarzadeh (2016),"A novel metaheuristic method for solving constrained engineering optimization problems: Crow search algorithm", Computers & Structures,ISSNno:0045-7949,Vol.169,Pages.1-12.https://doi.or g/10.1016/j.compstruc.2016.03.001
6. Wei Guo;Sujuan Qin; (2020), “Comments on “Provable Multicopy Dynamic Data Possession in Cloud Computing Systems”, Volume: 15, ISSN: 1556-6013, PP: 2584 – 2586.
7. Mahantesh N. Birje, (2019), “Determination of Trustworthiness of Cloud Service Provider and Cloud Customer”, ISSN: 2575-7288, 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), PP: 15-16.
8. Jose M. Benítez, (2019), “Delivering Data Mining Services in Cloud Computing”, ISSN: 2642-939X, 2019 IEEE World Congress on Services (SERVICES). PP: 8-13.
9. Kalyani Patel, (2019), “Pricing model for revenue generation using Recurrent Neural Network for Cloud service provider”, 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), PP: 23-25.
10. Alyani Ismail (2018),"A New Intrusion Detection System Based on Fast Learning Network and Particle Swarm Optimization",IEEE Access, vol. 6, pp. 20255-20261,doi:10.1109/ACCESS.2018.2820092.

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| Mobile No: +91-9940572462 | Whatsapp: +91-9940572462 | ijarasem@gmail.com |

www.ijarasem.com