STRATEGY FOR IMPROVING CLOUD COMPUTING PERFORMANCE WITH ADAPTIVE DATA REPLICATION

Bairysetti Prasad Babu
Research Scholar
Dept of Computer science & Engineering
Shri JJT University-Rajasthan

Dr. Prasadu Peddi
Research Supervisor
Dept of Computer science & Engineering
Shri JJT University-Rajasthan.

Dr. Suneel Pappala
Co-Supervisor
Associate Professor
Dept of Computer science & Engineering
Lords Institute of Engineering And Technology, Hyderabad.

ABSTRACT

Cloud computing is becoming a very popular word in industry and is receiving a large amount of attention from the research community. Replica management is one of the most important issues in the cloud, which can offer fast data access time, high data availability and reliability. By keeping all replicas active, the replicas may enhance system task successful execution rate if the replicas and requests are reasonably distributed. However, appropriate replica placement in a large-scale, dynamically scalable and totally virtualized data centers is much more complicated. To provide cost-effective availability, minimize the response time of applications and make load balancing for cloud storage, a new replica placement is proposed. The replica placement is based on five important parameters: mean service time, failure probability, load variance, latency and storage usage. However, replication should be used wisely because the storage size of each site is limited. Thus, the site must keep only the important replicas. We also present a new replica replacement strategy based on the availability of the file, the last time the replica was requested, number of access, and size of replica. We evaluate our algorithm using the Cloud Sim simulator and find that it offers better performance in comparison with other algorithms in terms of mean response time, effective network usage, load balancing, replication frequency, and storage usage.

Keywords: *Replica management, Cloud Sim, Cloud computing, storage size, large-scale, cloud storage, replica replacement.*

INTRODUCTION

Info Technology professionals need installation software development, nonetheless it's quite hard to construct it into the very competitive fast changing small business atmosphere of today. From spread to CC Hence, the facets of information culture are altered. Cloud is also a pure development of computing also the overall variant of virtualization support Oriented Architecture (SOA). Cloud highlights on support-oriented architectures. Computing is kind of computing which enables person's capacity to tap. By decreasing computing is demonstrating an added incentive to both programmer's IT customers. Computing systems identifies exactly that the way, exactly in which whatever stems out of ability infrastructure, into application, and also the company enterprise procedures. It's where it's necessary, sent as support whatever. NIST defines CC for a tool for empowering omnipresent, handy, on - demand system usage of some shared pool of computing tools (e.g. networks, server's storage, software which may be quickly invisibly discharged along with nominal direction endeavor service or s provider discussion. This CC version is consists of 5 fundamental faculties (i.e., on-demand capacities, extensive system accessibility,

resource allocation (rapid durability, and also quantified products solutions), about a few services designs (i.e., IaaS, SaaS, along with also PaaS), along with four installation designs (i.e., people, personal, hybrid vehicle, network). The value of its particular own adoption along with also CC can be explained regarding shipping, its faculties, installation types. Characteristics the models lie on very peak of each other forming a pile of some cloud. Amazon identifies CC the shipping of IT tools since, along with software through the Web with cover while you go prices. By encouraging the functions of enterprise software, CC offers accessibility IT tools. Using CC calculating, an individual will not will need to develop big upfront investments and also invest a whole good deal of time of handling hardware to the functioning. CC is defined by Even the Gartner Group as a manner of by which scalable IT - capacities are given as an agency working with the Web systems to clients that are outside.

LITERATURE REVIEW

Ragha Madhavi Gondu (2023) The purpose of this study is to provide a better knowledge of the cloud computing as well as to suggest relevant research paths in this growing field. Also, we will go through the future benefits of cloud computing and the upcoming possible challenges we will have. Intext Cloud, performance, cloud computing, architecture, scale-up, and big data are all terms used in this context. Cloud computing offers a wide range of architectural configurations, including the number of processors, memory, and nodes. Cloud computing has already changed the way we store, process, and access data, and it is expected to continue to have a significant impact on the future of information technology. Cloud computing enables organizations to scale their IT resources up or down quickly and easily, without the need for costly hardware upgrades. This can help organizations to respond more quickly to changing business needs and market conditions.

Leelakumar Raja Lekkala (2023) A total of 102 publications were analyzed, revealing six main topics in cloud computing in healthcare: telemedicine/ teleconsultation, medical imaging, public health and patient self-management, hospital management and information systems, therapy, and secondary use of data. Common features utilized in these applications included broad network access for data sharing and access, as well as rapid elasticity to meet computing demands. While some articles highlighted the cost-effectiveness of pay-per-use cloud services, only 14 articles reported successful implementations, with many publications focusing on conceptual or prototypic projects. Additionally, several articles equated cloud computing with internet-/web-based data sharing, failing to illustrate the unique benefits of the cloud computing paradigm.

Gabriel Mafura (2023) Currently, Kenyan research institutions Information Technology operations use external storage, within or without institutional network environments. This study presents a structured literature review of a cloud computing adoption strategy for research institutions in Kenya. The reviews objectives are: 1) to establish the characteristics/peculiarities and IT environments of research institutions in Kenya, 2) current cloud computing technologies, adoption approaches and drivers in the adoption of cloud computing. The findings are that there is need for an adoption strategy for Cloud

computing in the Kenyan research Institutions that will use a DOI and TOE combined approach while addressing technological, organizational and environmental factors. The findings are used to propose a cloud computing adoption strategy for the Kenyan Research Institutions in future research.

Weihua Zhang (2022) With the emergence of massive data inside and outside the enterprise, paying much attention to the processing and analysis of accounting big data can bring huge value-added value to enterprises, and adapt to complex and changing economic environment. Based on this, by analyzing the related theories of accounting big data, accounting information and cloud computing, we build a cloud computing storage module of big data analysis platform, and apply Apriori data mining algorithm based on association rules to deal with massive data of accounting. A comparative prediction of the financial status of a group shows that the maximum error is less than 8% compared with the actual results, thus verifying the reliability and superiority of the established accounting big data analysis platform based on cloud computing.

M. Humayun Kabir (2021) Cloud computing is a type of emerging computing technology that relies on shared computing resources rather than having local servers or personal devices to handle applications. It is an emerging technology that provides services over the internet: Utilizing the online services of different software. Many works have been carried out and various security frameworks relating to the security issues of cloud computing have been proposed in numerous ways. But they do not propose a quantitative approach to analyze and evaluate privacy and security in cloud computing systems. In this research, we try to introduce top security concerns of cloud computing systems, analyze the threats and propose some countermeasures for them. We use a quantitative security risk assessment model to present a multilayer security framework for the solution of the security threats of cloud computing systems.

Cloud Computing Origin

The phrase "Cloud" is most widely utilized in mathematics to spell out an important group of items which visually resemble by a column including a cloud defines some of stuff whose information are somewhat more scrutinized in a specific circumstance. The explanation is the apps attracted the system surrounded the icons of servers using a circle, along with also clusters of servers at a system diagram experienced overlapping circles that found a cloud. In reference into this usage that is favored, the term blur has been practiced as being a metaphor for its world wide web, also there has been a conventional contour used to denote a system of schematics. After a period that it had been utilized to portray the Web on your computer system. The cloud emblem was utilized to define components of equipment by 1977 at Arpanet's beginning, the CSNET from predecessors into the Web.

Cloud Basics

CC also gives a broad variety of computing tools through system using "cover as you go" coverage version. Availing services that were different out of cloud companies can reduce now-a-days the price tag of IT. Computing systems consists of four installation types, several services types along with 5 faculties. The most critical faculties are multi-tenancy,

scalable service oriented, Service Level Agreement (SLA) pushed, virtualized, self along with copy. Infrastructure for a Service (IaaS), Software as a Service (SaaS) Platform as a Service (PaaS) would be the fundamental services offered by computing. Manifested Cloud Personal Cloud, hybrid cloud group Cloud are just four installation designs of CC.

Cloud Adoption Barriers

Computing has produced a move around in data creation base run along with can be conducted shifting invention areas business. Be as it could reversal, you'll find cynics fans. Exchanging opportunity IT is also an increasingly assignment which integrates the authoritative technical dilemmas. The cloud is still just another world view which does not always have a sentence definition; yet it within this way switch to your method might possibly look befuddling, also integrates aspects. This many-sided caliber together with vulnerability helps make many hierarchical cloud adoption hindrances. As signaled by means of a review directed by KPMG, protection, price tag vulnerability, lack of command would be the most important about a few cloud-appropriation borders.

RESEARCH METHODOLOGY

Even the programmer itself includes a separate group of nodes for processing data, or PE for short. The inquiry method presupposes that the uniform chip velocity corresponding processing operation may be attained by the analysis of very analogous groups of classes and items. In computer architecture, dynamic scaling may be used to regulate energy consumption. In addition to protecting against overload, dynamic SLA violence is used to store electrical strength, most notably in laptop computers and other portable devices, with the energy stems initially improving battery life. Different cloud storage services provide their customers with a range of different price structures. Hosting companies for the cloud construct large data centers spread on the globe, each section housing a number of computers used for computation. To adapt to these shifting conditions, distributors will need to invest in new and better infrastructure. One cloud service that gives businesses a choice is the cloud storage service. Your data components are the only thing that will be challenging to handle, hence a separate section has been reserved for their storage. It is possible to use replication theory. As a result, it has been upgraded to a general-purpose cloud that makes it simple to manage who has access to what data. Data privacy from a semi-trusted cloud provider has to be protected while the category admin and all team members may store and edit the information anywhere.

RESULTS

Data of different types with size ranging from 1 GB to 10 GB are taken for the evaluation purpose. At the system build-up stage, the initial replica factor of all data is set as three. Data placement is done using the HDFS default rack aware data placement policy. The weighted fuzzy inference system is embedded as a component inside the HDFS file system. Data stored in HDFS is accessed randomly and the HDFS log file is periodically checked and analyzed to retrieve the relevant information. Table 1 shows the data type and size taken for experiment.

Table 1: The Data Type and Size

Data Type	Data Size
Text and document files	5 MB – 3 GB

Image files	5 GB – 10 GB
Audio Files	5 GB – 25 GB
Video	10 Gb – 30 GB
Log file	10 GB – 50 GB

The HDFS log file is analyzed on particular time intervals to retrieve information about data access of each data item. Based on this information the popularity degree and weight of each data is calculated. These values are used as input to the weighted fuzzy inference system to categorize the data. Table 2 shows the resultant popularity degree, weight and category of data. The result clearly shows that the WFDRM method intelligently categorizes data to hot, warm or cold compared to other methods.

Table 2: Weight, Popularity Degree Calculation and Fuzzy Data Classification

Data	No .of Accesses	No. of nodes accessed	Current replica Factor	Weight	Popularity Degree	Fuzzy Category
D1	10	1	3	1	3.33	COLD
D2	10	2	3	1	6.66	COLD
D3	10	5	3	3	50	COLD
D4	10	9	3	4	120	WARM
D5	10	10	3	4	133.33	WARM
D6	50	1	3	1	16.66	COLD
D7	50	2	3	1	33.33	COLD
D8	50	5	3	3	250	WARM
D9	50	9	3	4	600	HOT
D10	50	10	3	4	666.66	HOT
D11	100	1	3	1	33.33	COLD
D12	100	2	3	1	66.66	WARM
D13	100	5	3	3	500	WARM
D14	100	9	3	4	1200	HOT
D15	100	10	3	4	1333.33	HOT
D16	500	1	3	1	166.66	WARM
D17	500	2	3	1	333.33	WARM
D18	500	5	3	3	2500	WARM
D19	500	9	3	4	6000	HOT
D20	500	10	3	4	6666.66	HOT
D21	1000	1	3	1	333.33	WARM
D22	1000	2	3	1	666.66	WARM
D23	1000	5	3	3	5000	WARM
D24	1000	9		4	12000	HOT
D25	1000	10		4	13333.33	HOT

The result of replica factor calculation for WFDRM method is given in Table 3.

Table 3: Replica Calculation of WFDRM

Data	No.of Accesses	No. of nodes accessed	Current replica Factor	Category	New Replica Factor

D1	10		3	COLD	1
D2	10		3	COLD	1
D3	10		3	COLD	1
D4	10		3	WARM	3
D5	10		3	WARM	3
D6	50		3	COLD	1
D7	50		3	COLD	1
D8	50		3	COLD	1
D9	50		3	WARM	3
D10	50		3	WARM	3
D11	100		3	COLD	1
D12	100		3	COLD	1
D13	100		3	COLD	1
D14	100		3	WARM	3
D15	100		3	WARM	3
D16	500		3	COLD	1
D17	500		3	COLD	1
D18	500		3	COLD	1
D19	500		3	WARM	3
D20	500		3	HOT	3
D21	10000		3	COLD	1
D22	10000		3	COLD	1
D23	10000		3	WARM	2
D24	10000		3	HOT	3
D25	10000		3	HOT	3

Evaluation Criteria

The evaluation of the WFDRM strategy is carried out by evaluating the parameters storage consumption, replication frequency and average response time. The storage space consumption is evaluated by monitoring the storage space consumed by each data item in the specified time intervals. The mean storage space consumption in the system is calculated using the equation and average response time is calculated.

CONCLUSION

Replication strategies have been widely adopted in current cloud systems for data availability, reliability, and performance. The adaptation improves system resilience during disasters without any downtime. The cloud replication strategy trend to

preserve the geographically distributed huge data, hence, creates the need for optimal replication strategy for acceptable performance. We filter out the dynamic replication strategies and evaluate their optimization capabilities based on quantitative analysis of target objectives (Primary target objective, Secondary target objective, and Tertiary target objective) using different attributes that are addressed. We provide a critical quantitative analysis and a comprehensive performance evaluation based on target objectives. We perform a comparative parameter evaluation, along with the metrics comparison. The paper also discusses the challenges, issues, and future research directions. This study will be beneficial to researchers to identify the research problems of replication strategies in cloud computing configuration and will provide a depth in detail related to available dynamic replication strategies and target-oriented replication strategies. This research will open a new gate to develop the optimal dynamic replication strategy for clouds in the future.

REFERENCE

1. *Ragha Madhavi Gondu (2023),"The Future of Cloud Computing: Benefits and Challenges",International Journal of Communications, Network and System Sciences, ISSNno:1913-3723,Vol.16,No.4,Pages.53-65.*
2. *Gabriel Mafura (2023),"Adoption Strategy for Cloud Computing in Research Institutions:A Structured Literature Review",Journal of Computer and Communi-cations,ISSNno:2327-5227,Vol.11,No.4,Pages.63-78.*
3. *M. Humayun Kabir (2021),"A Multilayer Security Framework for Cloud Computing in Internet of Things (IoT) Domain",Journal of Computer and Communications, ISSNno:2327-5227,Vol.9,No.7,Pages.31-42.*
4. *Weihua Zhang (2022),"Research on the Construction of Enterprise Accounting Data Analysis Platform Based on Cloud Computing",Open Journal of Business and Management,ISSNno:2329-3292,Vol.10,No.6, Pages.3132-3141.*
5. *Leelakumar Raja Lekkala (2023),"Integration of Mobile Computing and Cloud Computing in Healthcare",Voice of the Publisher,ISSNno:2380-7598,Vol.9,No.3, Pages.119-128.*
6. *Abou_el_ela Abdou Hussein (2021),"Data Migration Need, Strategy, Challenges, Methodology, Categories, Risks, Uses with Cloud Computing, and Improvements in Its Using with Cloud Using Suggested Proposed Model (DMig 1)",Journal of Information Security,ISSNno:2153-1242,Vol.12,No.1,Pages.79-103.*
7. *Jean-Louis Tetchueng (2022),"Trust Assessment Model Based on a Zero Trust Strategy in a Community Cloud Environment",Engineering,ISSNno:1947-394X, Vol.14,No.11,Pages.479-496.*
8. *Sara I. Al-Issa (2023),"An Enhanced Healthcare Integration System Based on Cloud Computing during the Coronavirus Crisis",Open Access Library Journal, ISSN no:2333-9721,Vol.10,No.2,Pages.1-14.*
9. *Charalampos Apostolopoulos (2019),"Tweet Sentiment Analysis (TSA) for Cloud Providers Using Classification Algorithms and Latent Semantic Analysis",Journal of Data Analysis and Information Processing,ISSNno:2327-7203,Vol.7,No.4, Pages.276-294.*
10. *W. Scott Wayne (2021),"Implementation of Radial Basis Function Artificial Neural Network into an Adaptive Equivalent Consumption Minimization Strategy for Optimized Control of a Hybrid Electric Vehicle",Journal of Transportation Technologies,ISSNno:2160-0481,Vol.11,No.4,Pages.471-503.*