iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 Issue: VII Month of publication: July 2019

DOI: http://doi.org/10.22214/ijraset.2019.7137

The Cost Effective Data Storage & Multi-Cloud Environment System

Ms. Sonali A. Waman[1], Dr. Mininath R. Bendre[2]
[1, 2]Department of Computer Engineering, Pravara Rural Engineering College, Loni , Maharashtra, 413736

Abstract: Cloud can be utilized to store large application data sets. The pay-as-you-go model is functional for cost effective storing of large volume of generated data sets in clouds along with improvement of storage strategies and bench marking approaches. The main problem concern with it is that most of generated data sets are either impractical or costlier for storage.
In the system, a practical storage strategy is proposed to achieve a minimum cost of storing along with Regenerating of Data sets in Multiple Clouds i.e. whether data sets could be stored or deleted along with where to store or to regenerate whenever they are re utilized. It also shows cost effectiveness to gain best trade-off among computation, storage along with bandwidth costs in multiple clouds.
The theoretical usefulness of the system is assured by inclusive analysis and rigid theorems. In the given proposed strategy, it can automatically choose if at run time or not the storing of generated data must be done or not. The primary objective of this strategy is Local-optimization for the trade-off between computation and storage.
The Secondary objective is associated with the user's preference on storage. In this system we can able to accomplish original along with generated data storage. Also we can utilized data compression for the effectual cost effective data storage in Cloud.
Keywords: Data sets, storage, computation, Minimum Cost, cloud computing.

I. INTRODUCTION

In the current era, scientific research progressively depends on IT technologies, in which large scale and high performance computing systems (e.g., clusters, grids, and supercomputers). All these systems are applied by the communities of researchers to carry out their applications. The Scientific applications are usually computation along with data intensive. The generated data sets are frequently in terabytes or even peta bytes in size.

In the near future the scientific data will be twofold and it will double every year as reported by Szalay and Gray. In the process of scientific data generation there is involvement of large number of computation intensive tasks, e.g. scientific work flows. Therefore, it is taking a long time for its execution. The important intermediate or final results of the computation present in the generated data sets required proper storage as they are vital resources.

The reason required for this are scientists may reused the existing data or they re-analyze the results of it.

The scientists from different institutions shared the data for collaboration. The different computation results are shared for it.

The regeneration cost can be save by storing vital generated data sets , if they are reused , not to mention the delay induced by regeneration.

The Prime challenge for the scientific data sets storage is the large size of them. The motivation of the proposed system is GT-CS algorithm. This algorithm is used to Calculate the minimum cost storage along with the regeneration strategy for the data application in multiple clouds.

The novel GT-CSB algorithm used to achieve the minimum cost for storing and regenerating datasets. It is also used to find the best trade-off among computation, storage and bandwidth costs in clouds.

Toward achieving the Minimum Cost of Storing and Regenerating Datasets, we propose a novel GT-CSB algorithm that can find the best trade-off among computation, storage and bandwidth costs in clouds. This trade-off is characterized by the theoretical minimum cost strategy for storing along with regenerating application data among multiple cloud service providers.

For Cloud users the minimum cost is very essential reference in the following three aspects. 1) It can be applied to design minimum cost benchmarking approaches for the evaluation of the cost effectiveness in the clouds; 2) It helps to develop cost effective storage strategies by guiding cloud users for their applications; and 3) it can demonstrate the constitution of several costs in clouds and support users to understand the impact of different workloads on the total cost.

II. LITERATURE SURVEY

Maximizing Revenue with Dynamic Cloud Pricing: The Infinite Horizon Case In this paper, we presented an infinite horizon revenue maximization framework to tackle the dynamic pricing problem in an infrastructure cloud. The prices are charged on a usage time basis is the main technical challenge when compared with preceding pricing work which result the demand departure process. The demand departure process has to be clearly modelled. For the infinite horizon case an average reward dynamic program is formulated. Its optimal conditions along with structural results on optimal pricing policies were presented [1]. Counter-Intuitive Throughput Behaviors in Networks Under End-to-End Control The traffic sources aggregate congestion measure in their paths by adapting their rates and they implicitly maximize certain utility. To study some counter-intuitive throughput behaviors in such network related to either as fair allocation is always inefficient or increasing capacity always raises aggregate throughput, we have used this paper.A class of utility functions parameterized by a scalar used to define a bandwidth allocation policy. This can be interpreted as a quantitative measure of fairness. An Allocation is fair only if it is large while it is efficient if aggregate throughput is large. A Fair allocation is necessarily inefficient which is suggested by all examples present in the literature. We have specify exactly the trade-off between fairness and throughput in general networks.[2]. An Efficient & Cost Effective Data Storage and Regeneration of Files in Multi-Cloud Environment 3. A Highly Practical Approach toward Achieving Minimum Data Sets Storage Cost in the Cloud 2013 The unique features and requirements of data sets storage in computation along with the data intensive applications in the cloud is investigated by using this paper. A novel runtime local optimization based storage strategy is developed by us. This strategy is used to practically achieve the minimum data sets storage cost in the cloud. While the strategy is depends on the enhanced linear CTT-SP algorithm utilized for the minimum cost benchmarking by taking the consideration of users (optional) preferences. [3].Dynamic Pricing for Reusable Resources in Competitive Market with Stochastic Demand 2018. The dynamic pricing optimization problem regarding the service providers selling reusuable products was studied by us along with three main contributions. To capture the dynamic and competitive features of the market we have proposed a comprehensive model. A Providers optimal pricing policies as an AE formulated by us and we have developed an algorithm to solve it.[4]. minimizing storage cost in cloud computing environment 2014. In the project, the unique features and requirements of data sets storage in computation along with data intensive utilization in the cloud has been investigated.

A novel runtime local optimization based storage strategy has been developed for minimization of storage cost in the cloud computing environment.The minimum cost benchmarking is achieved by using this strategy by taking into the consideration of user preferences. The strategy is very cost effective which is indicated by theoretical analysis, general random simulations and specific case studies. The cost effectiveness is done by achieving close to or even the same as the minimum cost benchmark with highly practical runtime efficiency[5].

III. PROPOSED SYSTEM

This minimum cost is a very vital reference for cloud users in the following three aspects:

A. It can be utilized for design minimum cost bench-marking approaches regarding evaluations for the cost effectiveness in clouds;
B. It can guide cloud users to develop cost effective storage strategies for their applications; and
C. It can demonstrate the constitution of different costs in clouds and help users to understand the impact of different workloads on the total cost. To achieve the minimum cost benchmark in a practical manner, we propose a novel local optimization-based runtime strategy for storing the generated application data sets in the cloud. We have used a Cost Transitive Tournament Shortest Path (CTTSP)- based algorithm. This was utilized for static on-demand minimum cost benchmarking of data sets storage in the Cloud. We increase the CTT-SP algorithm by incorporating users (optional) preferences on storage that can offer users some flexibility.

IV. SYSTEM ARCHITECTURE

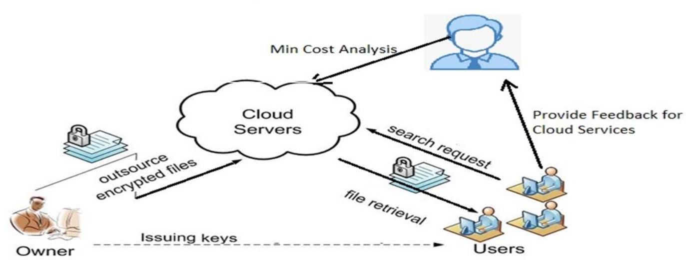


Fig. 1: System Architecture

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VII, July 2019- Available at www.ijraset.com

V. ALGORITHM

A. AES

1) KeyExpansions

a) For each round AES requires a separate 128-bit round key block plus one more.

2) InitialRound

a) AddRoundKeywith a block of the round key, each byte of the state is combined using bitwise xor.

3) Rounds

a) SubBytesin this step each byte is replaced with another byte.

b) ShiftRows for a certain number of steps, the last three rows of the state are shifted cyclically.

c) MixColumns a mixing operation which operates on the columns of the state, combining the four bytes in each column.

d) AddRoundKey

4) Final Round (no MixColumns)

a) SubBytes

b) ShiftRows

c) AddRoundKey.

B. Centrality

The centrality of a node in a graph provides the measure of the relative importance of a node in the network. The objective of improved retrieval time in replication makes the centrality measures more important. There are various centrality measures.

A. Closeness Centrality

A node is said to be closer with respect to all of the other nodes within a network, if the sum of the distances from all of the other nodes is lower than the sum of the distances of other candidate nodes from all of the other nodes. The lower the sum of distances from the other nodes, the more central is the node. Formally, the closeness centrality of a node v in a network is defined as: Where, N is total number of nodes in a network and d(v,a)represents the distance between node v and node a.

VI. RESULTS

To perform time analysis against the number of file size we have considered three strategies. Store all files, local optimization based strategy and proposed system. On Y axis we have taken time in seconds. And on X axis there is number of file count. Its clear that in proposed system it takes minimum time to store and regenerate the data set than other two strategies. And also there is very slight change in time even if file count get increased. So proposed scheme is efficient in terms of time. The figure 2 shows experimental analysis compared to existing system.

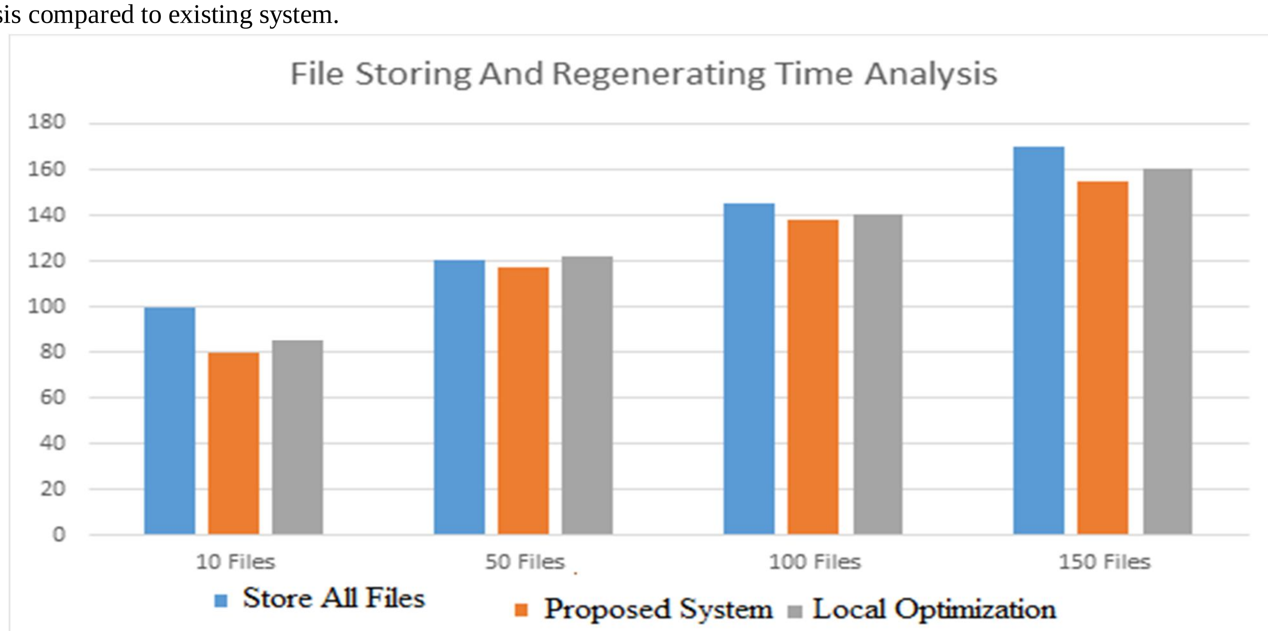


Figure 2. Experimental Analysis

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VII, July 2019- Available at www.ijraset.com

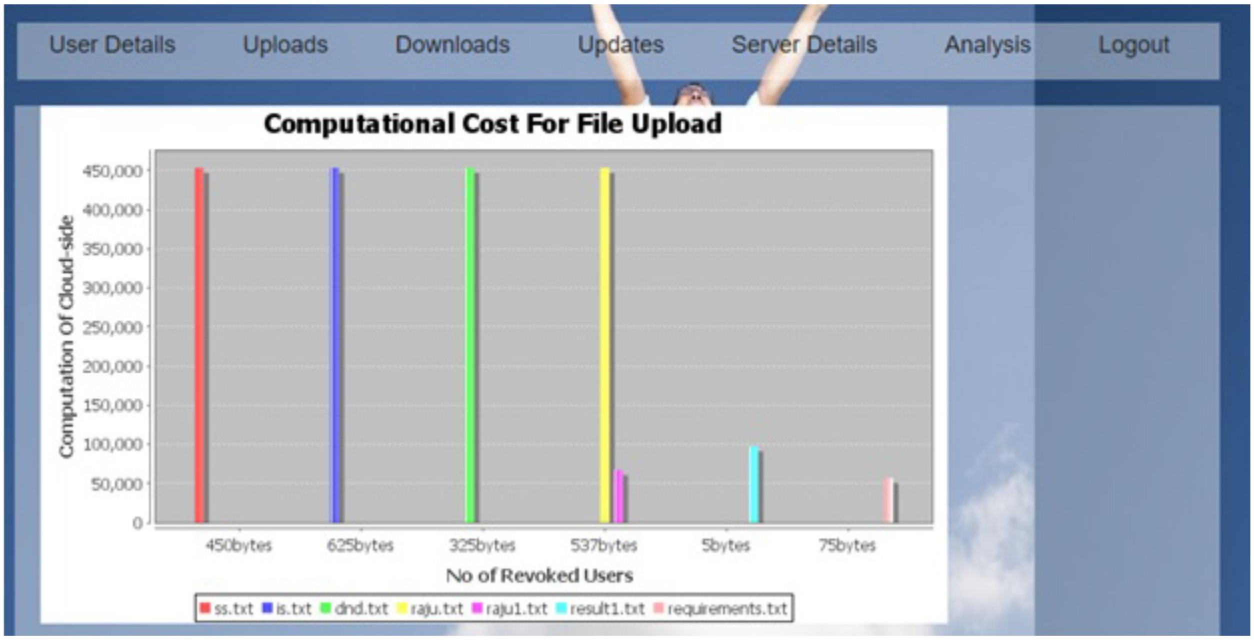


Figure 3. Project Output

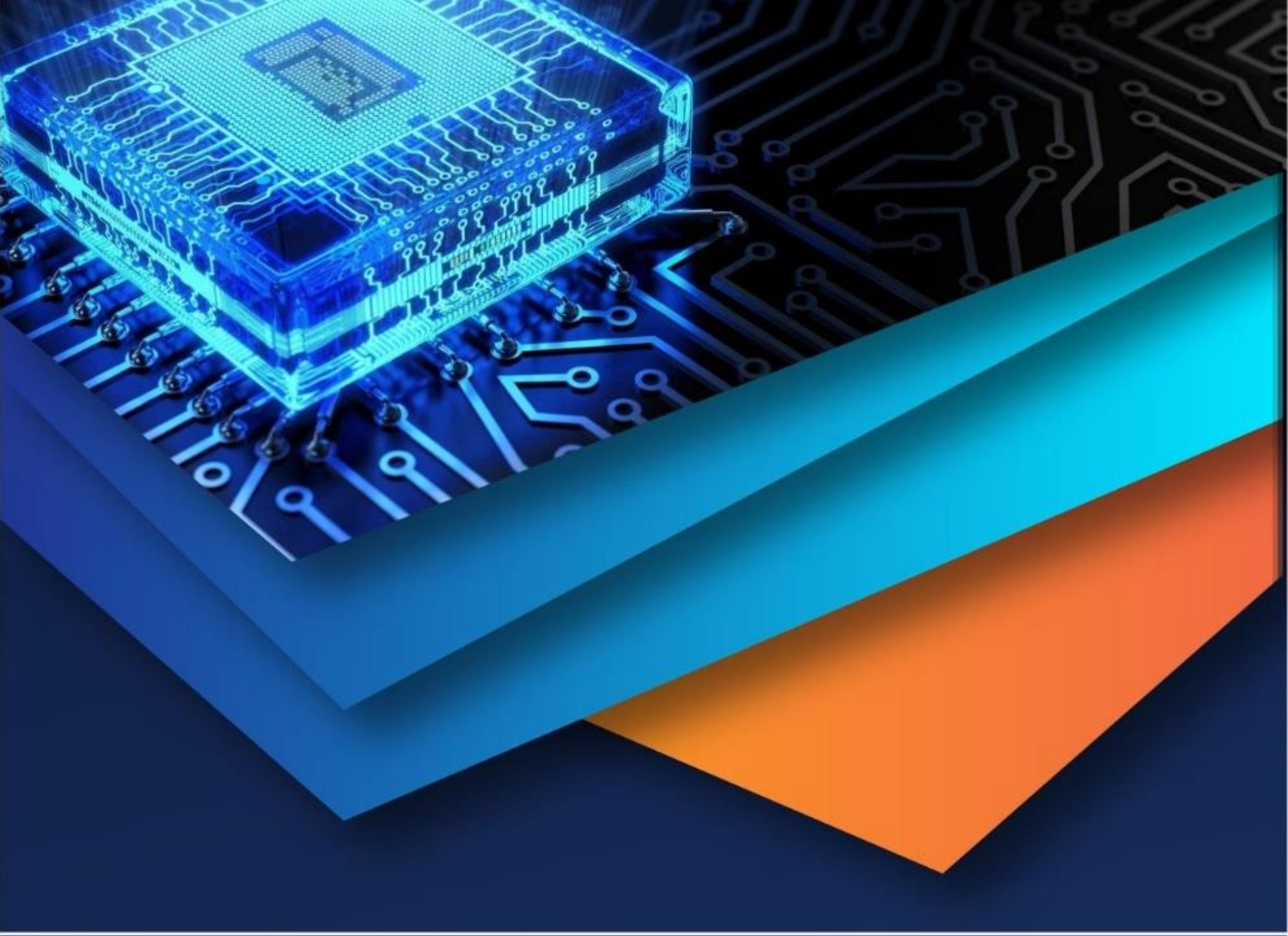
VII. CONCLUSION

The end of this decade is marked by a paradigm shift of the economic info technology towards a pay-per-use service business model called cloud computing. Cloud information Storage redefines the safety problems targeted on customer's outsourced information (data that's not stored/retrieved from the costumers own servers). In this work we have a tendency to ascertained that, from a customer's purpose of read, relying

upon a solo SP for his outsourced information isn't terribly promising, thus we have a tendency to square measure switch toward multi cloud. The cloud computing security is still a major issue in cloud computing environment in addition the loss of service availability and data integrity are the major problem for the customer.

REFERENCES

[1] Hong Xu, Baochun Li, "Maximizing Revenue with Dynamic Cloud Pricing: The Infinite Horizon Case, IEEE ICC 2012.[1]"
[2] Ao Tang, Student Member, "Counter-Intuitive Throughput Behaviors in Networks Under End-to-End Control, VOL. 14, NO. 2, APRIL 2006.[2]"
[3] A. Tang, J. Wang, and S. H. Low, Is fair allocation always inefficient, in Proc. IEEE INFOCOM, Mar. 2004.[3]"
[4] Dong Yuan, Yun Yang, "A Highly Practical Approach toward Achievin Minimum Data Sets Storage Cost in the Cloud, 2013.[4]"
[5] Jiang Rong, "Dynamic Pricing for Reusable Resources in Competitive Market with Stochastic Demand 2018.[5]"
[6] Sarika k b, "minimizing storage cost in cloud computing environment 2014.[6]"
[7] S. Kunniyur and R. Srikant, "End-to-end congestion control: utility functions, random losses and ECN marks, IEEE/ACM Trans. Netw., vol. 11, no. 5, pp. 689702, Oct. 2003.[7]"
[8] S. Agarwala, D. Jadav, and L.A., Adaptive Cost Optimization for Storage Clouds, IEEE Intl Conf. Cloud Computing, pp. 436-443, 2011.[8]"
[9] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, "A View of Cloud Computing, Comm. ACM, vol. 53, pp. 50-58, 2010[9]."
[10] R. Bose and J., "Lineage Retrieval for Scientific Data Processing: pp.m 1- 28, 2005[10]."



10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)