Achieving Minimum Transmitting and Caching Memory Cost On Data Staging Algorithm In Cloud

Prof.Sanchika Bajpai[1] ,Sujata Yadav[2] ,Reshma Ruke[3] ,Trupti Mahadik[4]

Department of Computer Engineering,BSIOTR ,Pune, India.

Abstract— **The strategies for efficiently achieving data staging and caching on a set of vantage sites in a cloud system with a minimum cost. Many time we require high bandwidth at that time cloud provider take more charge for uploads and downloads of customer data and this is the problem. To avoid this problem , we create a single or multiple copies of data item and that data item store on some site. When the ratio of transmission cost and caching cost is low, a single copy of every data item can powerfully serve to all the requested user , While in multi copy situation, we also consider the tradeoff between the transmission cost and caching cost by controlling the upper bounds of transmissions and copies this result is show under the homogeneous cost model.**

Index Terms— **Cloud computing, Data staging and caching, Resource constraints, Data Migration, Retainment, Excursion, Replication etc.**

I. INTRODUCTION

Now a days, data accessibility demand on cloud is increases because of that data availability maximization seems some important problem for accessing requested data . When user request for some data there is no guarantee to get data within minimum time. With increasing population of cloud user, the requested data available to the user is an important issue for Cloud server providers(CSP) to guarantee high quality services. To avoid this problem we store this requested data on some vantage site and cache the data from that site for some period of time. So quality of service can be greatly improved. We called this functionality as data staging.

Data staging problem by hold the dynamic programming (DP) techniques to replicate, migrate and cache the shared data items in cloud systems with or without some practical resource constraints in an well-organized way . Our first interest is memory cost which is a very flexible concept to reflect the qualities of various network features such as storage utilization, network bandwidth. The provision of the resources in cloud systems are usually based on pay-as-you-go fashion. Due to the optimality, our solutions are unique over other methods to provide the cloud-based services with the flexibility. They cannot decide the time of each data item to be cache data from some vantage sites.

II. RELATED WORK

Data staging is to facilitate a sequence of time variant accesses to cloud services necessarily builds upon work in a number of related areas. The requests made by users are to known in terms of their frequencies and positions in the network, while the least number and location of the servers to provide the requests with least total read and write costs are to be determined. Previous results can be viewed as an extension to this problem but it is not only considered the frequency of a particular request but also dealt with its time sequence. The cost of handling a sequence of requests is the same to the total distance (transmission cost) moved by the servers is the optimization goal.

The previous problem bears some similarity to the classic file allocation (FA) problem where multiple copies of a file are maintained by using caching, migrating, and replicating at the nodes of a network to minimize communication costs for read/write requests [4]. File copies could be created or deleted at with zero cost and file caching cost is also free. As a consequence there are only a transmission cost defined for file replication and write cost for file creation or update which is typically a nonlinear function of the number of copies present..

III. IMPLEMENTATION

With the heterogeneous cost model where each node has its own caching cost rate and transmission cost rate between any pair of nodes is not always identical. As this general case is overly complex However, by extending the DP algorithm in where Sk > Cij; I, j, k =1,2,….m is considered for optimal data staging solution, we can show that this heterogeneous case is still tractable under some restricted conditions. To satisfy a request for a particular data item we define the following five primitive operations to preform on the cached data item. This involve caching and transmission costs:

1. Retainment : In this operation ,cache the data item at a node pu from time t_u to t_v by paying $(t_v - t_u) S_u$, S_u is the rate of caching cost at node p_u, 1 <= u<=m.
2. Migration: In this operation ,move the data item from a node p_u to a node p_v at a cost equal to the distance C_{uv}.
3. Replication: In this operation , copy the item to the request node p_v from a node p_u at a cost of C_{uv}.
4. Excursion: In this operation , satisfy the request at a node p_v by using the copy at a node p_u without migration at a cost of E_{uv}.
5. Creation/Deletion: In this operation , create/delete the selected copies at some nodes without incurring any cost.

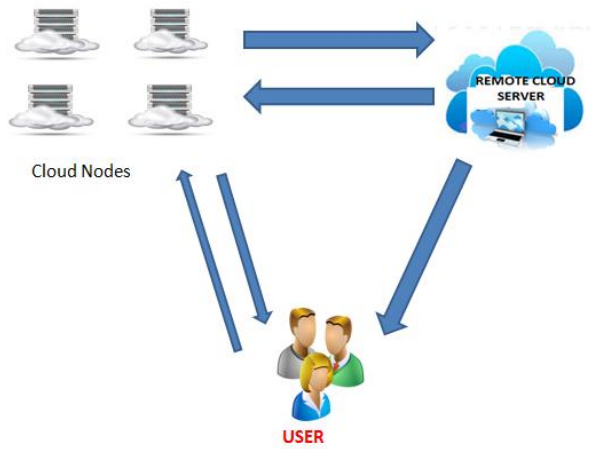

Fig.1 System Architecture

IV. A GREEDY ALGORITHM

The time complexity of the DP-based algorithm is prohibitively high even though we have imposed some constraints on the problem. In this , we suggest a simple greedy algorithm to the problem in its general form, whose staging cost is at most two times the optimum.

The basic idea is very easy. Given $= \sigma 1, \sigma 2 \ldots . \sigma n$, we first construct a service graph G= (V , E) in the following way:

1. V={1,2,. . . n} represents the n requests making at (pi, ti), i $i \epsilon [1, \ldots \ldots, n]$;
2. for each request i, compute its source point set E^i;
3. for each point $l \in E^i$, add an arc from *l* to i in G with the weight being equal to the cost of the shortest path from *l* to i.

Fig. 2(b) shows an example of the service graph which is constructed based on the request sequence in (a) [2]. Then, we compute a minimum spanning branchings (MSB) of the service graph to satisfy all request demands. Since $|E^i| <= m$, the total number of arcs is not greater than mn, and thus the time complexity of constructing the service graph is O(mn).Therefore, its MSB can be found either in O(mn log n) for as parse graph or in O(n^2) for a dense graph by Chu & Liu/ Edmonds' algorithm.

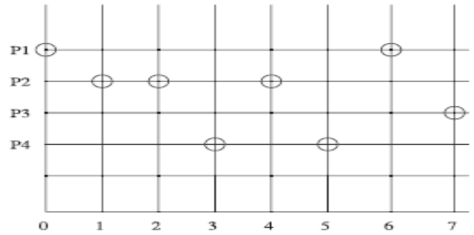

Fig. 2(a) Space time diagram of request Squence

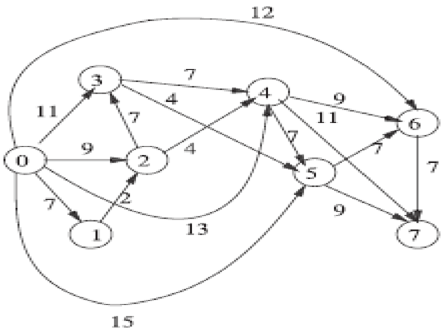

Fig. 2(b) Service graph for request sequence

V.EXPERIMENTAL RESULT

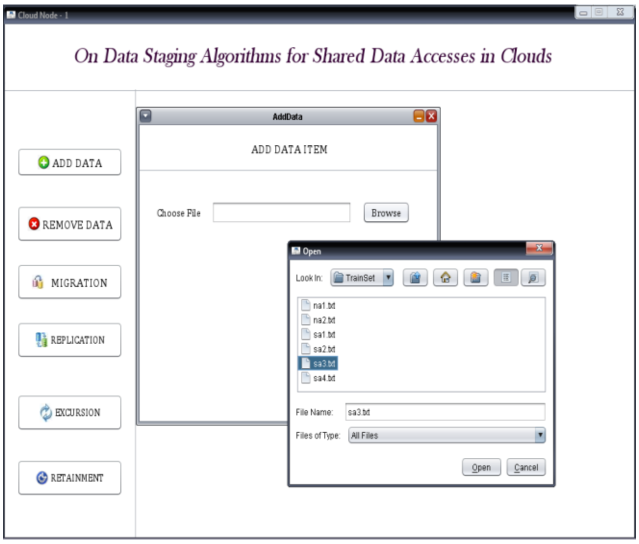

Fig.3 Add data into node

Above fig. shows that how data are added into the node. After adding this data into node this information is send to server and all other nodes.

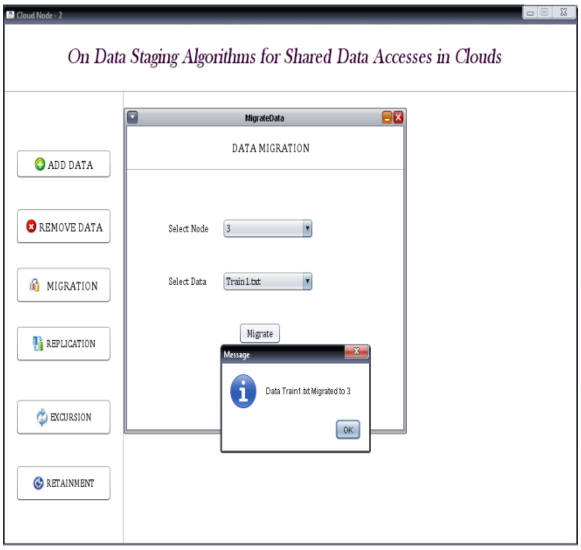

Fig.4 Migration of data

Above fig shows that we move the data from one node to another. Data migration is used to move the data onto the nearest node to minimize the cost and access time.

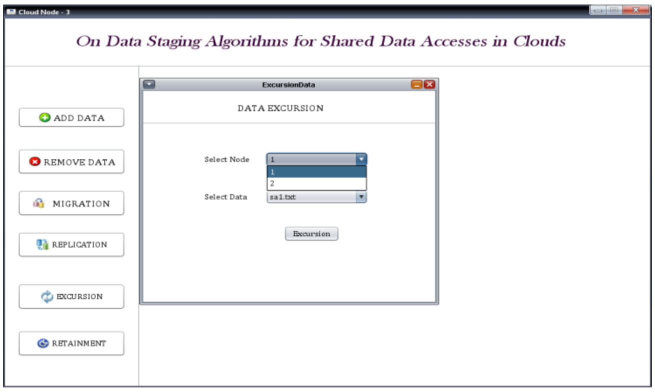

Fig. 5 Excursion of data

Above fig. shows to satisfy the request at a node pv by using the copy at a node pu without migration at a cost of Euv.

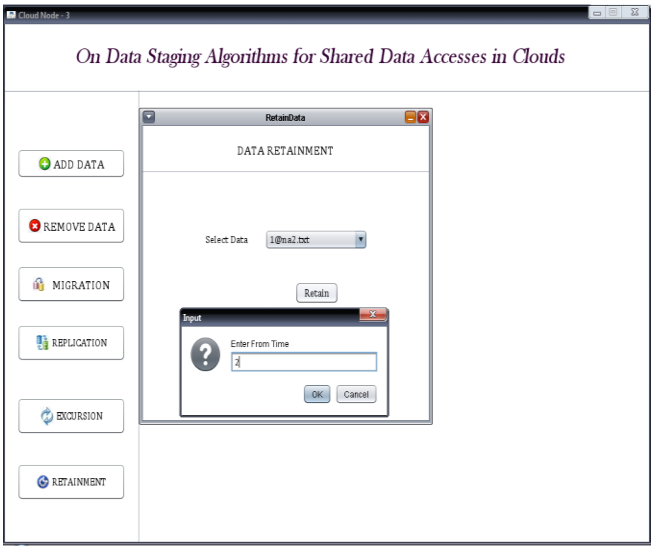

Fig.6 Retainment of data

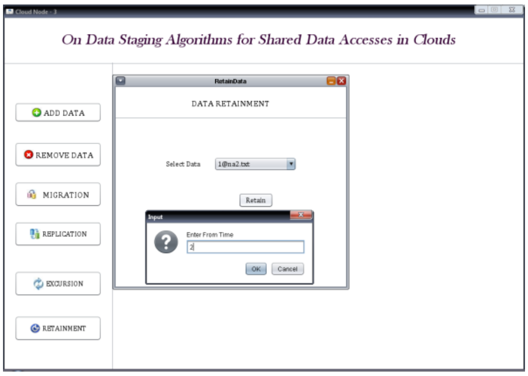

Fig.6 Retainment of data

Above fig 6(a) and (b) shows that we can use data of one node on another node for particular time

The Experiments conducted on simulated environment having Minimum 512 MB RAM With the Pentium IV Processor and above. working on windows 7 used for developing environment and java used as coding language.

VI. CONCLUSION

we conclude that, data staging is a set of distinct data items problem in a complete connected network to provide cloud-based services with monetary cost. In this we focused on the efficient staging strategies based on the cost models to reduce the total staging cost. We can also reduce the accessing time of requested data. So the required memory space also reduced.

REFERENCES

[1] Kalaiselvi.S1; Newlin Rajkumar.M2; ― Achieving Minimum Transferring And Storage Cost On Data Staging Algorithms For Shared Data Accesses In Clouds‖, PG Scholar, Regional Centre of Anna University, Coimbatore, India1

[2] Yang Wang,‖On Data Staging Algorithms for SharedData Accesses in Clouds‖,Bharadwaj Veeravalli, Senior Member, IEEE, and Chen-Khong Tham

[3] Bolan B. Veeravalli, ―Network Caching Strategies for a Shared Data Distribution for a Predefined Service Demand Sequence,‖ IEEETrans. Knowledge and Data Eng., vol. 15, no. 6, pp. 1487-1497, Nov.2003.

[4] V.U.D.P. Vardhani "Optimal Solution For Data Staging In Cloud Computing Shared Data Access" International Journal Of Professional Engineering Studies, Volume IV/Issue2/OCT2014

