

A Novel Method of Data Staging For Better Mobile User Accessibility through Cloud

Syeda Misba, M.Tech II year, Prof D. Jayaramaiah , HOD, ISE

The Oxford College of Engineering, Bangalore.

The Oxford College of Engineering, Bangalore.

Abstract: *In this paper, we identify and create vantage sites to stage shared data items for the ease of mobile user accesses. As mobile users will be charged on the usage of communication bandwidth and for the amount of the time the storage resources are utilized on their respective data, we aim to minimize cost for mobile users. As the identification of the access patterns to assist future requests are already known in advance, our aim is to efficiently stage the shared data items to predetermined vantage sites at specific time intervals thus minimizing the monetary costs for caching and transmitting the requested data items as well as managing some of the issues faced by Mobile cloud computing like limited bandwidth and high latency of network because the radio resource for wireless networks is much scarce as compared with the traditional wired networks and availability as mobile users may not be able to connect to the cloud to obtain a service due to traffic congestion, network failures, and the out-of-signal.*

Index terms: *Cloud computing, Mobile cloud computing, data staging, data caching, data placement and replication.*

I. Introduction

Mobile devices (e.g., smartphone and tablet PC) are increasingly becoming an essential part of human life as the most effective and convenient communication tools not bounded by time and place. The rapid progress of mobile cloud computing (MCC) becomes a powerful trend in the development of IT technology as well as commerce and industry fields. However, the mobile devices are facing many challenges in their resources (e.g., battery life, storage, and bandwidth) and communications. The limited resources significantly impede the improvement of service qualities.

Cloud computing (CC) has been widely recognized as the next generation computing infrastructure [1]. CC offers some advantages by allowing users to use infrastructure (e.g., servers, networks, and storages), platforms (e.g., middleware services and operating systems), and software's (e.g., application programs) provided by cloud providers (e.g., Google, Amazon, and Salesforce) at low cost. In addition, CC enables users to elastically utilize resources as an integration of CC into the mobile environment. MCC [2] brings new types of services and facilities mobile users to take full advantages of CC.

The mobile users can obtain some advanced services which are traditionally not widely available on the Internet due to the variety of query or request patterns, the size of requested data set and as well as the limited user budget. For example, a user-specific service may involve a large amount of data staging and efficient resource allocating to facilitate the data accessibility with minimum costs.

In general, different cloud-based mobile services may be initiated by various types of users, and their query patterns (e.g., time and location) could be known from the historical data and thus it is determined in advance to support the future access. Given this context, in this paper, we consider the problem of efficiently identifying vantage locations to stage a requested data item so that the data accessibility is improved to the mobile users with minimum costs. As well as to handle issues faced by MCC like network delay, limited bandwidth and availability.

Our strategy provides the Cloud Service Providers (CSPs) flexibility with which they can not only decide the time, location, and the duration of the data item to be cached[3] for the optimization but also hit a balance between communication cost and caching cost to meet the constraints imposed by the technical requirements or the user's budget.

II. Existing System

In existing system vantage sites (datacenters) are identified and created. In each of these vantage sites data items are staged which in turn has helped in reducing caching and communication cost. The focus was only on modeling the generation of the request workload and its access patterns. Some of the properties and features of network platform such as network traffic, bandwidth capacity, link latency are not modeled. And also problem under storage constraints are not explored.

III. Proposed System

The existing system is extended to mobile clients, such that they can access the information from anywhere and at any time. If the mobile users have to access data directly from the cloud it is very expensive, time consuming, requires more bandwidth and also increases the network traffic.

So, in order to overcome these problems and make it easy for mobile users in accessing data from cloud several vantage sites (datacenters) [6], [7] are created and data items are staged[8] in them.

In this project rather than replicating the cloud we are replicating, migrating[5] and storing only part of the data in the created vantage location for stipulated time based on the users demand and application. Synchronous replication technique is used for replicating data. Replica placement optimization problem[4] is handled using Closest Access Policy in which each user is allowed to be served only by the closest replica in the path from itself up to the source root.

Suppose there are k distinct shared data items initially stored at one node (vantage site), say p_1 and later replicated, and cached in a fully connected network with m nodes (p_1, p_2, \dots, p_m) to serve a sequence of requests $r = r_1, r_2, \dots, r_n$ in which each $r_i = (t_i, p_i, R_i)$, $1 \leq i \leq n$ which represents a request made for a data subset R_i by a node p_i at time t_i

In order to satisfy a request for a particular data item, the following primitive operations are performed on the cached data items, which may involve caching and transmission costs:

1. Retainment: cache the data item at a node p_u 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 , $1 \leq u \leq m$
2. Replication: copy the item to the request node p_v from a node p_u at a cost of C_{uv}
3. Excursion: 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}
4. Creation/Deletion: create/delete the selected copies at some nodes without incurring any cost

Each new request can trigger one or more operations, leading to the creation and deletion of the data items at arbitrary nodes. This procedure can be visually described by a space-time diagram in which the horizontal axis represents the time stages while the vertical axis represents the nodes, the requests are shown as circles, also referred as request points.

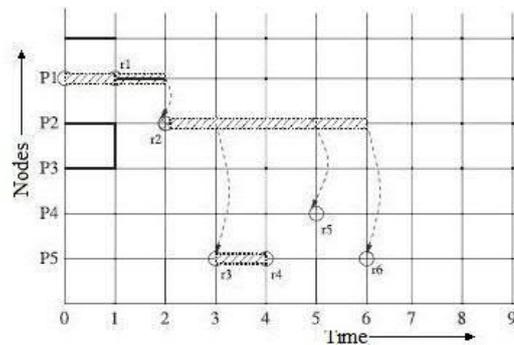


Fig. 1: Space-time diagram representation and a feasible caching and transmission schedule for a data item, which is initially located at p_1

In Fig. 1 the transmission of the shared data from one node to another at a stage is represented by a vertical directed arrow connecting two points on a stage between the respective nodes and the caching of the shared data at a node from one stage to another is represented by hashing the horizontal time line between corresponding stages at that node. The goal is to produce a data staging scheme by combining the primitive operations to satisfy the request sequence in a n efficient way while minimizing the total cost. The total cost for the staging(S) and communication(C) scheme shown in Fig. 1 is $2S_1 + 4S_2 + S_5 + C_{12} + 2C_{25} + C_{24}$. Since users are accessing information through mobile phones which are of different networks like 2G and 3G, comparison among them is done by considering some of the performance metrics like cost, latency, access time and availability.

A. System Parameters

- Specification of time to retain the data contents
- Identifying the type of data contents that has to be replicated
- Specifying the number of copies that has to be replicated
- Migration

B. System Architecture

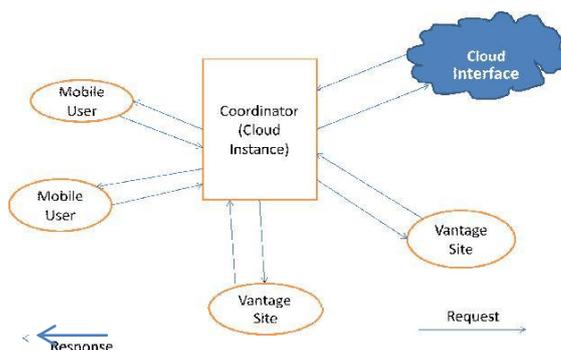


Fig.2: General Architecture

As shown in fig.2, the coordinator (Cloud Instance) is responsible for exporting the cloud services to mobile users. For every request that it receives from user it schedules the task for the nearest vantage site (datacenter) by using location based algorithm which calculates the distance and then finds the shortest path between the source of the request and the datacenter. After retrieving the requested data from the specified vantage site the cloud instance returns the data back to the requesting source. If the requested data is not available in the vantage site then the cloud instance sends the user request directly to the cloud.

IV. Study and Analysis

To verify and study the performance behavior of our concept, our proposed system is developed using NetBeans IDE 7.2 integrated with cloudsim-3.0 cloud simulator. It is implemented then in three phases. In first phase a network is created by establishing connection between nodes. Nodes include cloud instance, cloud interface, vantage sites and users. The data is then uploaded onto cloud interface. Second phase includes replicating some part of data to vantage sites and task scheduling for every user generated request by applying closest access policy based on the location of user and vantage site. Third phase deals with evaluation of performance metrics.

V. Results and Future work

- Efficiently stage the shared data items to predetermined sites
- Minimizing cost for caching and communication
- Minimizing latency
- Reducing network traffic
- Improving bandwidth efficiency

Providing comparison of performance metrics such as cost, latency, access time, and availability among different mobile networks while directly accessing cloud and when accessed to a vantage site.

VI. Conclusion

In this paper, we have proposed an adaptive way of staging a shared data item to vantage locations in a Cloud platform for the ease of mobile access with minimum communication and caching costs and also to overcome challenges of MCC such as latency, access time, data availability and bandwidth. In this paper, we have assumed that the accesses to the k distinct documents are independent with each other and the cost is defined on per-item basis. In this paper we have also overcome the problem that is faced under the storage constraints. As each node (vantage site) is equipped with fixed storage size, wastage of resources such as cache memory, processing time is avoided by replicating only those items that are more frequently been accessed by users from the cloud.

References

- [1] "Cloud Computing for Future Generation of Computing Technology" Proceedings of the 2012 IEEE International Conference on

- Cyber Technology in Automation, Control and Intelligent Systems May 27-31, 2012
- [2] "MOBILE CLOUD COMPUTING AS FUTURE FOR MOBILE APPLICATIONS – IMPLEMENTATION METHODS AND CHALLENGING ISSUES" Shahryar Shafique Qureshi , Toufееq Ahmad, Khalid Rafique2, Shuja-ul-islam, IEEE CCIS,2011
 - [3] H. Gupta and B. Tang, "Data Caching under Number Constraint," Proc. IEEE INFOCOM, 2006.
 - [4] A. Benoit, V. Rehn-Sonigo, and Y. Robert, "Replica Placement and Access Policies in Tree Network," IEEE Trans. Parallel and Distributed Systems, vol. 19, no. 12, pp. 1614-1627, Dec. 2008.
 - [5] M. Bienkowski, A. Feldmann, D. Jurca, W. Kellerer, G. Schaffrath, S. Schmid, and J. Widmer, "Competitive Analysis for Service Migration in Vnets," Proc. Second ACM SIGCOMM Workshop Virtualized Infrastructure System and Architectures (VISA), Sept. 2010.
 - [6] M. Al-Fares, A. Loukissas, and A. Vahdat, "A scalable, commodity data center network architecture," in *ACM SIGCOMM*, Aug. 2008.
 - [7] D. Li, C. Guo, H. Wu, K. Tan, Y. Zhang, S. Lu, and J. Wu, "Scalable and cost-effective interconnection of data-center servers using dual server ports," *IEEE/ACM Transaction on Networking*, vol. 19, no. 1, pp. 102–114, 2011.
 - [8] "Just-in-time Staging of Large Input Data for Supercomputing Jobs" Henry M. Monti and Ali R. Butt Virginia Tech Blacksburg, Sudharshan S. Vazhkudai Oak Ridge National Laboratory Oak Ridge, ©2008 IEEE