ISSN 2278-3091

**International Journal of Advanced Trends in Computer Science and Engineering**, Vol.2 , No.1, Pages : 259 – 263 (2013)
*Special Issue of ICACSE 2013 - Held on 7-8 January, 2013 in Lords Institute of Engineering and Technology, Hyderabad*

# Parameter Estimation of Price-Demand Model for Cloud Data Services

**Rambabu (M.Tech)**
11E21D5808
Vidya Vikas Institute of Technology
ramu.nandru89@yahoo.com

**Sai Krishna (M.Tech)**
11E21D5810
Vidya Vikas Institute of Technology
sai.krishnaΔappu@yahoo.com

**L.sunitha M.Tech,(Ph.D)**
Associate Professor
Vidya Vikas Institute of Technology
sunithaΔlingam@yahoo.com

**Abstract**—Cloud applications that offer data management services are emerging. Such clouds support caching of data in order to provide quality query services. The users can query the cloud data, paying the price for the infrastructure they use. Cloud management necessitates an economy that manages the service of multiple users in an efficient, but also, resource-economic way that allows for cloud profit. Naturally, the maximization of cloud profit given some guarantees for user satisfaction presumes an appropriate price demand model that enables optimal pricing of query services. The model should be plausible in that it reflects the correlation of cache structures involved in the queries. Optimal pricing is achieved based on a dynamic pricing scheme that adapts to time changes. This paper proposes a novel price-demand model designed for a cloud cache and a dynamic pricing scheme for queries executed in the cloud cache. The pricing solution employs a novel method that estimates the correlations of the cache services in a time-efficient manner. The experimental study shows the efficiency of the solution.

**Keywords:** Cloud data management, data services, cloud service pricing.

## 1. INTRODUCTION

The leading trend for service infrastructures in the IT domain is called cloud computing, a style of computing that allows users to access information services. Cloud providers trade their services on cloud resources for money. The quality of services that the users receive depends on the utilization of the resources. The operation cost of used resources is amortized through user payments. Cloud resources can be anything, from infrastructure (CPU, memory, bandwidth, network), to platforms and applications deployed on the infrastructure. Cloud management necessitates an economy, and, therefore, incorporation of economic concepts in the provision of cloud services. The goal of cloud economy is to optimize: 1) user satisfaction and 2) cloud profit. While the success of the cloud service depends on the optimization of both objectives, businesses typically prioritize profit. To maximize cloud profit we need a pricing scheme that guarantees user satisfaction while adapting to demand changes. Recently, cloud computing has found its way into the provision of web services Information, as well as software is permanently stored in Internet servers and probably cached temporarily on the user side. Current businesses on cloud computing such as Amazon Web Services [1] and Microsoft Azure [2] have begun to offer data management services: the cloud enables the users to manage the data of back-end databases in a transparent manner. Applications that collect and query massive data, like those supported by CERN [3], need a caching service, which can be provided by the cloud [4].The goal of such a cloud is to provide efficient querying on the back-end data at a low cost, while being economically viable, and furthermore, profitable. Fig. 1 depicts the architecture of a cloud cache. Users pose queries to the cloud through a coordinator module, and are charged on in order to be served. The cloud caches data and builds data structures in order to accelerate query execution. Service of queries is performed by executing them either in the cloud cache (if

ISSN 2278-3091

**International Journal of Advanced Trends in Computer Science and Engineering**, Vol.2 , No.1, Pages : 259 – 263 (2013)
*Special Issue of ICACSE 2013 - Held on 7-8 January, 2013 in Lords Institute of Engineering and Technology, Hyderabad*

necessary data are already cached) or in a back-end database. Each cache structure (data or data structures) has an operating (i.e., a building and a maintenance) cost. A price over the operating cost for each structure can ensure profit for the cloud. In this work, we propose a novel scheme that achieves optimal pricing for the services of a cloud cache.

### 1.1 Setting the Price for Cloud Caching Services

The cloud makes profit from selling its services at a price that is higher than the actual cost. Setting the right price for a service is a nontrivial problem, because when there is competition the demand for services grows inversely but not proportionally to the price. There are two major challenges when trying to define an optimal pricing scheme for the cloud caching service. The first is to define a simplified enough model of the price demand dependency, to achieve a feasible pricing solution, but not oversimplified model that is not representative. For example, a static pricing scheme cannot be optimal if the demand for services has deterministic seasonal fluctuations. The second challenge is to define a pricing scheme that is adaptable to 1) modeling errors, 2) time-dependent model changes, and 3) stochastic behavior of the application. The demand for services, for instance, may depend in a non predictable way on factors that are external to the cloud application, such as socioeconomic situations. A representative model for the cloud cache should take into account that the cache structures (table columns or indexes) may compete or collaborate during query execution. The demand for a structure depends not only on its price, but also on the price of other structures. For example, consider the query select A from T where B = 5 and C = 10. Out of the set of candidate indexes to run the query efficiently, indexes Ib = T(B), Ic = T(C), and Ibc = T(BC) are most important, since they can satisfy the conditions in the "where" clause. If the cache uses Ibc, then the indexes Ib and Ic, will never be used, since Ibc can satisfy both Conditions. Therefore, the presence of Ibc has a negative impact on the demand for Ib and Ic. Alternatively, if the ache uses Ib, then Ic can also improve query performance via index intersections, hence increasing the profit for the loud. Therefore, indexes Ib and Ic have positive impact on each other's demand. An appropriate estimation method is necessary to model price-demand correlations among cached structures. That the selling good is not a consumable product, but a persistent service. A consumable product diminishes with demand and has to be ordered, whereas a cloud cache service can satisfy infinite demand as long as it is maintained. Moreover, the demand for a cache service pauses if this service is not available. A consumable product may cost to maintain depending on the stored amount, whereas the maintenance cost of a cache service depends only on time. A big challenge for the cloud is to optimize the set of offered services, i.e., decide which services to offer and when, depending on their demand while they are available. Roughly, the cloud has to schedule online and offline periods of the offered services, which affects the maintenance and the setup cost. Furthermore, the optimization of the cloud profit has to be scheduled for a long period in time while it is flexible during this period to adjust to the real evolution of the service consumption. The long-term profit optimization is necessary in order for the cloud to schedule ahead associative actions for the maintenance of the cloud infrastructure and the cloud data. Moreover, the cloud can schedule the service availability according to the guarantees for the overall revenue estimated by the long term optimization.. The corrections may refer to the difference between the estimated and the actual price influence on the demand of services

ISSN 2278-3091

**International Journal of Advanced Trends in Computer Science and Engineering**, Vol.2 , No.1, Pages : 259 – 263 (2013)
*Special Issue of ICACSE 2013 - Held on 7-8 January, 2013 in Lords Institute of Engineering and Technology, Hyderabad*

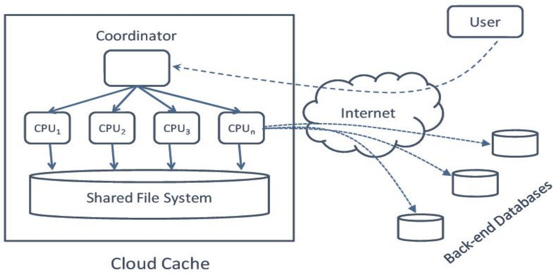


Fig1.A cloud cache

### 1.2 Present Proposal:

The cloud caching service can maximize its profit using an optimal pricing scheme. This work proposes a pricing scheme along the insight that it is sufficient to use a simplified price-demand model which can be reevaluated in order to adapt to model mismatches, external disturbances and errors, employing feedback from the real system behavior and performing refinement of the optimization procedure. Overall, optimal pricing necessitates an appropriately simplified price-demand model that incorporates the correlations of structures in the cache services. The pricing scheme should be adaptable to time changes. Simple but not simplistic price-demand modeling. We model the price-demand dependency employing second order differential equations with constant parameters. This modeling is flexible enough to represent a wide variety of demands as a function of price. The simplification of using constant parameters allows their easy estimation based on given price-demand data sets. The model takes into account that structures can be available in the cache or can be discarded if there is not enough respective demand. Optional structure availability allows for optimal scheduling of structure availability, such that the cloud profit is maximized. The model of price-demand dependency for a set of structures incorporates their correlation in query execution. Price adaptively to time changes. Profit maximization is pursued in a finite long-term horizon. The horizon includes sequential non overlapping intervals that allow for scheduling structure availability. At the beginning of each interval, the cloud redefines availability by taking offline some of the currently available structures and taking online some of the unavailable ones. Pricing optimization proceeds in iterations on a sliding time window that allows online corrections on the predicted demand, via reinjection of the real demand values at each sliding instant. Also, the iterative optimization allows for redefinition of the parameters in the price-demand model, if the demand deviates substantially from the predicted. Modeling structure correlations. Our approach models the correlation of cache structures as a dependency of the demand for each structure on the price of every available one. Pairs of structures are characterized as competitive, if they tend to exclude each other, or collaborating, if they coexist in query plans. Competitive pairs induce negative, whereas collaborating pairs induce positive correlation. Otherwise correlation is set to zero. The index-index, index column, and column-column correlations are estimated based on proposed measures that can estimate all three types of correlation.

## 2. Contributions:

This paper makes the following contributions:

- A novel demand-pricing model designed for cloud caching services and the problem formulation for the dynamic pricing scheme that maximizes profit and incorporates the objective for user satisfaction.
- An efficient solution to the pricing problem, based on nonlinear programming, adaptable to time changes.
- A correlation measure for cache structures that is suitable for the cloud cache pricing scheme and a method for its efficient computation.
- An experimental study which shows that the dynamic pricing scheme outperforms any static one by achieving 2 orders of magnitude more profit per time unit.

ISSN 2278-3091

**International Journal of Advanced Trends in Computer Science and Engineering**, Vol.2 , No.1, Pages : 259 – 263 (2013)
*Special Issue of ICACSE 2013 - Held on 7-8 January, 2013 in Lords Institute of Engineering and Technology, Hyderabad*

## 3. QUERY EXECUTION MODEL

The cloud cache is a full-fledged DBMS along with a cache of data that reside permanently in back-end databases. The goal of the cloud cache is to offer cheap efficient multiuser querying on the back-end data, while keeping the cloud provider profitable. Our motivation for the necessity of such a cloud data service provider derives from the data management needs of huge analytical data, such as scientific data for example physics data from CERN and astronomy data from SDSS. Furthermore, a viable, and moreover, profitable data service provider can achieve cost and time efficient management of smaller scientific collections or any type of analytical data, such as digital libraries, multimedia data, and a variety of archived data. Users pose queries to the cloud, which are charged in order to be served. Following the business example of Amazon and Google, we assume that data reside in the same data center and that users pay on-the-go based on the infrastructure they use, therefore, they pay by the query. Service of queries is performed by executing them either in the cloud cache or in the back-end database. Query performance is measured in terms of execution time. The faster the execution, the more data structures it employs, and therefore, the more expensive the service. We assume that the cloud infrastructure provides sufficient amount of storage space for a large number of cache structures. Each cache structure has a building and a maintenance cost. Fig. 2 presents at a high level the query execution model of the cloud cache. The names of variables and functions are self-explanatory. The user query is executed in the cache if all the columns it refers to are already cached. Otherwise it is executed in the back-end databases. The result is returned to the user and the cost is the query execution cost (the cost of operating the cloud cache or the cost of transferring the result via the network to the user). The cloud cache determines which structures (cached columns, views, indexes) S to build in order to accelerate query execution and reduce the query execution cost. Initially S is empty and gradually it is filled with structures that would have or have benefitted past queries. How S is populated and how the costs of building and maintaining cache structures as well as the query execution cost are computed is an input to the presented optimal pricing scheme. More details on these issues can be found in . Periodically (on predefined time intervals t[i]) the cloud performs the pricing scheme proposed in this work. The pricing scheme schedules the availability and sets the prices P of the structures S for a time horizon T as described in the rest of the paper. The goal is to maximize the provider's profit and at the same time ensure that the user is not overcharged.

## 4. SOLVING THE OPTIMAL PRICING PROBLEM

The problem of optimal pricing is an optimal control problem [6] with a finite horizon, i.e., the maximum time of optimization T is a given finite value. The free variables are the prices of the cache structures, Pis, called the control variables, and the dependent variables, called state variables, is the demand for the structures, is and the availability of the structures is. The problem is augmented with bounds on the values of both the control and the state variables and by a constraint on the dependency type of the state on the control variables.

### 4 .1 Designing the Solution

The objective function of the problem is the maximization of an integral, the optimality scope of the sought solution depends on the convexity of the objective function. The latter is bilinear w.r.t. the demand and the price. It is not possible to prove that the objective function is convex and, therefore, there is no guarantee of global optimality of the solution. Due to: 1) the nonlinearity of the objective function, 2 ) the presence of both integer inputs (the Δ is control binary variables) and continuous inputs and states (the Pis and the is , respectively),

**International Journal of Advanced Trends in Computer Science and Engineering**, Vol.2 , No.1, Pages : 259 – 263 (2013)
*Special Issue of ICACSE 2013 - Held on 7-8 January, 2013 in Lords Institute of Engineering and Technology, Hyderabad*

and 3) the potentially large scale of the system (when m is high), it is almost impossible to find an analytical solution to the optimization problem. This calls for numerical optimization techniques, such as mixed integer nonlinear programming (MINLP), which present the advantage of being implementable online. A way to implement dynamic optimization tools on real systems is to proceed as follows:

1. Solve the MINLP problem along a fixed prediction horizon to compute a sequence of values for the control variables.
2. Apply the first values to the system.
3. Slide the prediction horizon and go back to 1. This approach, referred to as Optimal Control with Receding Horizon or as Model Prective Control (for which a trajectory is tracked) in the control literature, has been successfully applied to a very large number of uncertain, complex, and nonlinear systems, in simulation as well at lab or industrial scales. This methodology has shown its ability to improve the performances of a large class of systems, despite the use of simplified models, the presence of uncertainty on model parameters, model mismatch, and process disturbances. We propose the division of the prediction horizon [0,T] into time intervals:.

## 5. CONCLUSION

This work proposes a novel pricing scheme designed for a cloud cache that offers querying services and aims at the maximization of the cloud profit. We define an appropriate price-demand model and we formulate the optimal pricing problem. The proposed solution allows: on one hand, long term profit maximization, and, on the other, dynamic calibration to the actual behavior of the cloud application, while the optimization process is in progress. We discuss qualitative aspects of the solution and a variation of the problem that allows the consideration of user satisfaction together with profit maximization. The viability of the pricing solution is ensured with the proposal of a method that estimates the correlations of the cache services in an time-efficient manner.

## 6. REFERENCES

1. http://aws.amazon.com/, 2011.
2. http://www.microsoft.com/azure/, 2011.
3. http:/www.cern.ch/, 2011.
4. Series of Meetings of the EPFL-IC-IIF-DIAS Lab with the Data Management Group of the European Organization for Nuclear Research (CERN) Started on the, Dec. 2008.
5. D. Dash, V. Kantere, and A. Ailamaki, “An Economic Model for Self-Tuned Cloud Caching,” Proc. IEEE Int’l Conf. Data Eng. (ICDE ’09), 2009.
6. I.E. Grossmann and Z. Kravanja, Large-Scale Optimization with Applications: Optimal Design and Control. Springer, 1997.