Differential Query Maintenance using Cost- Adequate Clouds

[1] Mohammed Riyazuddin [2] meeravali shaik [3] Varanasi Aruna [4]H.Balaji

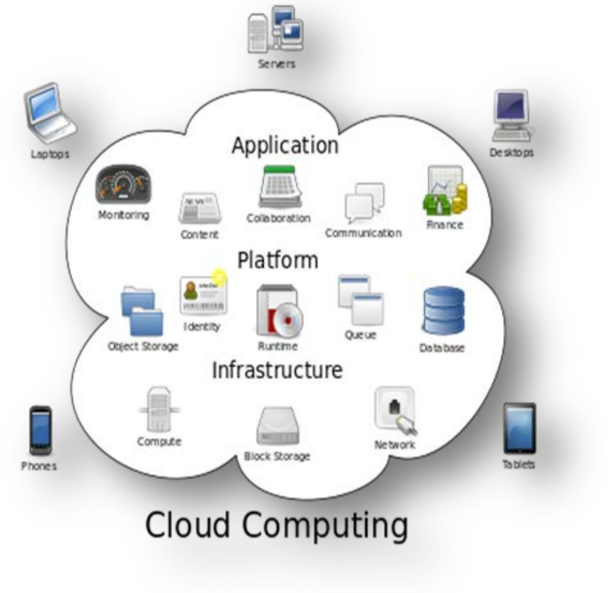
[1]M.Tech Student, Computer Science and Engineering, Sreenidhi Institute of Science and Technology, Yamnampet, Ghatkesar, Telangana State, India.
[2]Assistant Professor, Sreenidhi Institute of Science and Technology, Yamnampet, Ghatkesar, Telangana State, India.
[3]Head of Dept. and Professor, Sreenidhi Institute of Science and Technology, Yamnampet, Ghatkesar, Telangana State, India
[4]Associate Professor, Sreenidhi Institute of Science and Technology, Yamnampet, Ghatkesar, Telangana State,

ABSTRACT: Cloud computing as an rising technology trend is expected to reshape the advances in information technology. In a cost-efficient cloud environment, a user can tolerate a certain degree of hold-up while retrieving information from the cloud to reduce costs. In this paper we address two fundamental issues in such an environment: privacy and efficiency. We first review a private keyword-based file recovery scheme that was originally proposed by Ostrovsky. Their scheme allows a user to retrieve files of interest from an untrusted server without leaking any information. The main disadvantage is that it will cause a heavy querying overhead incurred on the cloud and thus goes against the original intention of cost efficiency. In this paper we present three efficient information retrieval for ranked query (EIRQ) schemes to decrease querying overhead incurred on the cloud. In EIRQ queries are classified into multiple ranks where a advanced ranked query can retrieve a higher percentage of matched files. A user can retrieve files on demand by choosing queries of different ranks. This feature is useful when there are a large number of coordinated files but the user only needs a small subset of them. Under different parameter settings extensive evaluation have been conducted on both analytical models and on a actual cloud environment in order to examine the effectiveness of our schemes.

INTRODUCTION

What is cloud computing?

Cloud computing is the use of computing capital (hardware and software) that are delivered as a check over a network (typically the Internet). The name comes from the common use of a cloud-shaped symbol as an concept for the complex infrastructure it contains in system diagrams. Cloud computing entrusts remote military with a user's data software and computation. Cloud computing consists of hardware and software capital made available on the Internet as managed third-party services. These military typically provide access to advanced software applications and high-end networks of server computers.



Structure of cloud computing

How Cloud Computing Works?

The aim of cloud computing is to apply established supercomputing or high-performance computing power normally used by military and research facilities to execute tens of trillions of computations per second in consumer-oriented applications such as financial portfolios to deliver personalized information to provide data storage or to power great immersive computer games. The cloud computing uses networks of large groups of servers typically running low-cost consumer PC technology with specialized connections to spread data-processing chores across them. This common IT infrastructure contains large pools of systems that are linked together. Often virtualization techniques are used to maximize the power of cloud computing.

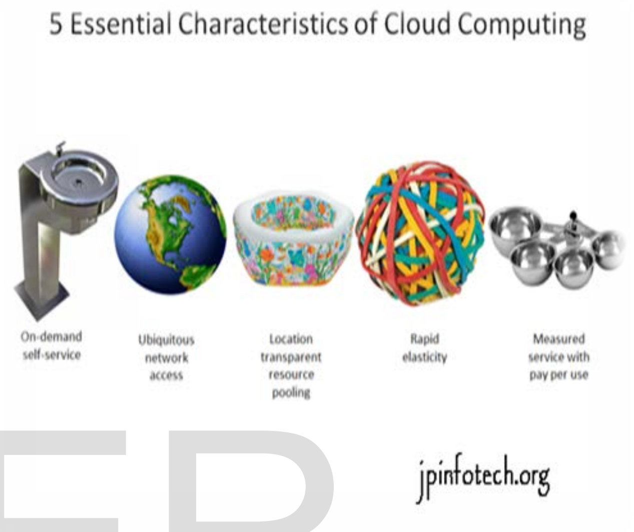
Characteristics and Services Models:

The most important characteristics of cloud computing based on the definitions provided by the National Institute of Standards and Terminology (NIST) are outlined below:

- **On-demand self-service**: A customer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring person interaction with each service's provider.
- **Broad network access**: Capabilities are available over the network and accessed through standard mechanisms that promote use by assorted thin or thick client platforms (e.g., cellular phone phones, laptops, and PDAs).
- **Resource pooling**: The provider's computing resources are pooled to serve multiple consumers using a

multi-tenant model with different physical and virtual capital dynamically assigned and reassigned according to consumer demand. There is a sense of location-independence in that the customer generally has no control or knowledge over the exact position of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or data center). Examples of resources contain storage, processing, memory, network bandwidth, and virtual machines.

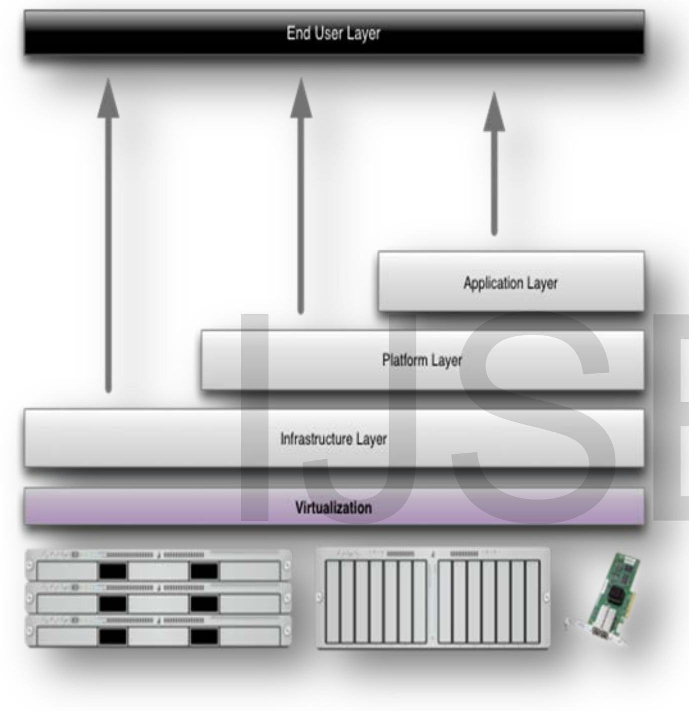
- **Rapid elasticity**: Capabilities can be rapidly and elastically provisioned, in some cases automatically to quickly scale out and fast released to quickly scale in. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be purchased in any quantity at any time.
- **Measured service**: Cloud systems repeatedly control and optimize resource use by leveraging a metering capability at some level of abstraction suitable to the type of service (e.g. storage space, processing, bandwidth and active user accounts). Resource usage can be managed, controlled and reported providing transparency for both the provider and shopper of the utilized service.


5 Essential Characteristics of Cloud Computing
On-demand self-service
Ubiquitous network access
Location transparent resource pooling
Rapid elasticity
Measured service with pay per use
jpinfotech.org

Characteristics of cloud computing

Services Models:

Cloud Computing comprises three different service models, namely Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS). The three examine models or layer are completed by an end user layer that encapsulates the end user perspective on cloud services. The model is exposed in figure below. If a cloud user accesses services on the infrastructure layer for instance she can run her own applications on the resources of a cloud infrastructure and stay put responsible for the support, maintenance and security of these applications herself. If she accesses a service on the application layer these tasks are normally taken care of by the cloud service provider.



Structure of service models

Benefits of cloud computing:

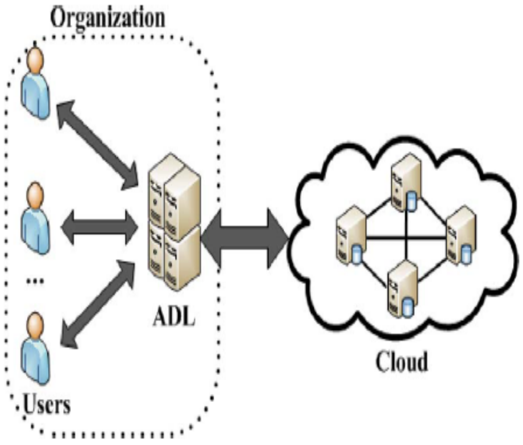
1. **Accomplish economies of scale –** increase volume output or productivity with fewer public. Your cost per unit, project or product plummets.
2. **Reduce expenses on technology infrastructure.** Maintain easy admission to your information with minimal sincere spending. Pay as you go (weekly, quarterly or yearly), based on demand.
3. **Globalize your labor force on the cheap.** People worldwide can admission the cloud, provided they have an Internet connection.
4. **reorganize processes.** Get more work done in less moment with less people.
5. **Reduce capital costs.** There's no need to waste big money on hardware, software or licensing fees.
6. **Improve accessibility.** You have access anytime, everyplace, making your life so much easier!
7. **Monitor projects more effectively.** Stay within resources and ahead of completion cycle times.
8. **Less personnel training is needed.** It takes fewer people to do more work on a cloud, with a minimum learning curve on hardware and software issues.
9. **Minimize licensing new software.** enlarge and grow without the need to buy expensive software licenses or programs.

10. **Improve flexibility.** You can change means without serious people or financial issues at stake.

Advantages:

1. **Price:** Pay for only the resources used.
2. **Security**: Cloud instances are inaccessible in the network from other instances for improved security.
3. **Performance:** Instances can be added instantly for improved performance. Clients have access to the total property of the Cloud's core hardware.
4. **Scalability:** Auto-deploy cloud instances when needed.
5. **Uptime:** Uses multiple servers for maximum redundancies. In case of server failure, instances can be repeatedly created on another server.
6. **Control:** Able to login from any location. Server snapshot and a software library lets you deploy custom instances.
7. **Traffic:** Deals with spike in traffic with quick deployment of extra instances to handle the load.

SYSTEM ARCHITECTURE



EXISTING SYSTEM:

Private searching was proposed by Ostrovsky et al. Which allows a user to retrieve files of attention from an untrusted server without leaking any information. Otherwise the cloud will learn that certain files without processing are of no interest to the user. industrial clouds follow a pay-as-you-go model where the customer is billed for different operations such as bandwidth, CPU time, and so on. Solutions that incur excessive computation and communication costs are unacceptable to customers. To formulate private searching applicable in a cloud environment, our previous work designed a cooperate private searching protocol (COPS) where a stand-in server called the aggregation and distribution layer (ADL) is introduced linking the users and the cloud. The ADL deployed inside an organization has two major functionalities: aggregating user queries and distributing search results. below the ADL the computation cost incurred on the cloud can be largely reduced since the cloud only needs to execute a joint query once no matter how numerous users are executing queries. Furthermore the communication cost incurred on the cloud will also be reduced, since files shared by the users require to be returned only once. Most importantly by using a series of secure functions COPS can defend user privacy from the ADL the cloud and other users.

DISADVANTAGES OF EXISTING SYSTEM

1. Ostrovsky scheme has a soaring computational cost while it requires the cloud to process the query on every file in a collection.

2. It will quickly become a performance bottleneck when the cloud needs to procedure thousands of queries over a collection of hundreds of thousands of records. We argue that subsequently proposed improvements, like also have the same drawback.

PROPOSED SYSTEM:

In this paper we introduce a novel concept differential query military to COPS where the users are allowed to personally decide how many matched files will be returned. This is motivated by the fact that under confident cases there are a lot of files matching a user’s query but the user is interested in only a certain percentage of matched files In the Ostrovsky scheme the cloud will have to return 2,000 records. In the COPS scheme the cloud will have to return 1,000 files. In our scheme the cloud only needs to return 200 files. Therefore by allowing the users to retrieve matched records on demand the bandwidth consumed in the cloud can be largely reduced. Efficient Information retrieval for Ranked Query (EIRQ) in which each user can choose the rank of his query to determine the percentage of coordinated files to be returned. The basic idea of EIRQ is to construct a privacy-preserving mask matrix that allows the cloud to filter out a certain percentage of matched records before returning to the ADL. This is not a trivial work, since the cloud needs to correctly filter out files according to the rank of queries without knowing anything about user solitude.

ADVANTAGES OF PROPOSED SYSTEM:

1. The cloud only needs to return 200 files. Therefore, by allowing the users to retrieve matched records on demand the bandwidth consumed in the cloud can be largely reduced.

2. We present two solutions to adjust related parameters one is based on the Ostrovsky scheme and the other is based on Bloom filters.

IMPLEMENTATION

MODULES

1. Differential Query military:
2. Efficient in rank Retrieval For Ranked Query:
3. Aggregation And Distribution Layer
4. Ranked Queries

MODULES DESCRIPTION

Differential Query Services:

We introduce a novel idea differential query services, to COPS, where the users are allowed to personally decide how many matched files will be returned. This is motivated by the detail that under certain cases there are a lot of files matching a user's query but the user is interested in only a certain percentage of matched files. To demonstrate let us assume that Alice wants to retrieve 2% of the files that include keywords "A, B", and Bob wants to retrieve 20% of the files that contain keywords "A, C". The cloud holds 1,000 records where *{F*1, . . . , *F*500*}* and *{F*501, . . . , *F*1000*}* are described by keywords "A, B" and "A, C", respectively. In the Ostrovsky scheme, the cloud will have to return 2, 000 records. In the COPS scheme the cloud will have to return 1, 000 files. In our scheme the cloud only needs to return 200 records. Therefore by allowing the users to retrieve matched records on demand the bandwidth consumed in the cloud can be mainly reduced.

Efficient Information Retrieval For Ranked Query:

We propose a scheme termed Efficient Information retrieval for Ranked Query (EIRQ) in which each customer can choose the rank of his question to determine the percentage of matched files to be returned. The basic idea of EIRQ is to construct a privacy preserving mask matrix that allows the cloud to filter out a certain percentage of matched records before returning to the ADL. This is not a trivial work since the cloud needs to correctly filter out files according to the rank of queries without knowing anything about user privacy. Focusing on dissimilar plan goals we provide two extensions the first extension emphasizes simplicity by requiring the least amount of modifications from the Ostrovsky scheme and the second addition emphasizes privacy by leaking the least amount of information to the cloud.

Aggregation And Distribution Layer :

An ADL is deployed in an organization that authorizes its employees to share data in the cloud. The staff members as the authorized users send their queries to the ADL which will aggregate user queries and send a combined query to the cloud. Then the cloud process the combined query on the file collection and returns a buffer that contains all of matched files to the ADL, which will distribute the search results to each user. To combined sufficient queries the organization may require the ADL to wait for a period of time before running our schemes which may incur a certain querying delay. In the supplementary records we will discuss the computation and communication costs as well as the querying delay incurred on the ADL.

Ranked Queries:

To further reduce the communication cost a differential query service is provided by allowing each user to retrieve matched records on demand. Specifically a user selects a particular rank for his query to determine the percentage of matched records to be returned. This feature is useful when there are a lot of records that match a user's query but the user only needs a small subset of them.

CONCLUSION

In this paper we proposed three EIRQ schemes based on an ADL to provide degree of difference query services while protecting user privacy. By using our schemes a user can retrieve different percentages of matched files by specifying queries of different ranks. By further dropping the communication cost incurred on the cloud, the EIRQ schemes make the private searching technique more applicable to a cost-efficient cloud environment. However, in the EIRQ schemes we only determine the rank of each organizer by the highest rank of queries it matches. For our future work we will try to design a bendable ranking mechanism for the EIRQ schemes.

REFERENCES

[1] P. Mell and T. Grance, ‘‘The NIST Definition of Cloud Computing (Draft),’’ in NIST Special Publication. Gaithersburg, MD, USA: National Institute of Standards and Technology, 2011.

[2] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, ‘‘Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions,’’ in Proc. ACM CCS, 2006, pp. 79-88.

[3] R. Ostrovsky and W. Skeith, ‘‘Private Searching on Streaming Data,’’ in Proc. CRYPTO, 2005, pp. 233-240.

[4] R. Ostrovsky and W. Skeith, ‘‘Private Searching on Streaming Data,’’ J. Cryptol., vol. 20, no. 4, pp. 397-430, Oct. 2007.

[5] J. Bethencourt, D. Song, and B. Waters, ‘‘New Constructions and Practical Applications for Private Stream Searching,’’ in Proc. IEEE SP, 2006, pp. 1-6.

[6] J. Bethencourt, D. Song, and B. Waters, ‘‘New Techniques for Private Stream Searching,’’ ACM Trans. Inf. Syst. Security, vol. 12, no. 3, p. 16, Jan. 2009.

[7] Q. Liu, C. Tan, J. Wu, and G. Wang, ‘‘Cooperative Private Searching in Clouds,’’ J. Parallel Distrib. Comput., vol. 72, no. 8, pp. 1019-1031, Aug. 2012.

[8] G. Danezis and C. Diaz, ‘‘Improving the Decoding Efficiency of Private Search,’’ Int’l Assoc. Cryptol. Res., IACR Eprint Archive No. 024, Schloss Dagstuhl, Germany, 2006.

[9] G. Danezis and C. Diaz, ‘‘Space-Efficient Private Search with Applications to Rateless Codes,’’ in Proc. Financial Cryptogr. Data Security, 2007, pp. 148-162.

[10] M. Finiasz and K. Ramchandran, ‘‘Private Stream Search at the Same Communication Cost as a Regular Search: Role of LDPC Codes,’’ in Proc. IEEE ISIT, 2012, pp. 2556-2560.

[11] X. Yi and E. Bertino, ‘‘Private Searching for Single and Conjunctive Keywords on Streaming Data,’’ in Proc. ACM Workshop Privacy Electron. Soc., 2011, pp. 153-158.

[12] B. Hore, E.-C. Chang, M.H. Diallo, and S. Mehrotra, ‘‘Indexing Encrypted Documents for Supporting Efficient Keyword Search,’’ in Proc. Secure Data Manage., 2012, pp. 93-110.

[13] P. Paillier, ‘‘Public-Key Cryptosystems Based on Composite Degree Residuosity Classes,’’ in Proc. EUROCRYPT, 1999, pp. 223-238.

[14] Q. Liu, C.C. Tan, J. Wu, and G. Wang, ‘‘Efficient Information Retrieval for Ranked Queries in Cost-Effective Cloud Environments,’’ in Proc. IEEE INFOCOM, 2012, pp. 2581-2585.

[15] S.Yu,C. Wang,K.Ren, andW. Lou, ‘‘Achieving Secure, Scalable, Fine-Grained Data Access Control in Cloud Computing,’’ in Proc. IEEE INFOCOM, 2010, pp. 1-9.

[16] G. Wang, Q. Liu, J. Wu, and M. Guo, ‘‘Hierarchical Attribute-Based Encryption and Scalable User Revocation for Sharing Data in Cloud Servers,’’ Comput. Security, vol. 30, no. 5, pp. 320-331, July 2011.

[17] M. Mitzenmacher, ‘‘Compressed Bloom Filters,’’ IEEE/ACM Trans. Netw., vol. 10, no. 5, pp. 604-612, Oct. 2002.

[18] D. Guo, J. Wu, H. Chen, and X. Luo, ‘‘Theory and Network Applications of Dynamic Bloom Filters,’’ in Proc. IEEE INFOCOM, 2006, pp. 1-12.

[19] A. Berl, E. Gelenbe, M. Di Girolamo, G. Giuliani, H. De Meer, M.Q. Dang, and K. Pentikousis, ‘‘Energy-Efficient Cloud Computing,’’ Comput. J., vol. 53, no. 7, pp. 1045-1051, Sept. 2010.