iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** IV **Month of publication:** April 2020

DOI: http://doi.org/10.22214/ijraset.2020.4042

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue IV Apr 2020- Available at www.ijraset.com

Cost Effective Cloud Server Provisioning for Predictable Performance of Big Data Analytics

Siddhant Mishra R[1] Gopinath R[2]
[1]Student, [2]Professor, K.S. Rangasamy College of Technology, Tiruchengode, Tamil Nadu

Abstract: Cloud storage provides easy-to-use, large, and cost-effective storage at a low cost, but data privacy is a major issue that prevents users from keeping files in trust. One way to improve privacy from the view of the data owner is to secretly encrypt files before exporting them to the cloud and to extract files after downloading them. However, data encryption is beyond the complexity of mobile devices, and the data recovery process creates a complex interaction between the data user and the cloud. Frequent with limited bandwidth and limited battery life, these problems present heavy computer and communication and high-power consumption of mobile device users, making encrypted searches over the mobile cloud a major challenge.

The main advantage of the Pack is its ability to load the TRE cloud server depletion clients, thus reducing the processing costs caused by the TRE algorithm In contrast to past arrangements, the Pack doesn't require the server to keep on keeping up client status. This makes the Pack more convenient for a multi-site environment including client-side migration and server migration to maintain cloud stability. The Pack depends on the novel TRE framework, which permits the customer to utilize recently obtained pieces to distinguish lump chains previously, which can be utilized as solid indicators for future lumps. You have introduced a fully functional Pack functionality, which is obvious for all TCP-enabled applications and net-work devices. Finally, the benefits of Pack users will be analyzed, using traffic from various sources.

Keywords: privacy enhancement, cloud server loading, trusted predictors.

I. INTRODUCTION

Offering assets to a perfect attachment and play model that streamlines framework arranging is the guarantee of "distributed computing. Two important benefits of this model are ease of use and cost efficiency. Although there are still questions about issues such as security and vendor locks, the benefits of this model are many. This work explores other basics of cloud interaction with a view to introducing features such as: The Realities and Dangers of Model Models in the Control and Utilization of the Model This project aims to provide model insights and explore the options available to fulfill your technology and infrastructure needs. Cloud computing is a compad paradigm, in which a large pool of applications are connected to private or public networks, providing a robust application infrastructure, data and file storage. With the introduction of this technology, the costs of integration, application management, content storage and delivery are greatly reduced. Cloud computing is an effective way to get the benefits of direct costs and has the potential to transform a data center from a complex base into a multi-priced environment. The concept of cloud computing is based on the most important principal of IT 're-engineering'. The difference that comes with cloud computing compared to traditional concepts of "grid computing", "computer computing", "utility computing", or "autonomic computing" is to expand the boundaries of an organization.

A. Cloud Providers Offer Administrations That Can Be Assembled Into Three Classifications

1) Software as a Service (SaaS): In this service, a complete application is made to the customer, such as the desired service. One example of a service is running in the cloud and many end users are created. On the consumer side, there is no need to invest further in servers or software licenses, while for a provider, the cost is reduced, since only one request needs to be handled and maintained.
2) Platform as a Service (Paas): Here, a software layer, or development environment is fenced and provided as a service, where some of the highest levels of service can be built. The client has the opportunity to construct the frameworks, which run on the supplier's foundation. To meet the requirements for application management and vulnerabilities, PaaS providers provide pre-defined combinations of OS servers and applications, such as LAMP platform (Linux, Apache, My Sql and PHP), J2EE, Ruby, etc. Google Operations Engine, Force.com, etc. some of the most popular PaaS examples.
3) Infrastructure as a Service (Iaas): IaaS provides basic storage and computing capabilities as standard services on the network. Servers, storage systems, networking machines, database center etc are built and available to carry workload. Usually the customer sends their software through the infrastructure. Some regular models are Amazon, Go Grid.

B. Public Cloud

Public clouds are owned and operated by third parties; they deliver the highest economies to customers, the expense of foundation is conveyed among blended clients, giving every customer an alluring ease, "Pay-more only as costs arise" model. All clients share a similar framework pool with constrained design, security insurance, and accessibility contrasts. These are overseen and upheld by the cloud supplier. One of the benefits of public cloud is that it can be larger than enterprise cloud, thus giving it the ability to scale seamlessly, on demand.

C. Hybrid Cloud

Hybrid clouds integrate public and private cloud models. With Hybrid Cloud, service providers can leverage 3rd party cloud providers in whole or in part thereby increasing computational flexibility. The Hybrid cloud environment can provide the desired, scaled-out scale. The capacity to make a private cloud with open cloud administrations can be utilized to deal with any startling reservations in remaining task at hand.

D. Private Cloud

Private clouds are created for only one business. They aim to address data security concerns and provide greater control, which often fails in the public cloud.

E. Mobile Cloud

In an increasingly connected world, users enter personal or shared data, stored in “cloud” by multiple devices. In addition to the popularity of cloud storage services, little work has focused on investigating cloud storage users' Hearing quality especially on mobile devices. In addition, it is unclear how the total number of users may affect QoE. It conducted an online survey with 349 cloud service users to gain insight into their use and cost. In a 2-week follow-up study, It looked at mobile cloud usage on tablets and smart phones, in real time using the Experience Sampling Way (ESM) mobile questionnaire. It collected 156 responses in the context of Drop box usage on mobile devices. It provides data on QoE-based cloud services by highlighting important mobile features (e.g. communication, location, social, tool), and how they affect users' experiences while using such services on their mobile devices. These devices provide a huge range of new costs for users and enable communication anywhere, anytime or any service. However, one disproportionate evil of these weights is, mobile devices have their ultimate storage capacity, and they limit their potential use. It is also common to find apps and apps running on different mobile devices and platforms.

II. EXISTING SYSTEM

A. To successfully support encrypted search systems with a high level of security over cloud data, It introduced a new technology It called G-SIR. This goal is to design an effective solution for secure search over mobile cloud storage. It first introduced the architecture concept and introduced the development of This protocol with the change of the traditional file search and cloud data restoration system. This program achieves the security and efficiency goals described above. After that, it discusses the reasons why Pack can achieves effectiveness.

B. The basic idea behind Pack is to load the calculations and the position of the relevant loads in the cloud. It is highlighted that loading some cold applications to the cloud can be a low-energy design philosophy that works well. Cloud providers can provide computer cycles, and users can use these cycles to reduce the amount of integration in the mobile system and save energy. However, at the same time, downloaded applications aim to increase the number of transfers and thus increase the power consumption from another feature. These two results prompt us to carefully reconsider the traditional file for shipment and the retrieval process. It first reviews the major processes of all file search processes and return programs. There are usually three main processes:

1) The authentication process is used by the data owner to authenticate users of the data.

2) The file set and its directory are stored in the cloud after encrypting by the data owner during the expansion and import of the index.

3) The data user searches for keywords related files by sending the request to the cloud server for search and retrieval processes.

It now presents a detailed design of how the Pack addresses energy efficiency and security challenges to streamline these processes. During the magnification and indexing phase, the data holder receives the TF table as references and uses the Order Preserving Encryption (OPE) to encrypt it. Because of this, the cloud server is able to calculate compliance measurements and set them without specifying an index. This rendering of computational load is safe and feasible. Therefore, the modified and retrieved search procedures for G-SIR.

III. PROPOSED SYSTEM

The main advantage of the Pack is its ability to load the TRE cloud server depletion clients, thus reducing the processing costs caused by the TRE algorithm. In contrast to past arrangements, the Pack doesn't require the server to keep on keeping up client status. This makes the Pack more convenient for a multi-site environment including client-side migration and server migration to maintain cloud stability. The Pack depends on the novel TRE framework, which permits the customer to utilize recently obtained lumps to recognize piece chains previously, which can be utilized as solid indicators for future pieces. You have introduced a fully functional Pack functionality, which is obvious for all TCP-enabled applications and net-work devices. Finally, the Pack benefits the users of the cloud, using traffic from various sources.

In this work, it proposes the Pack (Text Search for Traffic and Power), the bandwidth and power spent on the art of building a hidden surface on a mobile cloud. The proposed architecture spreads the load from mobile devices to the cloud, and also improves communication between mobile and cloud customers. It is shown that data privacy does not compromise when optimizing performance measures. This test show that the Pack reduces the computation time by 23% to 46% and saves the power consumption by 35% to 55% on file restoration, and, in fact, the network's smugglers during file recovery are also significantly reduced.

It uses This OPE for most users in the data owner module. It also rolled out the keywords to be searched by adding some sound to the data center module to help control keywords-files association leak. To find the right top-k files, it uses standard function to calculate the correct points in the cloud.

Given the key name in ORS, the cloud server is responsible for calculating the corresponding data user scores to find the corresponding top-k files. Therefore, It use both the abstract objects and the positions in the cloud server module.

In this work, it proposes the Pack (Text Search for Traffic and Power), the bandwidth and power spent with embedded architecture over a mobile cloud. The proposed architecture spreads the load from mobile devices to the cloud, and also improves communication between mobile and cloud customers. It is shown that data privacy does not compromise when optimizing performance measures. This test show that the Pack reduces the computation time by 23% to 46% and saves the power consumption by 35% to 55% on file restoration, and, in fact, the network's smugglers during file recovery are also significantly reduced.

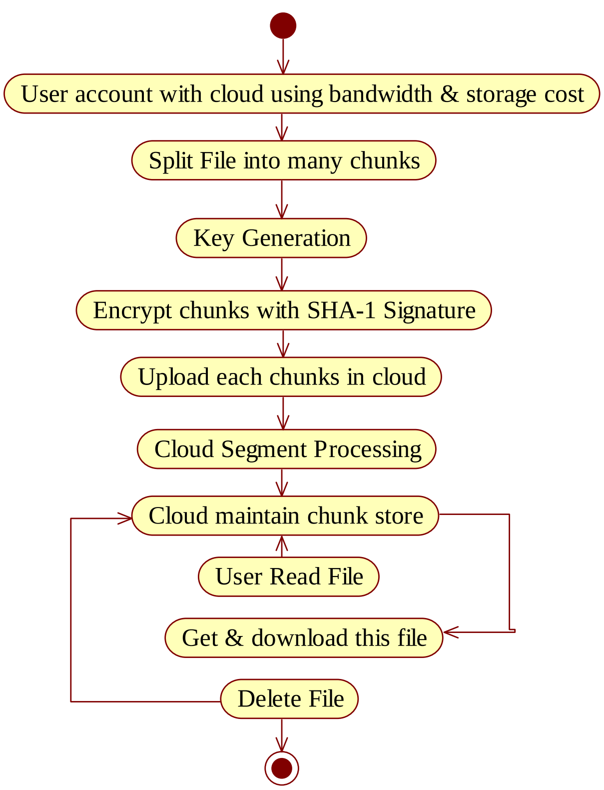

Figure 3.1 Flow Diagram

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue IV Apr 2020- Available at www.ijraset.com

A. Modified Process of Search and Retrieval

During the magnification and indexing phase, the data holder receives the TF table as references and uses the Order Preserving Encryption (OPE) to encrypt it. Because of this, the cloud server is able to calculate compliance measurements and set them without specifying an index. This rendering of computational load is safe and feasible. Therefore, the modified and retrieved search procedures for G-SIR.

$$\eta - \frac{E_{co}+E_{retr}}{2E_{co}+E_{rotr}+E_{w}} < 100\% \qquad (1)$$

As the image calculations, calculation is transferred to the cloud due to the OR-G-SIR method, it is completed. ORS combines file search and recovery processes in one (only one) cycle.) is the same for both TRS and ORS. Note that it is smaller than 1, and the energy consumption can be reduced with G-SIR. In addition, as it is quite large, it can predict that the power consumption of the mobile device is significantly reduced by G-SIR, which is clearly shown in This experiment.

B. Reducing File Search and Retrieval Time

Reducing the output time of file search and retrieval is important for the user experience. Note that beyond the 'single-cycle' time saved by G-SIR, the computational time is reduced by reloading the computational load from the user side to the cloud side.

Assume that the computing ability of the cloud server and of the mobile clients are denoted by C_{cs} and C_m (C_m<< C_{cs}), respectively. Then, the searching workload is Σ, and T_{retr} denotes the file retrieving time, while RTT represents one round trip time. The reduction ratio of the file search to the retrieval time, ρ, is given by Equation (2), which is smaller than 1 and proves the computing efficiency by offloading of G-SIR.

$$\rho = \frac{RTT+\frac{\Sigma}{C_{cs}}+T_{retr}}{2RTT+\frac{\Sigma}{C_{m}}+|T_{retr}} < 100\% \qquad (2)$$

C. Reducing Traffic Overhead

It is important to reduce the maximum communication to ensure that it does not cancel other performance improvements. In G-SIR, there is only one return link for each keyword search, and the selected index is not distributed between the cloud and the user as shown in the TRS case. This attribute greatly reduces any overload of communication.

IV. METHODLOGY

The following are the basic modules involved in the project:

1) *Reconstructing Data Owner Module:* It are changing the way to build an ORS support system for many of This OPEs and use it to control leak information. Authentication between data owner and user data is renamed to ensure G-SIR security. It now discusses the implementation of index construction, encryption functions and detail the authentication process.
2) *Design of a data Center Module:* The data user module is created on the mobile client side. Keyword wrap operation is performed to resolve keywords - file combinations. In the wrapping function, the stem, encryption and hash functionality are the same as the index construction algorithm. The file printing function corresponds to the encryption done by the data owner. The authentication function is used for authentication. It now has details of the operation of the wrap for this module.
3) *Cloud Module Design:* It will describe the functions that override keywords and measure the relevant points of the cloud module. These functions are used to find the correct top-k files according to the given keyword.

a) *Unwrap Function:* Note that the cloud server is less reliable, and therapy function can be run on the server. After receiving duplicate server calls get)) = h, searches the TF table, and sends the same files.
b) *Ranking Function:* The cloud server calculates the correct scores and returns the correct top-k files according to a search query from the data user. The calculation system is used in This system. Note that due to the preservation index, any other method of calculating the points can be employed. The Pack lists the relevant points as Equation

V. IMPLEMEMTATION

In This experiment, It use a data set of 1000 files of different sizes and a VM in the cloud with Dual vCPUs at 2.27GHz. A smart android phone with a 1GHz CPU sends queries as a Packet mobile client with an 8M wireless network. The Android system detects the user input and compiles it before it receives the hash value and then wraps it into a component sent to the mobile cloud server.

An alternative to installing this program is to retrieve files back to the mobile cloud server and separate them. In addition, It used TRS and PTS for comparison purposes. Since power consumption is important for mobile devices, It analyze the energy efficiency of the Pack in this subsection. It use the phone's power line to accurately measure the system's power consumption. Although small changes depending on the environment may occur, the comparison is as accurate as the randomized controlled trials. Note that power consumption is reduced from 0.08mAh to 0.036mAh when searching and retrieving 100KB size files, which means that ORS saves 55% power compared to TRS. When searching for and finding files of 1MB size, power consumption is reduced from 0.164mAh to 0.106mAh, which means energy savings of 35%. Therefore, the Pack provides the most efficient power consumption. For example, to take out This 1650mAh battery, ORS (for G-SIR) can operate at 22000 returns while TRS can only restore 13000 files in size 600KB.

A. *File Search and Retrieval Time*

It compares the File Searching and Retrieval Time (FSRT) for the three programs in this subsection as shown in Figure 10. It tests FSRT to find different files in size from 100KB to 1MB. It sees that the FSRT for PTS is too short because it doesn't have to do any of the security. The FSRT of the ORS is successfully reduced compared to that of the TRS. This difference is due to the advantages of Pack design in terms of calculating load points, and that results in a reduction in file search and retrieval process.

The FSRT value of ORS is very close to PTS, which means the lowest cost of security on a mobile device. For example, Pack saves FSRT by 46% compared to TRS for 100KB files, and 23% for 1MB files .The file recovery time depends only on file size and network bandwidth. Given the large bandwidth, the Pack works well since the download time of the files becomes a box for other applications. The file view time is the same for all schemes so it is not reasonable to estimate it.

	PTS	TRS	ORS
Request/Response	190ms	370ms	190ms
Stemming and Encryption	0	10ms	10ms
Hash and Wrap	0	145ms	150ms
Server file search	80ms	70ms	75ms
Client file search	0	260ms	0
Sum	270ms	855ms	425ms

B. *FSRT analysis of PTS, TRS and ORS*

An efficient FSRT for Pack is achieved by optimizing the process, since only one circle is used to communicate and escape points. The search procedure is analyzed in Table 1. Apart from the security service, PTS (Plain Text Search) does not spend any time in building and encrypting; and does not perform in hash and wrap. On the other hand, ORS and TRS provide search engines that are embedded in related quantities. As shown in Table 1, ORS can improve the "request / response" time more than TRS from 370ms to 190ms (saves 180ms), and reduce the "client file search" time by sending it to the server (260ms). Note that the "trace file search" ORS calculation load is 75ms, 5ms longer than TRS.

This is explained by the fact that the server takes the computed load count of the mobile user. In other words, the Pack eliminates the time to "search the client file" at the expense of the "server file search" time.

This proves that reloading works well (5ms vs. 260ms). In addition, ORS spends 5ms more on hash value than TRS to improve security. Note that the "server file search" time of PTS is higher than the other two schemes, since the server has to work with the hash function of the logical file search, while the hash functions are performed by the mobile data user in TRS and ORS. Overall, ORS is safe and effective.

C. *Throughput*

The calculation of uploads from the mobile device to the cloud data center reduces the computation time of the calculation article due to high server capacity. Therefore, this greatly enhances the system through the use of FSRT; It find that file access speeds are very effective when dealing with small files as the computation of the aptitude score is often done. For example, in a 100KB file, the access speed increased from 104KB / s to 194KB / s, almost doubling that box. Speed is still active when accessing files in size 1MB (29.6% speed). ORS output is not much smaller than that of PTS.

Used to carry out encrypted searches on a mobile cloud. Pack security research shows that it is safe enough for a mobile cloud computer, and a series of experiments have highlighted its effectiveness. Pack is a bit more time consuming and powerful than keyword searches over text, but at the same time saves a lot of power compared to traditional strategies with the same level of security. Based on G-SIR, this function can be extended. It have proposed a single keyword search program to make the data encryption search more efficient. However, there are still some extensions for This remaining work. It wish to propose a multi-keyword search system to carry out hidden searches of data on the mobile cloud in the future Since This OPE algorithm is simple, another extension is to find a powerful algorithm that will not compromise efficiency.

VI. CONCLUSION & FUTURE SCOPE

In this work, It developed a new architecture, PACK as the first attempt to design a keyword search tool for traffic and power embedded over mobile cloud storages. It started with the introduction of a basic program that compared it to the encrypted search tools of the cloud computer and demonstrated its inefficiency in the mobile cloud environment. After that It built a great implementation to get the search embedded in the mobile cloud. Pack security research shows that it is safe enough for a mobile cloud computer, and a series of tests have highlighted its effectiveness. Pack is more time and energy efficient than keyword searches over embedded text, but at the same time retains more power compared to traditional strategies with the same level of security. Based on G-SIR, this function can be extended. It has proposed a single keyword search program to make the data encryption search more efficient. However, there are still some extensions for This remaining work. It wishes to propose a multi-keyword search system to carry out hidden searches of data on the mobile cloud in the future Since This OPE algorithm is simple, another extension is to find a powerful algorithm that will not compromise efficiency.

A. Future Work

Many different scenarios, tests, and tests are left to the future due to lack of time (e.g. real-time test data is often very time-consuming, it even requires days to complete a single run). Future work addresses in-depth analysis of specific processes, new suggestions for trying different approaches, or simply knowing.

It may be interesting to look at the regions in this model and the data images of different importance, depending on their size or specification on the recognition process. This approach may have helped to isolate the most complex problems that are regions that are important to be identified, which are sometimes seen, and those that do not.

VII. ACKNOWLEDGEMENT

It acknowledges DST-File No. 368, DST--FIST (SR/FIST/College-235/2014 dated 21Nov 2014) for financial support and DBT-STAR-College-Scheme-ref.no: BT/HRD/11/09/2018 for providing infrastructure support.

REFERENCES

[1] Agrawal R., Kiernan J., Srikant R., and Xu Y., “Order preserving encryption for numeric data,” in Proceedings of the 2004 ACM SIGMOD international conference on Management of data. ACM, 2004, pp. 563–574.

[2] Blei D. M., Jordan M. I., and Ng A. Y., “Latent dirichlet allocation,” the Journal of machine Learning research, vol. 3, pp. 993–1022, 2003.

[3] Boneh D., Di Crescenzo G., Ostrovsky R., and Persiano G., “Public key encryption with keyword search,” in Advances in Cryptology- Eurocrypt 2004. Springer, 2004, pp. 506–522.

[4] Caceres J., Lindner M., Rodero-Merino L., and Vaquero L., “A break in the clouds: towards a cloud definition,” ACM SIGCOMM Computer Communication Review, vol. 39, no. 1, pp. 50–55, 2008.

[5] Cao N., Li M., Lou W., Ren K., and Wang C., “Privacy-preserving multi-keyword ranked search over encrypted cloud data,” Parallel and Distributed Systems, IEEE Transactions on, vol. 25, no. 1, pp. 222–233, 2014.

[6] Cao N., Li M., Lou W., Ren K., and Wang C., “Privacy-preserving multi-keyword ranked search over encrypted cloud data,” in INFOCOM, 2011 Proceedings IEEE. IEEE, 2011, pp. 829–837.

[7] Cao N., Lou W., Ren K., and Wang C., “Enabling secure and efficient ranked keyword search over outsourced cloud data,” Parallel and Distributed Systems, IEEE Transactions on, vol. 23, no. 8, pp. 1467–1479, 2012.

[8] Chai Q. and Gong G., “Verifiable symmetric searchable encryption for semi-honest-but-curious cloud servers,” in Communications (ICC), 2012 IEEE International Conference on. IEEE, 2012, pp. 917– 922.

[9] Chang Y. and Mitzenmacher M., “Privacy preserving keyword searches on remote encrypted data,” in Applied Cryptography and Network Security. Springer, 2005, pp. 391–421.

[10] Cooke E., Flinn J., Jahanian F., Oberheide J., and Veeraraghavan K., “Virtualized in-cloud security services for mobile devices,” in Proceedings of the First Workshop on Virtualization in Mobile Computing. ACM, 2008, pp. 31–35.

[11] Curtmola R., Garay J., Kamara S., and Ostrovsky R., “Searchable symmetric encryption: improved definitions and efficient constructions,” in Proceedings of the 13th ACM conference on Computer and communications security. ACM, 2006, pp. 79–88.

[12] Hou Y. T., Li M., Lou W., Ren K., and Yu S., "Toward privacy assured and searchable cloud data storage services," Network, IEEE, vol. 27, no. 4, pp. 56–62, 2013.
[13] Huang D., "Mobile cloud computing," IEEE COMSOC Multimedia Communications Technical Committee (MMTC) E-Letter, 2011.
[14] Jahanian F. and Oberheide J., "When mobile is harder than fixed (and vice versa): demystifying security challenges in mobile environments," in Proceedings of the Eleventh Workshop on Mobile Computing Systems ACM, 2010, 43 48.
[15] Kamara S. and Lauter K., "Cryptographic cloud storage," in Financial Cryptography and Data Security. Springer, 2010, pp. 136– 149.
[16] Moffat A., T. C. Bell et al., Managing gigabytes: compressing and indexing documents and images. Morgan Kaufmann Pub, 1999.
[17] Moffat A., and Zobel J., "Inverted files for text search engines," ACM Computing Surveys (CSUR), vol. 38, no. 2, p. 6, 2006.
[18] Nejdl W., Olmedilla D., Siberski W., and Zerr S., "Zerber+ r: Topk retrieval from a confidential index," in Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology. ACM, 2009, pp. 439–449.
[19] Perrig A., Song D., and Wagner D., "Practical techniques for searches on encrypted data," in Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on. IEEE, 2000, pp. 44– 55.
[20] Ramos J., "Using tf-idf to determine word relevance in document queries," Technical report, Department of Computer Science, Rutgers University, 2003.