International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2014): 5.611

Secure Dynamic Fragment and Replica Allocation in Large-Scale Distributed File Systems

M. Kanchana[1], R. Ruhin Kouser[2]

[1]ME-CSE, Kingston Engineering College, Vellore, India

[2]Assistant Professor -CSE, Kingston Engineering College, Vellore, India

Abstract: ***The Division and Replication of Data in the Cloud for Optimal Performance and Security (DROPS) that collectively approaches the security and performance issues. Data to a third-party administrative control, as is done in cloud computing, gives rise to security concerns. The data compromise may occur due to attacks by other users and nodes within the cloud. Therefore, high security measures are required to protect data within the cloud. To protected the data in cloud by using fragmentation and replication. In DROPS methodology, divide a file into fragments, and replicate the fragmented data over the cloud nodes. Each of the nodes stores only a single fragment of a particular data file that ensures that even in case of a successful attack, no meaningful information is revealed to the attacker.***

Keywords: Centrality, cloud security, fragmentation, replication, performance.

1. Introduction

Cloud computing is characterized by on-demand ,self-services, network accesses, resource pooling, elasticity, and measured services. The aforementioned characteristics of cloud computing make it a striking candidate for businesses, organizations, and individual users for adoption. The benefits of low-cost, negligible management (from a users perspective), and greater flexibility come with increased security concerns is one of the most crucial aspects among those prohibiting the wide-spread adoption of cloud computing. The data outsourced to a public cloud must be secured. Unauthorized data access by other users and processes must be prevented. Any weak entity can put the whole cloud at risk. In such a scenario, the security mechanism must substantially increase an attacker's effort to retrieve a reasonable amount of data even after a successful intrusion in the cloud.

2. Literature Survey

A body of literature has been conducted by several authors and a list of them is given below;

1. Energy-Efficient Data Replication in Cloud Computing Datacenters .
Cloud computing is an emerging paradigm that provides computing resources as a service over a network. Communication resources often become a bottleneck in service provisioning for many cloud applications. Therefore, data replication, which brings data closer to data consumers is seen as a promising solution. It allows minimizing network delays and bandwidth usage. In this paper we study data replication in cloud computing data centers. Unlike other approaches available in the literature, we consider both energy efficiency and bandwidth consumption of the system, in addition to the improved Quality of Service as a result of the reduced communication delays. The evaluation results obtained during extensive simulations help to unveil performance and energy efficiency tradeoffs and guide the design of future data replication solutions.

2. Data Security Issues in Cloud Computing.
Cloud computing is an enticing technology which is a combination of many existing technologies such as parallel computing, grid computing, distributed computing and others. It offers services like data storage, computing power, shared resources at low cost to its users over internet at anytime from anywhere. Costing model on cloud computing is based on pay as you go method, hence companies are saving millions by adopting this technology. As more and more individuals and companies are relying on cloud for their data, the question arises here is how secure cloud environment Though cloud computing has many advantages, it also have some security problems.

3. On the Characterization of the Structural Robustness of Data center Networks.
A Data Center Network (DCN) constitutes the communicational backbone of a data center, ascertaining the performance boundaries for cloud infrastructure. The DCN needs to be robust to failures and uncertainties to deliver the required Quality of Service (QoS) level and satisfy Service Level Agreement (SLA). In this paper, analyze robustness of the state-of-the-art DCNs. Our major contributions are: (a) we present multi-layered graph modeling of various DCNs; (b) we study the classical robustness metrics considering various failure scenarios to perform a comparative analysis; (c) The present the inadequacy of the classical network robustness metrics to appropriately evaluate the DCN robustness; and (d) The propose new procedures to quantify the DCN robustnessCurrently, there is no detailed study available centering the DCN robustness. Therefore, we believe that this study will lay a firm foundation for the future DCN robustness research . Motivated by the question of access control in cloud storage, we consider the problem using Attribute-Based Encryption (ABE) in a setting where users' credentials may change and cipher may be stored by a third party.

4. Secure Overlay Cloud Storage with Access Control and Assured Deletion

This paper describes outsource data backups off-site to third-party cloud storage services so as to reduce data management costs. However, we must provide security guarantees for the outsourced data, which is now maintained by third parties. We design and implement FADE, a secure overlay cloud storage system that achieves fine-grained, policy-based access control and file assured deletion. It associates outsourced files with file access policies, and assuredly deletes files to make them unrecoverable to anyone upon revocations of file access policies. To achieve such security goals, FADE is built upon a set of cryptographic key operations that are self-maintained by a quorum of key managers that are independent of third-party clouds. In particular, FADE acts as an overlay system that works seamlessly atop today's cloud storage services. We implement a proof-of-concept prototype of FADE atop Amazon S3, one of today's cloud storage services. We conduct extensive empirical studies, and demonstrate that FADE provides security protection for outsourced data, while introducing only minimal performance and monetary cost overhead. Our work provides insights of how to incorporate value-added security features into today's cloud storage services.

5. Security and Privacy Issues in Cloud Computing Environment

Cloud computing is emerging as a powerful architecture to perform large-scale and complex computing. It extends the information technology (IT) capability by providing on-demand access to computer resources for dedicated use. The information security and privacy are the major concerns over the cloud from user perspective. This paper surveys and evaluates the architecture, data security and privacy issues in cloud computing like data confidentiality, integrity, authentication, trust, service level agreements and regulatory issues. The objective of this paper is to review comprehensively the current challenges of data security and privacy being faced by cloud computing and critically analyse these issues.

6. Dike: Virtualization-aware Access Control for Multitenant File systems

This paper describes in a virtualization environment that serves multiple customers (or tenants), storage consolidation at the file system level is desirable because it enables data sharing, administration efficiency, and performance optimization .The scalable deployment of file systems in such environments is challenging due to intermediate translation layers required for purposes of networked file access or identity management. Analyzes the security requirements in multitenant file systems. Then we introduce the Dike authorization architecture, which combines native access control with tenant namespace isolation that is backwards compatible to object-based file systems. We experimentally evaluate a prototype implementation that we developed, and show that our solution incurs limited added performance overhead.

7. Static and adaptive distributed data replication using genetic algorithms

Fast dissemination and access of information in large distributed systems, such as the Internet, has become a norm of our daily life. However, undesired long delays experienced by end-users, especially during the peak hours, continue to be a common problem. Replicating some of the objects at multiple sites is one possible solution in decreasing network traffic. The decision of what to replicate where, requires solving a constraint optimization problem which is NP-complete in general. Such problems are known to stretch the capacity of a Genetic Algorithm (GA) to its limits. Nevertheless, we propose a GA to solve the problem when the read/write demands remain static and experimentally prove the superior solution quality obtained compared to an intuitive greedy method. Unfortunately, the static GA approach involves high running time and may not be useful when read/write demands continuously change, as is the case with breaking news. To tackle such case we propose a hybrid GA that takes as input the current replica distribution and computes a new one using knowledge about the network attributes and the changes occurred. Evaluate these algorithms with respect to the storage capacity constraint of each site as well as variations in the popularity of objects, and also examine the trade-off between running time and solution quality.

8. Addressing cloud computing security issues.

The recent emergence of cloud computing has drastically altered everyone's perception of infrastructure architectures, software delivery and development models. Projecting as an evolutionary step, following the transition from mainframe computers to client/server deployment models, cloud computing encompasses elements from grid computing, utility computing and autonomic computing, into an innovative deployment architecture. From a security perspective, a number of unchartered risks and challenges have been introduced from this relocation to the clouds, deteriorating much of the effectiveness of traditional protection mechanisms. As a result the aim of this paper is twofold; firstly to evaluate cloud security by identifying unique security requirements and secondly to attempt to present a viable solution that eliminates these potential threats. This paper proposes introducing a Trusted Third Party, tasked with assuring specific security characteristics within a cloud environment.

9. Comparison and analysis of ten static heuristics-based Internet data replication techniques

Compares and analyses 10 heuristics to solve the fine-grained data replication problem over the Internet. In fine-grained replication, frequently accessed data objects (as opposed to the entire website contents) are replicated onto a set of selected sites so as to minimize the average access time perceived by the end users. The paper presents a unified cost model that captures the minimization of the total object transfer cost in the system, which in turn leads to effective utilization of storage space, replica consistency, fault-tolerance, and load-balancing. The set of heuristics include six A-Star based algorithms, two bin packing algorithms, one greedy and one genetic algorithm. The heuristics are extensively simulated and compared using an experimental test-bed that closely mimics the Internet

infrastructure and user access patterns. GTITM and Inlet topology generators are used to obtain 80 well-defined network topologies based on flat, link distance, power-law and hierarchical transit–stub models. The user access patterns are derived from real access logs collected at the websites of Soccer World Cup 1998 and NASA Kennedy Space Centre. The heuristics are evaluated by analysing the communication cost incurred due to object transfers under the variance of server capacity, object size, read access, write access, number of objects and sites. The main benefit of this study is to facilitate readers with the choice of algorithms that guarantee fast or optimal or both types of solutions.

10. Enhanced dynamic credential generation scheme for protection of user identity in mobile-cloud computing.
In this paper, to improve the resource limitation of mobile devices, mobile users may utilize cloud-computational and storage services. Although the utilization of the cloud services improves the processing and storage capacity of mobile devices, the migration of confidential information on untrusted cloud raises security and privacy issues. Considering the security of mobile-cloud-computing subscribers' information, a mechanism to authenticate legitimate mobile users in the cloud environment is sought. Usually, the mobile users are authenticated in the cloud environment through digital credential methods, such as password. Once the users' credential information theft occurs, the adversary can use the hacked information for impersonating the mobile user later on. The alarming situation is that the mobile user is unaware about adversary's malicious activities. In this paper, a light-weight security scheme is proposed for mobile user in cloud environment to protect the mobile user's identity with dynamic credentials. The proposed scheme offloads the frequently occurring dynamic credential generation operations on a trusted entity to keep minimum processing burden on the mobile device. To enhance the security and reliability of the scheme, the credential information is updated frequently on the basis of mobile-cloud packets exchange.

3. Conclusion and Future Enhancement

The DROPS methodology, a user has to download the file, update the contents, and upload it again. It is strategic to develop an automatic update mechanism that can identify and update the required fragments only. The aforesaid future work will save the time and resources utilized in downloading, updating, and uploading the file again. Moreover, the implications of TCP in cast over the DROPS methodology need to be studied that is relevant to distributed data storage and access.

4. Acknowledgement

I would like to take this opportunity to express my profound gratitude and deep regard to my guide, *Prof R.RUHIN ROUSER Kingston Engineering College*, for his exemplary guidance, valuable feedback and constant encouragement in completing this paper. His valuable suggestions were of immense help in getting this work done. Working under his, was an extremely knowledgeable experience. Also, I would like to extend my sincere gratitude to my parents for their constant support and encouragement in completing this paper.

References

[1] K. Bilal, M. Manzano, S. U. Khan, E. Calle, K. Li, and A. Zomaya, "On the characterization of the structural robustness of data center networks," IEEE Transactions on Cloud Computing, Vol. 1, No. 1, 2013.
[2] Y. Deswarte, L. Blain, and J-C. Fabre, "Intrusion tolerance in distributed computing systems," In Proceedings of IEEE Computer Society Symposium on Research in Security and Privacy, Oakland CA,110-121, 1991
[3] B. Grobauer, T.Walloschek, and E. Stocker, "Understanding cloud computing vulnerabilities," IEEE Security and Privacy, Vol.9, No. 2, 2011.
[4] W. K. Hale, "Frequency assignment: Theory and applications," Proceedings of the IEEE, Vol. 68, No. 12, 1980,
[5] K. Hashizume, D. G. Rosado, E. Fernndez-Medina, and E. B. Fernandez, "An analysis of security issues for cloud computing," Journal of Internet Services and Applications, Vol. 4, No. 1,2013,
[6] W. A. Jansen, "Cloud hooks: Security and privacy issues in cloud computing," In 44th Hawaii IEEE International Conference onSystem Sciences (HICSS), 2011,
[7] A. Juels and A. Opera, "New approaches to security and availability for cloud data," Communications of the ACM, Vol. 56, No. 2, 2013.
[8] L. M. Kaufman, "Data security in the world of cloud computing," IEEE Security and Privacy, Vol. 7, No. 4, 2009, pp. 61-64.
[9] L. Qiu, V. N. Padmanabhan, and G. M. Voelker, "On the placement of web server replicas," In Proceedings of INFOCOM 2001, Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies, Vol. 3, pp. 1587-1596, 2001.
[10] D. Sun, G. Chang, L. Sun, and X. Wang, "Surveying and analyzing security, privacy and trust issues in cloud computing environments," Procedia Engineering, Vol. 15, 2011, pp. 2852 -2856.
[11] J. J. Wylie, M. Bakkaloglu, V. Pandurangan, M. W. Bigrigg, S. Oguz, K. Tew, C. Williams, G. R. Ganger, and P. K. Khosla, "Selecting the right data distribution scheme for a survivable storage system," Carnegie Mellon University, Technical Report CMU-CS-01-120, May 2001.
[12] A. R. Khan, M. Othman, S. A. Madani, S. U. Khan, "A survey of mobile cloud computing application models," IEEE Communications Surveys and Tutorials, DOI: 10.1109/SURV.2013.06261

Author Profile

M. Kanchana is a Post-graduate student in the Computer Science Department, Kingston Engineering College, and Vellore, India. She received M.S degree in 2012 from Sathyabama University, Chennai, India.

Her research interests are Cloud Computing, data mining and data structure.

R.Ruhinkouser Assistant Professor, Department of Computer Science, Kingston Engineering College. She received her B.E degree from Bharathidasan Engineering College. She then completed her M.E degree at Sapthagiri Engineering college . Her area of interest includes Data mining, Computer Networks Security, Wireless Networks, software testing, Network simulation, Sensor Networks, Data Structure and Cloud Computing. She has published the 6 book and has attended 6 workshop and 6 conferences.