Capacity Management for Virtualized Data Centers using ECIES and Scheduling

Anil Kumar Raghuwanshi
Department of Computer Science
OCT College
Bhopal, India

Kavita Burse, Ph.D.
OCT College
Bhopal, India

ABSTRACT
With the growth Internet development cloud computing is novel technique to serve better and secure services. E-business is growing rapidly with the development of Internet. The cloud computing provides on demand self service methodology that authorizes users to request resources dynamically as a best benefit. The use of Cloud Computing is ahead reputation due to its mobility and massive availability in minimum cost. Here in this paper an efficient Capacity Management of user's data on datacenters is proposed using attribute and scheduling techniques. The proposed technique provides much efficient use of Virtualized data as compared to the existing technique.

Keywords
Cloud Computing, Public Verifiability, Cloud Storage, Cloud Security, Virtualization.

1. INTRODUCTION
In the modern era e-business is mostly conducted over the internet thereby with the help of emerging technology named as Cloud computing data storage, platform, and various IT services are provided over internet. Cloud computing deployed efficiently is capable of managing multiple data centers. With the help of virtual machines (VM's) inside the data centers diverse workload can be performed but it also causes high and low utilization of resources which requires enhanced virtualization mechanism technique for solving the related problem [1]. Cloud Computing enables the enterprises to offer location sovereign resource pooling, rapid resource elasticity, ubiquitous network access, usage-based pricing, transference of risk, on-demand self-service etc. [2].

There are some services associated with Cloud computing:

i. **Infrastructure as a service (IaaS):** in this service vendor's of cloud computing share their dedicated resources with clients on pay per use payment. IaaS is basically a single tenant cloud layer.

ii. **Software as a service (SaaS):** works upon virtualized and pay-per-use costing model. SaaS vendors leases the software applications to contracted organizations.

iii. **Platform as a service (PaaS):** is build upon IaaS. With the help of PaaS clients can access basic operating software. It provides optional services like database access and payment services for using and deploying the software applications. Thereby removing the requirement of purchasing and managing the computing infrastructure. There are various cloud models apart from the services provided by cloud computing:

i. **Public Cloud:** With the help of mainstream web browsers users can access cloud through interfaces.

ii. **Private Cloud:** is usually within an **organization's** internal enterprise datacenter. The setup is deployed inside.

iii. **Hybrid Cloud:** as the name indicates it is a combination of more than one cloud.

iv. **Community Cloud:** This cloud is used by a specific community. The community is of consumers who are from organizations that consist of shared apprehensions [2].

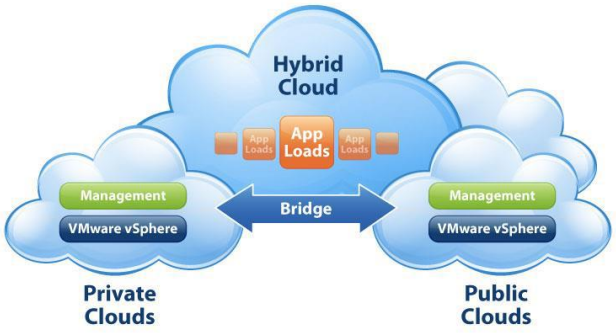

Figure 1: Types Of Clouds In Cloud Computing [3]

Virtualization is capable of protecting the attack by users which are directed to one another or on cloud infrastructure. Virtualization software enables the use of multiple operating systems and applications simultaneously in a computer [3] and is also able to keep local and global resources hidden that are attached with the system. The server, workstation storage etc. does not depend physical hardware layer due to virtualization. Virtualization also reduces number of physical system availability and multiple system environments can be used inside a single system. Now focusing on cloud security which implies the security of data, networking components and security of other resources from viruses, hackers, intruders, worms etc. The security issues are of cloud service providers and of the users who use it. Cloud Security is based upon some parameters that are: authentication, , personal and physical security, privacy of the content and the user availability and application security.

Cloud computing usually used for the on demand services that permit users to access the demanded resources dynamically. This is the best benefits of the cloud computing. Cloud computing server handles the security of data when it is stored upon remote server and cloud. Data security is maintained in publicly auditable cloud, the providers of cloud storage provides trusted third party auditor (TPA) which verifies data integrity thus ensuring security of the data [4].

The service quality is monitored on objective and autonomous view by TPA. The integrity verification tasks of clients are assigned to TPA by public audit capability. This is done when clients cannot perform computation resources that perform continuous verifications [5]. But to be considered individual auditing of TPA also is tiresome and lack in efficiency [6]. In cloud environment the clients are sometimes unreliable and the overhead of performing frequent truthfulness verification is unaffordable for the clients. Considering the realistic use of the technique verification protocol with public verifiability equipment is balanced in nature thereby through public verifiability economic scale for Cloud Computing is achieved [7].

Verifiability can also authenticate and validate the client and can even validate TPA. It is categorized as private verifiability and public verifiability. Public Verifiability is defined as a technique in which user is forcefully denied to upload private data in the cloud storage and it is also ensured that the data that is to be stored should correctly be stored. Private Verifiability as compared with public verifiability is more secure. It is generally used for higher efficiency [8].

Cloud data storage is affected when cloud data storage provider is untrusted and is malicious. The identity and data of the user in the cloud is essential but there is also a need to preserve its privacy. The rest of paper is organized as follows. In Section II describes about background of cloud environment. Section III describes about related work in public verifiability, cloud security, virtualization. Section IV describes about proposed method. Section V describes about experimental result algorithm followed by a conclusion in Section VI.

2. BACKGROUND

Online service on internet provides large storage space and computing resources that are customizable like computing stage shift. But this has also removed local machines liability which maintains the data at the same time. Hence the security of data is essential to maintain service quality as users are dependent on cloud service providers. Virtualization can hide physical resources which are used in cloud computing environment. For storing valid data it depends upon the factors involving user and the type of data being used. In social networks users also upload private data and share the data with other users. Thus the privacy of the data needs to be preserved every time.

3. RELATED WORK

This section describes about related work in fields of cloud storage, cloud security, virtualization and public verifiability.

Cloud capacity management (CCM) consists of multiple low overhead techniques. CCM operates on practical on field observations achieving scalability allocation. The architecture of CCM consists of levels which are top level cloud manager, mid level super cluster managers and cluster managers at the lowest level. Clusters formed by logically grouped hosts are basically at bottom level and are bounded tightly to the capacity manager and the corresponding capacity manager monitor the clusters formed. Cloud level capacity manager is a collection of super clusters under which the other clusters work. Capacity manager monitors black box VM CPU thereby aggregating it and analyzing the usage information of memory. With the help of black box monitoring and allocation, CCM perform capacity allocation for a broad class of applications. The management cost is reduced by monitoring and changing the resources in intervals that are not so frequent while moving up on the hierarchy. CCM generally analyzes computing estimated demand of a cluster and super cluster respectively at the super cluster and cloud level [1].

Cloud storage providers that are auditable in public have data owners who look upon the third party auditor for verification of data integrity of the data which is obtained from a source for ensuring the security. They adopted a homomorphic authenticator technique which provides public audit ability without burdening the data owner. Homomorphic authenticators are extraordinary metadata which is obtained from individual data blocks and securely aggregating and providing guarantee to the verifier informing that linear combination of data blocks computation is proper by verifying aggregated authenticator. The linear combination is masked with randomness obtained from the server. The combination is obtained from the sampled blocks in response from the the server [4].

They analyzed and resolve the problem of providing the ability to simultaneously audit public and data dynamics that remotely check the data integrity in Cloud Computing. They [5] offered a protocol that supports fully dynamic data operations and support block insertion. With the help of cloud large data files can be stored on remote servers and the clients can be freed from the storage concern, calculation and the problems. There are some concerns associated to clients like assurance of correct storage of data and its maintenance. If the local copies are absent client should be able to verify the remote data and its correctness with the help of a security measure and the clients should also be able to interact with cloud servers for accessing and retrieving pre-stored data. The client performs block level operations on data files multiple times therefore for supporting public auditability efficiently and not allowing the retrieval of the data blocks by themselves defined by and explained through homomorphic authenticator technique. During the process of verification block less approach and authenticating the block tags is done and the original blocks are not considered. For the block tag confirmation they manipulated classic Merkle Hash Tree construction for achieving efficient data dynamics and improving the existing proof of storage models. They analyzed bilinear aggregate signature technique and presented their result up to a multiuser setting for supporting competent handling of multiple auditing tasks and TPA concurrently performing multiple auditing tasks [5].

D. Srinivas proposed that in the duration of auditing process efficiency a guarantee is provided of TPA not gathering information about the data which is stored on the cloud server. This is done through homomorphic non linear authenticator and random masking this thereby reduces user's burden of auditing task which is pricey and tedious providing the user the security of his data which is outsourced. Privacy preserving public auditing protocol is extended to a multi-user scenario in which TPA is able to execute multiple auditing tasks in batch manner providing high efficiency and security [6].

Zhu, Yan et al [7] focused and proposed PDP scheme for hybrid clouds which provides privacy protection and dynamic scalability constructing Cooperative Provable Data Possession (CPDP) using Homomorphic Verifiable Responses (HVR) and Hash Index Hierarchy (HIH). The mechanism thereby convinces the clients of data possession without making them know about machines and the geographical locations of the residence of their files. They proposed a new hash index hierarchy for the clients that allows easy storage and management of the resources in hybrid clouds. The results given by them also validate the effective approach of their construction [7].

For acceptance of integration requirements from clients to share data and maintain its privacy in the cloud a privacy preserving repository is proposed. In the scheme data is collected and integrated from data sharing services giving the integration results to users. The query plan wrapper scrutinizes integration requirements and constructs query plans which are then used by query plan executor [9].

Architecture of cloud database storage is used to avoid the local administrator. Additionally Ulrich Greveler et al [10] studied about the cloud administrator and outsourced database content. The cloud data should be protected from external attackers as well as internal attacks also. This system follows the information-centric scheme to make cloud data self-intelligent. According to the scheme cloud data are encrypted and packaged with a usage policy. After using the policy data should be created, virtualized and virtualization and try to review the trustworthiness of the data environment [10]. Data integrity and verifiability that is remotely accessed performs dynamic operations on data and such new method is recommended by Qian Wang et al [11]. This scheme is competent enough to identifies the troubles and potential security problems of the well known extensions and it is fully dynamic data updates. This scheme gain efficient data dynamics. It is also improves the retrieve ability by manipulating the classic Merkle Hash Tree (MHT) construction i.e. used for block tag validation [11].

As far as data sharing serices are concerned with respect to data warehouses they can capable to update and control the access and limit the usage of their shared data. They are like substitute of submitting data to establishment. Their [12] repository was supported data sharing and addition. Data warehouse model cannot be used to generate other results except that of the specified data addition request. Due to that the cooperation of warehouse can only reveal the results of the specified data integration demand. The cooperation of central establishment will expose all data and recommended a privacy preserving storage area to incorporate data from numerous data distribution services. This repository or data storage collects data from data sharing services. These services are based on users' integration requirements rather than all the data from the data sharing services as existing central establishment. The central establishment (existing) has full control of the collected data; the warehouse capability is controlled to computing the integration results demanded by users and cannot get other information about the data or use it for other work [12].

Michael Armbrust et al [13] worked to discovering top technical and non-technical obstacles and various opportunities of cloud computing. Virtualization is one of them. It is primary security mechanism of cloud computing with powerful defense method. Multiple virtual machines (VMs) are used to share CPUs and main memory surprisingly well in cloud computing. Virtualization is essential component to enhance architectures and operating systems to professionally virtualized interrupts and I/O channels. It has been known to contain bugs allowed by virtualized code to break loose" to some level. Incorrect virtualization may lead the access of sensitive portions of the provider's infrastructure. Sometimes the resources of other users multiple virtual machines (VMs) also share CPUs and main memory unexpectedly glowing in cloud computing [13].

A system with the black-box and gray-box information from individual VMs to detect and alleviate resource hotspots via VM migrations was suggested by Wood et al [14]. Preferably most resources are having intensive parts of a host-move operation due to the need to evict all VMs currently running on the host. Due to this reason the duration of the operation is administrated by factors to facilitate VM size and active memory dirtying rate [15]. Effective management is gained without demanding expensive or costly fine-grained monitoring of workload VMs at large scales.

4. PROPOSED ALGORITHM

The proposed methodology works in the following phases:

1. Provides authentication when ever any new user is send data to datacenters at broker.

2. Each time virtualized data center is created a dictionary entry for scheduling of these data centers at broker.

3. Data is always access using the concept of query based encryption and decryption.

Annotations Used

Table 1. Annotations used in the methodology

Ui	User of cloud
Brki	Local broker of the cloud
DCi	Datacenter of the cloud
Pkt	Packet or data to be send
Ci	Capacity of the storage
Ri	Resource of the cloud
Atti	Attribute for each Pkt

1. If 'N' of packets are send from Ui to DCi.

   Ui (N) {Pkt} → DCi

2. For each Pkt send from Ui → Brki
3. Generate a Atti and encrypt the Pkt using Enc (Pkt).
4. Send the Tupple (Enc (Pkt), Atti) to the local Broker Brki.
5. Scheduling of this Data Pkt is done at the local broker for the access of the resource Ri.
6. Create virtal data centers at the time of request of the Pkt to access.
7. The receiver needs to authenticate at the local broker Brki and Dec (Enc (Pkt)).

5. RESULT ANALYSIS

The table shown below is the analysis and comparison of the communication overhead during the transmission from cloudlet user to the local broker.

Table 2. Comparison of the comparison Overhead

	Latency		
Schemes	**Cloudlet**	**Broker**	**Communication Overhead**
Existing	n	1	nS+nlognH
Proposed	logn	1	((n-1)+lognH))S

The figure shown below is the analysis of communication overhead for the existing and the proposed work. The proposed methodology provide less overhead as compared to the existing work.

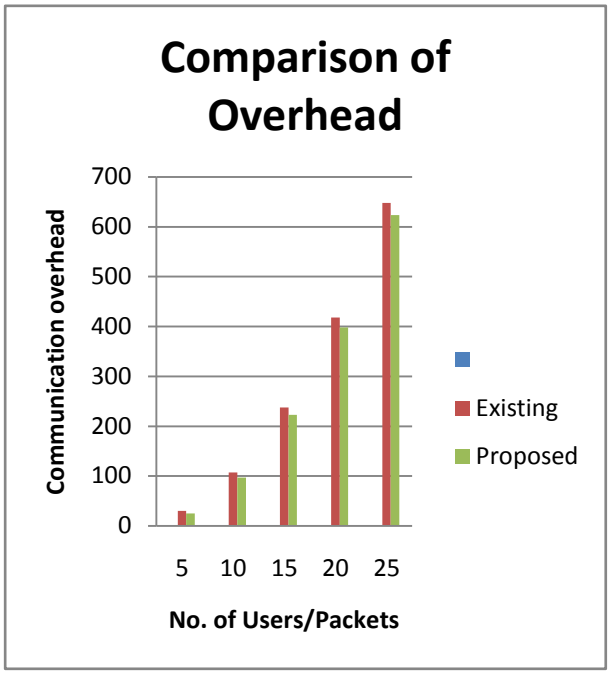

Figure 1. Comparison of Overhead

The table shows below is the varios attack parameters that the proposed methodology is able to prevent.

Table 3. Prevention from various Attacks

Security Parameter	**Prevented by proposed technique**
Public verifiability	YES
Password impersonation	YES
Insider attack	YES
Outsider attack	YES
Password guessing attack	YES
Denning sacho attack	YES

The figure shown below is the comparison of the utilization of the CPU between the existing and the proposed work.

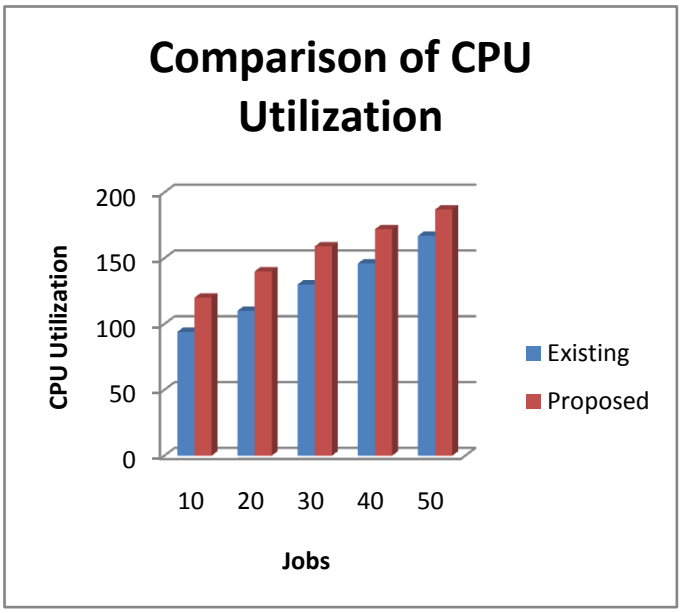

Figure 2. Comparison of CPU Utilization

The table shown below is the comparison of access of various application in cloud environemnt and their respective performance metrics.

Table 4. Comparison of Application Performance

Applications	**Existing**	**Proposed**
Nutch	106	87
Linpack	468	526
Voldemort	9.8	9.2
Olio	274	382

6. CONCLUSION & FUTURE WORK

The proposed technique implemented here for the Management of Datacenters so that the capacity of the users can be increased. The result analysis shows the performance of the proposed methodology. Also the methodology provides less storage and less communication overhead as compared to the existing technique. Although the technique is efficient in terms of CPU Utilization and Performance but further enhancement is required in the enhancement of the methodology of virtualization of datacenters.

Although the technique implemented here provides low storage cost and low communication overhead and provides better capacity management but further enhancements can be done in the field of applying better scheduling and also clustering is done at the central Authority to improve the performance of the methodology.

7. REFERENCES

[1] Kesavan, Mukil, Irfan Ahmad, Orran Krieger, Ravi Soundararajan, Ada Gavrilovska, and Karsten Schwan "Practical Compute Capacity Management for Virtualized Datacenters", IEEE Transactions On Cloud Computing, Vol. 1, No. 1, pp. 88 – 100, 2013.

[2] P. Mell and T. Grance, "The NIST definition of cloud Computing", online available at: http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf, 2012.

[3] Swathi, T., K. Srikanth, and S. Raghunath Reddy. "Virtualization In Cloud Computing." International Journal of Computer Science and Mobile Computing, ISSN 2320–088X Vol.3 Issue.5, pp. 540-546, May- 2014 .

[4] Swathi Sambangi "Cloud Data Storage Services Considering Public Audit for Security", Global Journal of Computer Science and Technology Cloud and Distributed, ISSN: 0975-4172, Vol. 13, Issue 1, pp. 1 – 6, 2013.

[5] Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing," IEEE Trans. Parallel and Distributed Systems, vol. 22, no. 5, pp. 847-859, May 2011.

[6] Srinivas, D. "Privacy-Preserving Public Auditing In Cloud Storage Security." International Journal of computer science and Information Technologies,ISSN: 0975-9646, vol. 2, no. 6, pp. 2691-2693, 2011.

[7] Zhu, Yan, Huaixi Wang, Zexing Hu, Gail-Joon Ahn, Hongxin Hu, and Stephen S. Yau. "Efficient provable data possession for hybrid clouds." In Proceedings of the 17th ACM conference on Computer and communications security, pp. 756-758. ACM, 2010.

[8] Qian Wang, Cong Wang, Jin Li1, Kui Ren, and Wenjing Lou, "Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing" Proceedings of the 14th European conference on Research in computer security (ESORICS'09), pp. 355-370, 2009.

[9] Mishra, Ranjita, Sanjit Kumar Dash, Debi Prasad Mishra, and Animesh Tripathy. "A privacy preserving repository for securing data across the cloud." In IEEE 3rd International Conference on Electronics Computer Technology (ICECT-2011), vol. 5, pp. 6-10, 2011.

[10] Greveler, Ulrich, Benjamin Justus, and Dennis Loehr. "A Privacy Preserving System for Cloud Computing." In IEEE 11th International Conference on Computer and Information Technology (CIT- 2011), pp. 648-653, 2011.

[11] Qian Wang, Cong Wang, Jin Li1, Kui Ren, and Wenjing Lou, "Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing" Proceedings of the 14th European conference on Research in computer security(ESORICS'09), pp. 355-370, 2009.

[12] Stephen S. Yau, Fellow And Yin Yin "A Privacy Preserving Repository For Data Integration Across Data Sharing Services", IEEE Transactions On Services Computing, Vol. 1, No. 3, July-September 2008.

[13] Armbrust, Michael, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee et al. "A view of cloud computing." Communications of the ACM, vol. 53, no. 4, pp. 50-58, 2010.

[14] T. Wood et al., "Black-Box and Gray-Box Strategies for Virtual Machine Migration," Proceedings of Fourth USENIX Conf. Networked Systems Design and Implementation (NSDI '07), pp. 17-17, 2007.

[15] Clark, Christopher, Keir Fraser, Steven Hand, Jacob Gorm Hansen, Eric Jul, Christian Limpach, Ian Pratt, and Andrew Warfield "Live migration of virtual machines", In Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation,Vol. 2, pp. 273-286, 2005.

[16] Liu, Haikun, Hai Jin, Cheng-Zhong Xu, and Xiaofei Liao "Performance and energy modeling for live migration of virtual machines", Cluster computing, vol. 16, no. 2, pp. 249-264, 2013.