ISSN: 0976-3104
SPECIAL ISSUE: Computer Science

ARTICLE

SECURE DATA STORAGE AND SHARING IN CLOUD: VM SCHEDULING

R.G. Babukarthik[1*], J. Satheesh Kumar[2], J. Amudhavel[3]

[1]*Part time Research scholar (Category-B) R&D, Bharathiar University, Coimbatore, Tamil Nadu, INDIA*
[2]*Department of Computer Application, Bharathiar University, Coimbatore, Tamil Nadu, INDIA*
[3]*Department of Computer Science and Engineering, KL University, Andhra Pradesh, INDIA*

ABSTRACT

Performance of cloud storage is comparatively high compare to other storage devices in terms of consistency, efficiency improvement and execution time. Based on workflow application VM resource type is chosen dynamically in the cloud. A pure isolation hypervisor simply divides a machine into partitions, and permits sharing of resources between the partitions, a load balancing scheme is based on dynamic resource allocation policy for virtual machine cluster, running under para virtualization mode on a cluster of physical machines (PM) in shared storage architecture. In this paper we proposed a mechanism that makes possible for the data owners to enforcement their security policies to ensure data confidentiality and integrity, which enable trusted data sharing through untrusted cloud providers. The computation time for PSO technique is calculated and is compared with various scheduling algorithm such as Round Robin (RR), Random, Heros, Green, Randens and Bestdens.

KEY WORDS
VM Virtual Machine, PM Physical Machine, PSO Particle Swarm Optimization, RR Round Robin

Received: 3 June 2017
Accepted: 20 July 2017
Published: 12 Sept 2017

INTRODUCTION

Workflow applications evolved as an eye-catching paradigm for programming in distributed computing, it is widely used in bioinformatics, scientific computing and physics [1]. Workflow application emerged as a huge data application, due to raising problems of scientific computing systems, leading to vast infrastructures for executing the application within a stipulated time [2]. Thus cloud computing infrastructures pays a special need [3], distributed computing focusing on efficient and cost-effective system operation. Cloud infrastructure is provided on the basis of pay-per-use, facilitating dynamic scaling for large workflow applications. Cloud storage performs well compare to other storage in terms of consistency, efficiency improvement and execution time [4]. Based on workflow application VM resources and its type are chosen dynamically in the cloud. Even though huge amount of resources are available in the cloud, users need to focus on economical cost and its policy [5]. Cost is calculated using time unit model and not on the basis of usage of resources [6, 7]. Generally task of all real world applications differs a lot, exhibiting heterogeneous behaviors (Memory-intensive, data-intensive and computation-intensive) leading to choose an unique VM types [8,9] example E2C of Amazon delivers various instance of VM, such as computing optimized, storage optimized and memory optimized thereby rendering various VM types [10]. Security arise a major concern for various application on cluster [11, 12, 13] heterogeneous distributed computing systems [14, 15], grid computing [16] and cloud computing [17]. International Data Corporation states that cloud security arise as a major concern [18]. This is mainly due to huge amount of user is allowed to execute various untested third party applications, thus applications and users need to be scrutinized [19]. Still many cloud computing environments yet to apply security to tackle the security threats [20]. Hence there is an emerging need to implement security and privacy in cloud data center.

There are very few studies regarding the wearing and laundering of lab coats in hospitals and medical practice. This study highlights the role of lab coats acting as vector for transmitting health care infections to the patients and the common areas where contamination occurs.

RELATED WORKS

Task scheduling is a basic functional unit of cloud architecture from security aspect, it consider various security requirements mainly due to sensitivity of the data such as medical image analysis, image storage and stock photo service. Very high security level is needed in medical image analysis, where as in traditional scheduling consider only makespan and minimizing energy consumption for optimized load balancing [21, 22]. As a result complexity of problems increases gradually; a model encompassing various approaches is stated for security based scheduling [23]. Cloud security services is grouped into three types, application service in the application layer, secure process enabling service (infrastructure layer) and secure physical service (physical layer) based on tools ensuring security (security as a service) [24].

Hypervisor which reduce the overhead of virtualization and provide security: There are two classes of hypervisors that must be considered when examining the technical implications of MLS for hypervisors; they are pure isolation hypervisors and sharing hypervisors. A pure isolation hypervisor simply divides a machine into partitions and permits no sharing of resources between the partitions (other than CPU time and primary memory). Implementing a pure isolation hypervisor is very easy, because the only security policy to be enforced is isolation. IBM's EAL5- evaluated PR/SM system [25] for the z/Series mainframes and it is a good example of a pure isolation hypervisor. The idea of a secure sharing hypervisor originated with Madnick and Donovan [26]. The best examples of such secure sharing hypervisors are KVM/370 [27] and Digital's A1-secure VMM [28, 29, 30]. The most critical feature of a secure sharing hypervisor is a secure shared file store11. The secure shared file store allows a high level partition to have read-only

***Corresponding Author**
Email:
r.g.babukarthik@gmail.com
Tel.: +91-9043108042

IIOAB JOURNAL
ISSN: 0976-3104

access to low-level data, while a low-level partition gets read-write access to the same data. This avoids the clumsy one-way networking approaches only a single copy of the data is required and updates are visible immediately to all partitions [31, 32, 33].

Integrity Using a Virtual Machine Verifier to assure the integrity during the Virtual Machine Mobility: Distributed computing architectures, such as grid and cloud computing, depend on the high integrity execution of each system in the computation [34, 35]. While integrity measurement enables systems to generate proofs of their integrity to remote parties, current integrity measurement approaches are insufficient to prove runtime integrity for systems in these architectures. Integrity measurement approaches that are flexible enough have an incomplete view of runtime integrity, possibly leading to false integrity claims and approaches that provide comprehensive integrity is used only for computing environments that is too restrictive. Proposed architecture is used to build comprehensive runtime integrity proofs for general purpose systems in distributed computing architectures [36, 37]. In this architecture, they strive for classical integrity, using an approximation of the Clark-Wilson integrity model as our target. Key for building such integrity proofs is a carefully crafted host system whose long-term integrity is justified easily using current techniques and few new component called VM verifier, which comprehensively enforces our integrity target on VMs. Building a prototype based on the Xen virtual machine system for SELinux VMs and to find the distributed compilation is implemented, thereby providing accurate proofs of our integrity target with less than 4% overhead [30].

PROPOSED WORK

Towards secure data storage and sharing in the cloud: End-users of the cloud store their data in the provider's infrastructure; a critical concern is the security and privacy of these data. End-users want to know where their data is stored and who has control of the data in addition to the owners. They also want to be protected against illegal access to the data by the cloud provider, or other third parties. Secure access and storage of data in the cloud is addressed through the following tasks.

Data leakage prevention and privacy with 3rd party service providers: Shifting data storage to off-premises providers has two consequences: First, data owners have only limited control over the data stored in the cloud. Second, cloud providers have excessive privileges, giving them extensive control over the cloud user's data. Combining, this leads to a low level of trust between the end-user and the cloud provider with respect to keeping and sharing business critical data in the cloud. Mechanisms that make it possible for the data owners to enforcement their security policies to ensure data confidentiality and integrity, mechanisms that enable trusted data sharing through untrusted cloud providers.

Information source authentication: Algorithms that guarantee the authenticity of data stored in the cloud. This provide authentication and trust in the acquired information to avoid situations where the user's data may be altered without the owner's consent.

Methods to enable free inter cloud data movement: End-users buy services offered by the cloud providers without knowing where the cloud resources are located. The location might be beyond the borders of a legislative entity and can cause problems when disputes happen, which might be beyond the control of cloud provider. Furthermore, entrusting significant amounts of data to a cloud provider creates a risk of data lock-in. Technical solutions that ease the implementation of free inter- cloud data movement and a policy specification to standardize that process between providers.

Policies for data retention: Data retention is defined as storing recorded data for a period of time that is longer than the time necessary to perform the tasks; this remains the reason for recording the data. E.g. Amazon stores the list of all previous purchases for each individual even though it is not necessary for practical or legal reasons (order completion, accounting etc.). Currently, data retention is usually regulated in the terms of a service agreement between the provider and the user. Due to the complexity and frequent changes of such agreements they are usually not read by the user before they are accepted, developed policies for data retention and technical tools for enforcing these policies.

End results and significance: The security solutions increase the trust between cloud customers and cloud providers, thereby increasing the security of the cloud services and infrastructure. Moreover, sensitive data can be securely stored, shared and processed in the cloud. This allows businesses to reap the full benefits of the cloud, and a business critical decision is made with all relevant data available.

PSO VM scheduling algorithm is used for optimal scheduling not only focusing on makespan and energy efficient scheduling but also on the security and privacy principles. The number of servers is given as input keeping the fixed number of user and specific topology. The first step starts with the initialization of server, user, topology and data center components. The second step is setting of parameters such as power models, core switch and aggregate switch. The execution of task is taken place using PSO scheduling algorithm and with other scheduling algorithm the performance of proposed scheduling algorithm is compared. The computation time and energy is calculated.

IIOAB JOURNAL
ISSN: 0976-3104

PSO_VMscheduling_Algorithm
Input: No. of servers, topology, user.
Output: Energy, computation time, memory.
Step 1: Initialization
Server, topology, user, data center components.
Step 2: Parameters Setting
Non-linear power model, core switch, aggregation switch, task.
Step 3: Execution
Scheduler
PSO, Round Robin, Random, Heros,
RandDENS, BestDENS
Topology
Three-tier debug
Energy model
eDVFS, DNS.
Step 4: Report
Data center load
Individual server load
Individual VM load
Load of individual links
Step 5: Display
Total energy consumed
Energy of servers

EXPERIMENTAL EVALUATION AND ANALYSIS

The experiment is carried out using the green cloud simulation tool, the input is given as number of server, for a given fixed number of user and topology of the network, the computation time for PSO technique is calculated and is compared with various scheduling algorithm such as Round Robin (RR), Random, Heros, Green, Randens and Bestdens. It is clear that computation time for green scheduling is minimum compared to the all other scheduling algorithm at minimum number of server (tasks), whereas as the number of server increases PSO really out performed well. The [Table 1] shows the computation time for various scheduling algorithm. [Fig. 1] show the comparison of the computation time for various scheduling algorithm and thus PSO really out performed well, underlying below the all scheduling algorithm with minimum computation time.

Table 1: Computation time comparison of various scheduling algorithm

Sl. No.	Server	RR	Random	Heros	Green	Randens	Bestdens	PSO
1	30	0.5	0.41	1	0.36	0.81	0.89	0.42
2	60	1	1	2	1	1	2	1
3	90	2	2	4	2	2	3	2
4	120	3	3	7	4	3	5	3
5	150	4	4	10	5	4	8	4
6	180	4.5	5	13	6	5	10	4.4
7	210	5	5.7	17	8	6	13	5
8	240	6	6	22	9	8	16	6
9	270	7	7	28	11	9	20	7
10	300	8	9	35	13	10	23	8

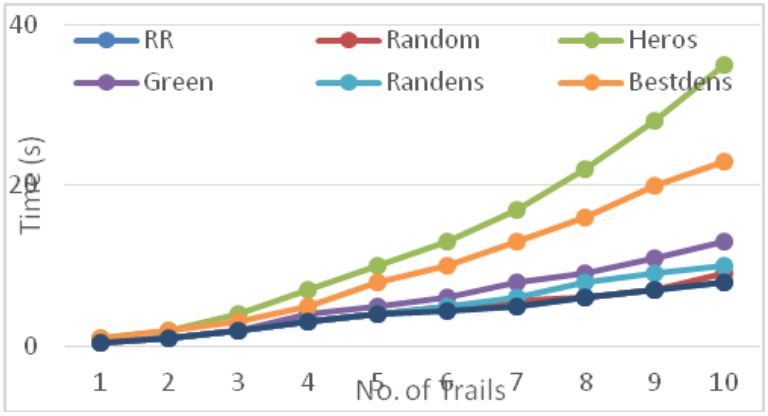

Fig. 1: comparison of scheduling algorithm.

CONCLUSION

Cloud infrastructure is provided on the basis of pay-per-use, facilitating dynamic scaling for large workflow applications. Secure access and storage of data in the cloud is addressed by Data leakage prevention and privacy with 3rd party service providers, Information source authentication, various methods to enable free inter cloud data movement, Policies for data retention and End results significance. The computation time for PSO technique is calculated and is compared with various scheduling algorithm, green scheduling has minimum computation time compared to the all other scheduling algorithm with minimum number of server (tasks), whereas as the number of server increases PSO really out performed well.

CONFLICT OF INTEREST
There is no conflict of interest.

ACKNOWLEDGEMENTS
None

FINANCIAL DISCLOSURE
None

REFERENCES

[1] Juve G, Chervenak A, Deelman E, Bharathi S, Mehta G, Vahi K.[2013] Characterizing and profiling scientific workflows, Future Generation Computer System 29(3): 682 - 692.
[2] Kashlev A, Lu SY.[2014] A system architecture for running big data workflows in the cloud, 2014 IEEE International Conference on Services Computing, SCC pp. 51-58.
[3] Buyya R, Yeo CS, Venugopal S, Broberg J, Brandic I,[2009] Cloud computing and emerging it platforms: vision, hype, and reality for delivering computing as the 5th utility ,Future Generation Computer System 25(6): 599 -616.
[4] Chang V, Wills G.[2015] A model to compare cloud and non-cloud storage of Big Data, Future Generation Computer Systems.
[5] Calheiros RN, Buyya R.[2014] Meeting deadlines of scientific workflows in public clouds with tasks replication, IEEE Transactions on Parallel and Distributed Systems 25(7) :1787-1796.
[6] Rodriguez MA, Buyya R.[2014] Deadline based resource provisioning and scheduling algorithm for scientific workflows on clouds, IEEE Transactions on Cloud Computing 2(2):222-235.
[7] Fard HM, Fahringer T, Prodan R.[2013] Budget - constrained resource provisioning for scientific applications in clouds, 2013 IEEE 5th International Conference on Cloud Computing Technology and Science, Cloud Com pp. 315-322.
[8] Thomas G, Williams AB.[2009], Sequential auctions for heterogeneous task allocation in multiagent routing domains, IEEE International Conference on Systems, Man and Cybernetics, SMC pp. 1995-200.
[9] Iturriaga S, Nesmachnow S, Luna F, Alba E.[2015] A parallel local search in CPU/GPU for scheduling independent tasks on large heterogeneous computing systems, Journal of Supercomputing 71 (2):648 -672.
[10] Amazon EC2, http://aws.amazon.com/ec2/, 2015.
[11] Amin A, Ammar R, El Dessouly A.[2004] Scheduling real time parallel structures on cluster computing with possible processor failures, Ninth International Symposium on Computers and Communications, ISCC pp. 62 -67.
[12] Apvrille A, Pourzandi M.[2004] XML distributed security policy for clusters, Computers & Security 23(8):649-658.
[13] Xie T, Qin X.[2006] Scheduling security -critical real -time applications on clusters, IEEE Transactions on Computers 55(7) :864 -879.
[14] Song S, Hwang K, Kwok YK. [2006] Risk-resilient heuristics and genetic algorithms for security -assured grid job scheduling, IEEE Transactions on Computers 55(6):703 –719.
[15] Xie T, Qin X.[2007] Performance evaluation of a new scheduling algorithm for distributed systems with security heterogeneity, Journal of Parallel and Distributed Computing 67(10) :1067 -1081.
[16] Tang X, Li K, Zeng Z, Veeravalli B. [2011] A novel security -driven scheduling algorithm for precedence constrained tasks in heterogeneous distributed systems, IEEE Transactions on Computers 60(7):1017-1029.
[17] Zeng LF, Veeravalli B, Li XR. [2015] SABA: a security -aware and budget-aware workflow scheduling strategy in clouds, Journal of Parallel and Distributed Computing 75: 141 -151.

[18] Gens F., IT cloud services user survey, pt.2: Top benefits & challenges (October 2008). URL http://blogs.idc.com/ie/?p=210
[19] Yurcik W, Meng X, Koenig G, Greenseid J. [2004] Cluster security as a unique problem with emergent properties, Fifth LCI International Conference on Linux Clusters: The HPC Revolution 2004, May 2004.
[20] Behl A, Behl K.[2012] An analysis of cloud computing security issues, 2012 World Congress on Information and Communication Technologies, WICT 109 -114.
[21] Magoulès F, Pan J, Teng F. [2012] Cloud Computing: Data-intensive Computing and Scheduling, CRC press,
[22] Kolodzie j, Xhafa F.[2011] Meeting security and user behavior requirements in grid scheduling, Simul. Modell.Pract. Theory19(1):213–226,doi: 10.1016/j.simpat.2010.06.007.
[23] Khan AN, Mat Kiah ML, Khan SU, Madani SA. towards secure mobile cloud computing: a survey, Future Gener. Comput. Syst.29 (5) (2013) 1278–1299, doi: 10.1016/j.future.2012.08.003.
[24] A Furfaro, A Garro, A Tundis, towards security as a service (secaas): On the modelingof security services for cloud computing, in: 2014 International Carnahan Conference on Security Technology (ICCST), 2014, pp. 1–6, doi: 10.1109/CCST.2014.6986995.
[25] Certification Report for Processor Resource/ System Manager (PR/SM) for the IBM eServer zSeries 900, BSI-DSZ-CC-0179-2003, 27 February 2003, Bundesamt für Sicherheit in der Information stechnik: Bonn, Germany. URL:http://www.commoncriteriaportal.org/public/files/epfiles/0179a.pdf
[26] Madnick SE, Donovan JJ. Application and Analysis of the Virtual Machine Approach to Information System Security. In Proceedings of the ACM SIGARCH-SIGOPS Workshop on Virtual Computer Systems. 26-27 March 1973, Cambridge, MA: Association for Computing Machinery. p. 210-224. URL: http:// portal.acm.org/ citation.cfm?id=803961.
[27] M. Schaefer,, B. Gold, R. Linde, and J. Scheid. Program Confinement in KVM/370. In Proceedings of the 1977 ACM Annual Conference. 16-19 October 1977, Seattle, WA: p. 404-410.
[28] Karger PA, PA ME Zurko, Bonin DW, AH Mason, Kahn CE, A Retrospective on the VAX VMM Security Kernel. IEEE Transactions on Software Engineering, November 1991. 17(11): 1147-1165.
[29] Joshua Schiffman, Thomas Moyer, Christopher Shal [2009]Justifying Integrity Using a Virtual Machine Verifier" IEEE Annual Computer Security Applications Conference
[30] Anusha B, Noah, Sivaranjani C, Priyanka S, [2015.]Predictive analysis of movie reviews using hybrid approach", International Research Journal of Advanced Engineering Sciences and Technologies, ISSN: 2455 - 8907, 1(1): 1-7
[31] Dharshini G, Subhasri V, Sujitha G, Ganesan M, [2016] "Secure Information Retrival for Decentralised Disruption Tolerant Military Networks using CP-ABE", International Research Journal of Advanced Engineering Sciences and Technologies, ISSN: 2455 - 8907, 2(1): 1-6
[32] Govindharaj I, Karthiga S, Manishalakshmi R.[2016] R Mary Silvia Theodore, "Home Power Analyzer with Smart Power Monitoring using IoT", International Research Journal of Advanced Engineering Sciences and Technologies, ISSN: 2455 - 8907, 2(1): 7-13
[33] Ahilandeswari T, Nandhini S, Sivasankari P, Rajalakshmy M.[2016] Intensifying the Generic Middleware for Smart Environment, International Research Journal of Advanced Engineering Sciences and Technologies, ISSN: 2455 - 8907, 2(2): 1-5,
[34] Gayathri R, Indumathi K, Githanjali P, Roobini V, [2016] Securing Multimedia using Data Lineage in Malicious Environment: A Survey, International Research Journal of Advanced Engineering Sciences and Technologies, ISSN: 2455 - 8907, 2(.2): 22-29.
[35] Shanmugam M, Dhivya S, Lavanya B, Keerthana V.[2017] Free Voice Calling in Wi-Fi Network using Android, International Research Journal of Advanced Engineering Sciences and Technologies, ISSN: 2455 - 8907, 3(1):1-8
[36] Rajadurai R, Amelia, Aubrey, A.Anusha, Danapriya P. Geethashnee D.[2017] Efficient Data Leakage Prevention Strategy using Key Distribution", International Research Journal of Advanced Engineering Sciences and Technologies, ISSN: 2455 - 8907, 3(1): 9-16
[37] Sasidevi V, Hannah D Sathiyan, Rajadurai R.[2017] Classification Algorithm: A Survey", International Research Journal of Advanced Engineering Sciences and Technologies, ISSN: 2455 - 8907, 3(1): 7-21