2021 7th International Conference on Advanced Computing & Communication Systems (ICACCS)

Elastic Resource Allocation, Provisioning and Models Classification on Cloud Computing A Literature Review

Bharanidharan.G
Department of Computer Science, VISTAS,
Pallavaram, Chennai
bharanmca@gmail.com

Jayalakshmi.S
Dept. of Computer Applications, VISTAS,
Pallavaram, Chennai.
jai.scs@velsuniv.ac.in

2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) | 978-1-6654-0521-8/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICACCS51430.2021.9442018

***Abstract*— Cloud Computing (CC) system faces new challenges every day, due to the complex structure of system clusters and high volume of data processed by the systems. The ability of acquiring resources in an elastic manner is considered as the primary rationale for adopting CC system. Elasticity mainly supports the facility to grow and shrink the virtual resources dynamically according to the requirement of cloud users. Elastic resource allocation plays a vital part in managing cloud resources efficiently according to user demands as well as cloud service provider's capacity. For gaining an efficient allocation of resource allocation in cloud, this paper has discussed the study of improving the load distribution and resource utilization with elasticity. However, literature shows that several works in CC and its benefits but there is a dearth of survey in analyzing detail about elasticity present in the cloud. This study has an attempt for fulfilling an existing available gap and introducing this survey based on adapting elasticity in CC. This paper has involved with various elasticity mechanism inclusive of definition, tools to measure, elasticity evaluation and existing elastic solutions with some issues faced. Finally, this paper has presented few open issues and new solutions. To the best knowledge, this study is novel to identify the issues faced in the model of elasticity solution using systematic review method.**

Keywords— Elasticity, Systematic Review, Metrics, Strategies, Elastic Resource Allocation, Cloud Service Provider, Reactive, Proactive.

I. INTRODUCTION

Cloud Computing is more familiar during last few years and has established an immense attention deal from both academia and industry. The major aspects have motivated usage of cloud services in their capability of delivering infrastructure based on the requirement of customers or the requirement of elastic provisioning and divest. Hence, the elasticity is a major advantage of CC which automatically shifts the amount of allocated resources amount for satisfying the variations occur in demands of work load [1]. In general, the virtualization based method has utilized cloud provider for creating their stack. The possibility of virtualization is to execute various applications and Operating System (OS) at the same time and on same server. This creates an abstract layer which hides the complexities of working environment such as hardware and software. The CC framework has facilitates quick implementation and scaling of workloads through the accelerated provisioning of virtualized services. However, the implementation is done by Virtual Machines (VMs) through virtualization using hypervisors. Therefore, the hypervisor perform the virtualization technique which involve multiple OS for sharing a single hardware host which has considered that each OS perform to be its own independent resources. Thus, the VM ware ESX, Xen, Hyper-V and KVM are the example of hypervisor utilized in worldwide.

One of the other modes of virtualization is container-based virtualization, or OS level virtualization in which the virtualization layer operates within the same OS as an application. Containers are performed as a lightweight solution which provide for less overhead and fast start-up time [2]. Hence, virtualization act as a core component of CC which improves elasticity whereas cloud elasticity has addressed in the sense of both containers and VMs. Elasticity plays a vital role for incorporating Quality of Service (QoS) in cloud services. It enables providers without interference in adding or deleting features for controlling the load variance at runtime. There are several researchers from academia and industry who have investigated elasticity. Several virtualization solutions that focus on the cloud continue to grow. Therefore, considerable attempts have been made for enabling cloud structures to function with elastic nature and several applications continue to emerge. Thus, it has inspired the researcher to provide a detailed and expanded classification to CC elasticity. In general, the elasticity is frequently related to system scaling but there are some differences. The system capability of adding more resources to satisfy a larger workload requirement is known as scalability [3]. Elasticity contains resource extraction and reduction based on the workload, when the scalability is considered only for growth.

Moreover, the subset of the problem set is cloud auto scaling available in the CC. Hence, this survey has specifically focused on cloud elasticity and its associated literature. Most of the reviewed survey papers are innovative and highly overlay on elasticity features needs like method, model, scope and policy. The scope of cloud environment is wider that represent overall elasticity of cloud research but it lack in detail about essential implementation technique of elastic solution. Since, there are various review papers that have utilized several elasticity features for classifying entire cloud elasticity. In this classification, the basic implementation method is performed as an attribute and hence there is no classification represented in existing literature that has accomplished the attributes of specific implementation method. According to this literature, the definition, mechanism, method, open issues and new solution for elasticity is discussed associated with industrial and organizational work.

978-1-6654-0521-8/21/$31.00 ©2021 IEEE

II. Elasticity Feature on CC

The system degree has ability for adapting in modifying the workload by provision and divests resources within an autonomic way is defined as elasticity. At extend each point in time, the resources availability get matched with the recent demand as near as probable. There are many definitions available in the literature for elasticity [4]. For CC, the elasticity is a dynamic property and it consists of two types namely

II.I Horizontal Elasticity (HE)

II.II Vertical Elasticity (VE)

HE involves adding or removing instances of application-associated computing tools. Vertical elasticity involves in increase or decrease, moving, resizing features of computing resources namely CPU time, memory, network bandwidth and cores to large or small Virtual Machines (VM's).

The categorizing of application and infrastructure level is done by Galante and De Bona[5] and the taxonomy consists of various characteristics such as purpose, scope, evaluation proposed and decision making mechanism [6]. There is a provision of taxonomy which can able to focus even on the aspects of application provider [7].The researcher has focused on the aspects of action type, strategy and architecture [8] in which the adaptability view of larger-scale computing resources are involved with the concepts such as node adaptation and migrations of VMs available [9]. However, the adaptation methods have utilized to study the literature as one of the dimensions. Hence, function of elasticity like reactive replication, reactive resizing, reactive migration and proactive replication are utilized for classification [10]. Nevertheless, the elasticity control can be performed or executed from the platform of its own applications. Thus, the elasticity controller is made embedded to an application or executing environment that merged with cloud based on the resource request or release. In these cases, the basic infrastructure needs to be supported by elasticity characteristics. This method is typically implemented in Platform as a Service (PaaS) clouds whereas the resources present in the applications are managed automatically using containers [11]. In general, the APIs is used as an interface for the cloud provider to interact user with the system. There are some public provider namely Microsoft Azure, GoGrid and Rackspace and the framework of Elastin, work Queue, etc. are the systems that manages the resources manually [12].
The utilization of monitoring services of the system results to determine the resource scale in or scale out.

III. Elasticity Resource Allocation in CC

Elasticity is a central facet in CC that maintains good rapport with cloud users and providers. Rajkumar Buyya, et al. [14] focuses over handling and monitoring of cloud resources energy consumption based on elasticity. Bacigalupo, et al. [15] described the resource scaling methods with the help of queuing models which contain three levels such as database, storage disk and application. Each level had analyzed the result which contains throughput, response time and resource utilization of a cloud server. Based on the analysis, elastic cloud resource allocation algorithm was suggested to scale applications in dynamic environment. Elasticity handles the postulation for new cloud resources and relinquishes idle resources. Elastic resource management scaling techniques [16] were classified as predictive or reactive approach.

TABLE I

ELASTICITY SOLUTIONS CLASSIFICATION

Service Providers	Policy	Scope	Method	Evaluation
GoGrid	Manually	Infrastructure	Replication	Performance
AmazonWeb Service (AWS)	Automatic Reactive/Predictive	Infrastructure	Replication	Performance
Rackspace	Manually	Infrastructure	Replication	Performance
Rightscale	Automatic Reactive	Infrastructure	Replication	Performance
Scalr	Automatic Reactive	Infrastructure	Replication	Performance
Tide	Automatic Reactive	Infrastructure	Replication	Performance
Press	Automatic predictive	Infrastructure	Resizing	Performance
Elastic VMs	Automatic predictive	Infrastructure	Resizing	Performance
Chp	Automatic Reactive	Infrastructure	Replication	Increase infrastructure capacity

Predictive approach refers to the elastic cloud system which attempts to forecast succeeding resource demands in order to guarantee adequate cloud resources. It was based on heuristic proficiency, control theory, strengthener learning, time series analysis and analytical techniques. The number of cloud servers needed to handle the predicted system workload was decided using the workload predictor and execution model. The predictive based techniques are proposed for automatically scaling of cloud resources. In the Reactive approach, the system responded to workload changes without pre-planning and focused on the target measured time based resource scaling. Typically, cloud environment has an elastic controller which handles cloud user requisite conversion, monitoring of applications and scaling of the cloud resources [17]. But the controller was desirable only for client-server applications to manage workload. In the scheduling techniques elastic clouds for minimizing cost to

both cloud provider and user were proposed, which conceived service level agreement (SLA) violations and job priorities. Villegas, et al. [18] exhibited four resource allocation policies, eight resource provisioning techniques and cost analysis in a detailed study for Infrastructure as a Service (IaaS) clouds.

Containers have enabled the implementation of applications that can be quickly adapted by horizontal and vertical scaling. However, the horizontal elasticity causes the container for increasing and decreasing the data which means scale-in and scale out. The scale up and scale down of resources computation designated for every application is done through Vertical elasticity whereas several existing solution is horizontal or vertical elasticity. The complete utilization of elasticity applications can perform much quicker reaction to minimum workload variation through fine grained vertical scaling and also for rapid workload peaks by horizontal scaling. However, there are very few works have discussed the advantages of integrating the two dimensions of elasticity in the container-based applications. Based on the recent and evolving developments in hardware is involved with usual increase in server capability and the rack-scale computing emergence. Similarly, the vertical elasticity perform as a promising field for dynamic application scaling and even an initial selection to elastic scaling before attempting horizontal scaling [19]. Vertical elasticity has the capability in resizing the application existed dynamically in containers or VMs [20], [21].

III.I. Horizontal Elasticity

In this CC, workload distribution has been resolved using soft computing based stochastic hill climbing method by Brototi Mondala et al. [22]. The researcher has aimed for an entire algorithm response time but not the elastic load distribution. The Genetic Algorithm (GA) is utilized for load balancing technique to minimize the workload timing by Kousik Dasgupta, et al. [23]. This algorithm discuss about load balancing which doesn't depend upon resource elasticity and its utilizations. L.D Dhinesh Babu, et al. has introduced an algorithm as Honey Bee Behavior inspired Load Balancing (HBB-LB) to the task of non-preemptive independent [24]. Similarly, resource scheduling algorithm of an adaptive VMs is introduced by Weiwei Kong, et al. [25] with the consideration of auction date, network bandwidth and cost. The algorithm consisted of two steps namely the cloud property evaluation and the cloud consumer payment process. The architecture of Cloud Load Balancing (CLB) and the recent algorithm are suggested from Shang-Liang Chen, et al. for improving load in CC services whereas the processing power and machine loading is considered in this CLB architecture [26].

III.II. Vertical Elasticity

Roy, et al. has proposed an algorithm of predicted based VM allocation that can be utilized by second rate autoregressive travelling average prediction method to optimize the benefit of cloud application in an average prediction horizon [27]. The procedural method has introduced by Buyya, et al. to simulate large-scaled networks and network communications using CloudSim in CC environments [28].Virtualization technologies has been illustrated by Di Costanzo, et al. for acknowledging resource allocation between the grids and even allocation of resources over inter-grid mechanism in accordance with time constraint [29]. The agents of cloud coordinators is introduced by Rodrigo N. Calheiros et al. is utilized for maintaining accuracy, scalability and performance of elastic cloud applications in intercloud [30]. Therefore, instead of traditional process efficiency, the evaluation and cost evaluation of costs and profit to resource utilization and provision policies has a major impact [31].

The reference gap present in the HE is about existing algorithms that did not focus on premature converge and further, the parameter settings for convergence of the algorithm was also not concentrated. Hence, in order to overcome the demerits of existing work, the technique named Horizontal Elastic Resource Allocation Technique (HERAT) is proposed. Similarly in the case of VE the result obtained from the existing works seems to be promising, they fail to avert local optimum problem. These existing works fail to fix infeasible solutions generated by an algorithm. We concentrated on the resource allotment challenges from the perspective of resource provider, where the admission control policies and profit based scheduling are focused. In order to overcome such flaws and to provide better resource allocation, the technique named Vertical Elastic Resource Allocation Technique (VERAT).

IV. Model for Elasticity

STRATOS act as a broker service introduced by Pawluk et al. for promoting cross cloud application of topology platform building and modification of runtime based on the implement objectives [32]. However, the STRATOS has facilitated cross-provider with replica server's instantiation to determine an issue of multi-criteria optimization. The meth of management and virtualization namely SLA, multi-tendency and dynamic provisioning are illustrated by Mauch et al.[33]. The CC environment with higher performance has been introduced by researcher using interconnection of high speed cluster like In-finiBand. This method gets acknowledge with rapid migration of VMs for maintaining QoS and network performance. The method proposed by Dawoud et al. is used for performing operations of vertical scaling and even presented the comparison among their architecture of multi-instance and implementation is shown in Table 2 [34].This comparison involves analytical queuing model and its implementation features for every architecture. They dispute vertical scalability which involves low response time and less overhead provisioning that may resulted in less SLAs violation.

The probabilistic method presented by Konstanteli et al. on the virtualized physical resources for optimal service allocation issues based on the requirement of horizontal elasticity [35]. The establishment of probabilistic admission control test can be formulated using optimization model. This research has utilized prediction models based

on statistical information which has the ability to forecast the resource provision with assist of historical monitoring data. ElastiCat is a loop model for Control Analyze Plan Execute (CAPE) which used as a framework of load rebalancing has been presented by Qin et al. [36]. When the rebalancing costs get minimized, the framework has focused for eliminating the load imbalance. ElastiCat is used for two interfering aware prediction model namely predicting data migration performance and time impact by statistical ML to build a cost model by providing trade-offs among them.

The elasticity controller is developed by Ali-Eldin et al.has modified the allocation of VMs in accordance with reactive actions and subsequence load predictions [37]. This method is based on queuing theory and it is utilized for creating a hybrid controller namely reactive adaptive controller for sudden reaction to the rapid load change which inhibit premature resources release by considering the heterogeneity workload and prevent oscillations. The results illustrated that proposed hybrid controller has reduced oscillations of resource like reduction of SLA violation but there is a trade-off in over provisioning among 20% and 30%. Similarly, the algorithm of cost effective database placement has been produced by Yu et al. with detailed database placement and migration plan based on constraints in resources and user using system preferences[38].For reducing the migration cost and increase the use of cloud resources from researcher by combining reactive solution with migration planning and it creates a system with prototype on top of IBM cloud platform in this case study. The ElasTraS is used as scalable, multitenant DBMS to transact workload process and Albatross technique in the case of lightweight migration of live database are presented by Das et al.[39].Albatross gets advanced from ElasTraS in load balancing of elasticity by accumulating more servers in the case of high load and stabilized to some servers.

TABLE II

MODELS, METHOD AND MECHANISM OF ELASTIC SOLUTIONS AND ITS DISADVANTAGES

Author	Model	Method	Mechanism for elastic solution	Disadvantages
Fito et al. [40]	Reactive	Replication	Cloud Hosting Provider used to scale the implemented web applications. Scalability has outsourced to third party agencies who handle the operations and managements of web server requirements.	Scale up and scale down from the several backend servers based on input load becomes critical operation of the CHP. Another major issue is fault tolerance has not considered.
Perez-Sorrosal et al.[41]	Reactive	Replication	One of the recent multi version caches is Elastic SI-Cache which obtains high performance and reliability in multi-tier systems. Elastic SI-Cache can able to endure scalable replication of various tiers whereas replicas can be accumulated or eliminated dynamically with good scalability.	In the case of cache miss, data present in the cache has become wasted. The result ensues due to no updates in the cache when data gets modified over database. When move on to a new node, failure of node or scale out may cause missing data which may occur until data are added or updated on the database.
Pawluk et al.[32]	Reactive	Replication	STRATOS is a broker service for facilitating cross-cloud, constructing platform of application topology and runtime modification. STRATOS' has the capability in facilitating the acquisition of cross-cloud resource based on implement directives.	The highly complicated set up of STRATOS testing has involved provision of several SMI attributes. To reject the requirement in configuring specifications and focused an acquisition process of application driven resources.

He et al. [42]	Reactive	Migration	The managing solution of effective resource has specifically planned to assist small and medium sized IaaS cloud providers for good utilization in their hardware resources with less operational cost. The model of economic intensive has presented for encouraging cloud users to shut down its VMs in improving the performance of resource reallocation.	The resource migration of cross IDC is complex in provider-level profit or resource to share agreement. This is cost sensitive whereas analyzing of the operational cost improvement to optimize resource utilization.
Mauch et al.[33]	Reactive	Migration	The elastic virtual cluster provision as suitable pay-as-you-go resources get avoided the initial cost in hardware with physical maintained and acknowledge great flexibility in scalability for customers and their application. InfiniBand has been supported using HPC IaaS which acknowledge the configuration of individual	InfiniBand VMs of live migration configuration is until unsolved issues due to the protocol of machine specific offloading over adapter hardware. The effective benefit of the inadequate resources of QoS that may limit the maximum isolated virtual cluster number over similar physical context.
			network with QoS mechanisms.	
Dawoud et al.[34]	Reactive	Resizing	Elastic VM has reduced the overhead of provision and results with less SLOs violation. It has scaled the applications such as Databases with less cost and low complexity.	In the pattern of several web application and workload demand. There may be complexity in optimized the elastic VMs packing into physical hosts.
Calheiros et al.[43]	Proactive	Replication	This model aim is to suit targeted QoS associated with service time, requesting rejection rate and available resources utilized. This model is Simulation based provisioning which has ability to detect the modification in arrival pattern and demands of resource demands that may ensue over time and resource allocations.	The queueing model has acknowledged modeling composite services and access for cloud storage required for development. The SLA management issue for QoS trade-offs among various requests, possibly with several significances and incentives.
Konstanteli et al.[35]	Proactive	Replication	This optimization model establishes an admission of probabilistic control test. It has exploited statistical information	The computation time in solving the problem is high because it is in need of developing heuristic method.

			about the elastic workload requirements of horizontally scalable services to reduce the respectively booked resources.	
Bruneo [44]	Proactive	Migration	In CC, the exact evaluations of these parameters are needed for quantifying the obtainable QoS and opportunely manage SLAs.	The representation of PaaS and SaaS Cloud systems are required for integrating the required mechanisms needed in capturing VM migration to envelop a critical role in policies of energy saving.

V. Open challenges and new solution in elasticity of resource allocation

In the subsequent decade, several enterprises may shift to the strategy of hybrid cloud based on the announcement of Gartner's. In a hybrid cloud, there are few major trends associated for maintaining integration, elasticity and performances over hybrid cloud. Consequently, the virtualization of container based may be an alternate for reducing the overhead of virtualization and hence it improved the benefits of resources in datacenters. Containers can able to minimize the replicas start-up time after it consider in an application and its dependencies which can be utilized for automating implements, application scaling backend services and databases that are better features in PaaS environments. The event between the academic efforts and academic may proceed for driving innovation and results development. This literate review has found out various elastic solutions whereas cloud provider and academic have improved various solution but few existing problems are not discussed completely. There are some challenges associated with CC elasticity. According to this review, the challenges are found out are represented in review papers.

This research study focus on the solution of these challenges has been valuable to improve an efficient method of auto scaling which has the capability for managing several application conditions resulting from unexpected modification in traffic. In order to combine the method of reactive and proactive that represented to be a better idea. The automatic scaling issues has been addressed by various models namely reinforcement learning, control theory, time-series analysis, threshold based policies and queuing theory.

VI. Conclusion

The pay-as-you-go models and on-demand service with elasticity can be provided by CC for maintaining the variation of workload and elasticity has enabled the cloud system for scaling its resource by up or down computation without disturbances. However, in this system elasticity plays a major part in handling the infrastructure, quality establishment and development of resource usage. This paper has illustrated the survey on CC elasticity based on adoption of traditional systematic review and it find out the idea available in the literature for elasticity of CC. Moreover, the main idea presented has utilized in CC environment namely strategies, metrics and tools to maintain. This paper also provides unique ideas into future methods to further understanding in the concept of elasticity and reasoning. In general, there are several CC is enclosed the allocation of resources architecture, models and tools for common purposes, service cost definition over cloud and benefits present in heterogeneous environments. The limited taxonomy of existing elastic solutions have focused on the scalability mechanisms, models, and processes, which addressed several open problems in several areas, like metrics, resources, tools, availability, autonomic computation, elastic methods, and techniques.

References

[1] L. Badger, T. Grance, R. Patt-Corner, and J. Voas, "Draft cloud computing synopsis and recommendations," NIST special publication, vol. 800, p. 146, 2011.

[2] C. Pahl, "Containerization and the Paas Cloud," IEEE Cloud Computing, vol. 2, no. 3, pp. 24–31, 2015.

[3] Islam, S., Lee, K., Fekete, A., Liu, A.: How a consumer can measure elasticity for cloud platforms. In: Proceedings of the third joint WOSP/SIPEW international conference on Performance Engineering, ICPE '12, pp. 85-96. ACM, New York, NY, USA (2012).

[4] E. F. Coutinho, F. R. de Carvalho Sousa, P. A. L. Rego, D. G.Gomes, and J. N. de Souza, "Elasticity in Cloud Computing: aSurvey," Annals of Telecommunications, pp. 1–21, 2015.

[5] Galante, G., De Bona, L.C.E.: A survey on cloud computing elasticity. In: Proceedings—2012 IEEE/ACM 5th International Conference on Utility and Cloud Computing, UCC 2012, pp. 263–270 (2012).

[6] Naskos, A., Gounaris, A., Sioutas, S.: Cloud elasticity: a survey. In: Algorithmic Aspects of Cloud Computing, pp. 151–167. Springer, Cham (2016).

[7] Shoaib, Y., Das, O.: Performance-oriented cloud provisioning: taxonomy and survey (November):1–14 (2014). arXiv:1411. 5077 21.

[8] Najjar, A., Serpaggi, X., Gravier, C., Boissier, O.: Survey of elasticity management solutions in cloud computing. In: Continued Rise of the Cloud, pp. 235–263. Springer, London (2014) 22.

[9] Hummaida, A.R., Paton, N.W., Sakellariou, R.: Adaptation in cloud resource configuration: a survey. J. Cloud Comput. 5(1), 1–16 (2016) 23.

[10] Coutinho, F.E., De Carvalho Sousa, F.R., Rego, P.A.L., Gomes, D.G., De Souza, J.N.: Elasticity in cloud computing: a survey. Ann. Telecommun. 70, 289–309 (2015).

[11] E. Caron, L. Rodero-Merino, and A. M. F . Desprez, " Auto-scaling, load balancing and monitoring in commercial and open-source clouds, " INRIA, Tech. Rep. 7857, 2012.

[12] D. Rajan, A. Canino, J. A. Izaguirre, and D. Thain, "Converting a high performance application to an elastic cloud application, " in Proceedings of the 3rd Intl. Conference on Cloud Computing T echnology and Science, ser . CLOUDCOM '11. IEEE, 2011, pp. 383–390.

[13] R. N. Calheiros, C. V ecchiola, D. Karunamoorthy, and R. Buyya, "The aneka platform and qos-driven resource provisioning for elastic applications on hybrid clouds, " Future Generation Computer Systems, vol. 28, no. 6, pp. 861–870, June 2011.

[14] R. Buyya, R. Ranjan, and R. Calheiros, "Intercloud: Utility-oriented federation of cloud computing environments for scaling of application services." In ICA3PP 2010: The 10th International Conference on Algorithms and Architectures for Parallel Processing, pp. 19–24, 2010.

[15] D.A. Bacigalupo, Jinn-Tsong Tsai, Jia-Cen Fang, Jyh-Horng Chou, "Optimized task scheduling and resource allocation on cloud computing environment using improved differential evolution algorithm", Computers & Operations Research, Vol. 40, No. 12, pp. 3045-3055, 2013.

[16] T. Lorido-Botran, J. Miguel-Alonso and J.A. Lozano, "Auto scaling Techniques for Elastic Applications in Cloud Environments", Technical Report EHU-KAT-IK-09-12, 2012.

[17] S.K. Tesfatsion, E. Wadbro, J. Tordsson, "A combined frequency scaling and application elasticity approach for energy-efficient cloud computing", Sustainable Computing: Informatics and Systems, Vol. 4, No. 4, pp. 205-214, 2014.

[18] D. Villegas, A. Antoniou, et al. "An Analysis of Provisioning and Allocation Policies for Infrastructure-as-a-Service Clouds", Cluster, Cloud and Grid Computing (CCGrid), 2012 12th IEEE/ACM International Symposium on, May 13-16, pp. 612-619, 2012.

[19] E. B. Lakew, C. Klein, F. Hernandez-Rodriguez, and E. Elmroth, "Towardsfaster response time models for vertical elasticity," in Proceedingsof the 2014 IEEE/ACM 7th International Conference on Utility andCloud Computing. IEEE Computer Society, 2014, pp. 560–565.

[20] Y. Al-Dhuraibi, F. Paraiso, N. Djarallah, and P. Merle, "Autonomicvertical elasticity of docker containers with elasticdocker," in 10th IEEEInternational Conference on Cloud Computing, IEEE CLOUD 2017,2017.

[21] E. Kalyvianaki, T. Charalambous, and S. Hand, "Adaptive resource provisioningfor virtualized servers using kalman filters," ACM Transactionson Autonomous and Adaptive Systems (TAAS), vol. 9, no. 2, p. 10, 2014.

[22] Brototi Mondala, Kousik Dasgupta, Paramartha Dutta, "Load Balancing in Cloud Computing using Stochastic Hill Climbing-A Soft Computing Approach", Procedia Technology, Vol. 4, pp. 783–789, 2012.

[23] Kousik Dasgupta, Brototi Mandal, Paramartha Dutta, Jyotsna Kumar Mondal, Santanu Dam, "A Genetic Algorithm (GA) based Load Balancing Strategy for Cloud Computing", International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA) 2013, Procedia Technology, Vol. 10, pp. 340–347, 2013.

[24] Dhinesh Babu L.D., P. Venkata Krishna, "Honey bee behavior inspired load balancing of tasks in cloud computing environments", Applied Soft Computing, Vol. 13, pp. 2292–2303, 2013.

[25] Weiwei Kong, Yang Lei, Jing Ma, "Virtual machine resource scheduling algorithm for cloud computing based on auction mechanism", Optik, Vol. 127, pp. 5099–5104, 2016.

[26] Shang-Liang Chen, Yun-Yao Chen, Suang-Hong Kuo, "CLB: A novel load balancing architecture and algorithm for cloud services", Computers and Electrical Engineering, Vol. 58, pp. 154-160, 2016.

[27] A.-F. Antonescu, T. Braun, "Modeling and simulation of concurrent workload processing in cloud-distributed enterprise information systems", in: ACM SIGCOMM Workshop on Distributed Cloud Computing, DCC 2014, 2014.

[28] Alexandru-Florian Antonescu, Torsten Braunb, "Simulation of SLA-based VMscaling algorithms for cloud-distributed applications", Future Generation Computer Systems, Vol. 54, No. C, pp. 260-273, 2015.

[29] A. Haldorai and A. Ramu, "Security and channel noise management in cognitive radio networks," Computers & Electrical Engineering, vol. 87, p. 106784, Oct. 2020. doi:10.1016/j.compeleceng.2020.106784

[30] A. Haldorai and A. Ramu, "Canonical Correlation Analysis Based Hyper Basis Feedforward Neural Network Classification for Urban Sustainability," Neural Processing Letters, Aug. 2020. doi:10.1007/s11063-020-10327-3

[31] D. Devikanniga, A. Ramu, and A. Haldorai, "Efficient Diagnosis of Liver Disease using Support Vector Machine Optimized with Crows Search Algorithm," EAI Endorsed Transactions on Energy Web, p. 164177, Jul. 2018. doi:10.4108/eai.13-7-2018.164177

[32] Pawluk, P., Simmons, B., Smit, M., Litoiu, M., Mankovski, S.: Introducing stratos: Acloud broker service. In: Cloud Computing (CLOUD), 2012 IEEE 5th InternationalConference on, pp. 891{898 (2012).

[33] Mauch, V., Kunze, M., Hillenbrand, M.: High performance cloud computing. Future Generation Computer Systems (0), { (2012).

[34] Dawoud, W., Takouna, I., Meinel, C.: Elastic vm for rapid and optimum virtualizedresources' allocation. In: Systems and Virtualization Management (SVM), 20115th International DMTF Academic Alliance Workshop on, pp. 1-4 (2011).

[35] Konstanteli, K., Cucinotta, T., Psychas, K., Varvarigou, T.: Admission control for elasticcloud services. In: Cloud Computing (CLOUD), 2012 IEEE 5th International Conferenceon, pp. 41-48 (2012).

[36] Qin, X., Wang, W., Zhang, W., Wei, J., Zhao, X., Huang, T.: Elasticat: A load rebalancingframework for cloud-based key-value stores. In: High Performance Computing(HiPC), 2012 19th International Conference on, pp. 1-10 (2012).

[37] Ali-Eldin, A., Kihl, M., Tordsson, J., Elmroth, E.: Efficient provisioning of burstyscientific workloads on the cloud using adaptive elasticity control. In: Proceedingsof the 3rd workshop on Scientific Cloud Computing Date, ScienceCloud '12, pp.31-40. ACM, New York, NY, USA (2012).

[38] Ali-Eldin, A., Kihl, M., Tordsson, J., Elmroth, E.: Efficient provisioning of burstyscientific workloads on the cloud using adaptive elasticity control. In: Proceedingsof the 3rd workshop on Scientific Cloud Computing Date, ScienceCloud '12, pp.31-40. ACM, New York, NY, USA (2012).

[39] Das, S., Agrawal, D., El Abbadi, A.: Elastras: An elastic, scalable, and self-managingtransactional database for the cloud. ACM Trans. Database Syst.38(1), 5:1{5:45 (2013).

[40] Fito, J., Goiri, I., Guitart, J.: Sla-driven elastic cloud hosting provider. In: Parallel,Distributed and Network-Based Processing (PDP), 2010 18th Euromicro InternationalConference on, pp. 111 {118 (2010).

[41] Perez-Sorrosal, F., Pati~no Martinez, M., Jimenez-Peris, R., Kemme, B.: Elasticsi-cache: consistent and scalable caching in multi-tier architectures, TheVLDB Journal 20(6), 841-865(2011).

[42] He, S., Guo, L., Guo, Y.: Real time elastic cloud management for limited resources.In: Cloud Computing (CLOUD), 2011 IEEE International Conference on, pp. 622- 629(2011).

[43] Calheiros, R., Ranjan, R., Buyya, R.: Virtual machine provisioning based on analyticalperformance and qos in cloud computing environments. In: Parallel Processing (ICPP),2011 International Conference on, pp. 295 {304 (2011).

[44] Bruneo, D.: A stochastic model to investigate data center performance and qos in IaaS cloud computing systems, pp. 1-1 (2013).