International Conference on
Contemporary Technological Solutions towards fulfillment of Social Needs

Survey of Cloud Virtualization Management and Placement Methods

Rajesh Sharma [1], Ravi Singh Pippal [2],
[1,2]Department of Computer Science & Engineering,
RKDF University, Bhopal, India

Abstract: Cloud models enables on-demand network access conveniently from shared computing resources such as application, services, servers computing and storage which can be quickly provided and delivered with minimum management effort or less interaction of the service provider. Surrogation placement techniques are necessary for content delivery to determine content to be stored into another distributed surrogate servers geographically in the cloud network to fulfill end-user requirements with quality of service (QoS). Cloud computing services host large number of virtual machines (VMs) for demanding conditions as virtualization where surrogation placement in required. In this review, an logical perspective on cloud computing in virtualization by surrogation placement techniques are provided for economical utilization of energy consumption and efficient resource management.

Keywords: Cloud Computing, Surrogation Placement, Virtualization, Content Delivery Network.

I Introduction

Cloud computing is basically a distributed computing network model for enabling convenient and accessible, on-demand network to a shared bunch of configurable and reliable computing resources such as servers, networks, applications, storage and services which can be rapidly supplied and brought out with minimal management effort or interaction of service providers [1]. It rapidly provide resources with minimal management effort and service provider interaction [2]. The underlying concept of cloud computing is to bring the specific services from shared resources with the help of virtualization technology [3]. The aim of cloud computing model is to produce a powerful usage of distributed resources, assign them along to make high turnout and to control large-scale computation efficiently and economically [4]. Cloud computing architecture as shown in Figure 1 is classified into three levels of model [5]. computing services .

- **Infrastructure-as-a-Service (IaaS):** IaaS is basically hardware and software resources that operate it all like servers, networks, storage and operating systems. Cloud computing substitute primarily hardware resources. Users of IaaS layer manage to support applications and functional systems, however there is no requirement for server, networking hardware and storage, so it is an information center to control the hardware. Popular example of this supplier corporation is Amazon [6].

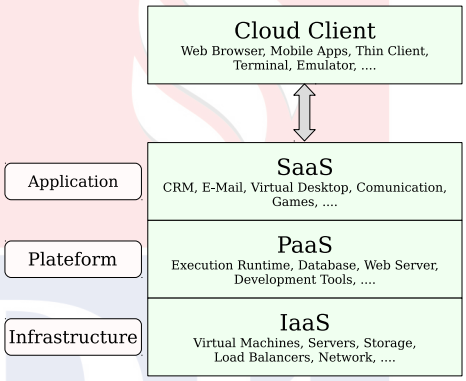

Figure 1: Cloud Computing Architecture

- **Platform-as-a-Service (PaaS):** PaaS is a collection of services and tool developed to create coding and deploy applications fast and efficiently. Cloud computing replaces machine language by providing the system to execute software system of the user. The example of PaaS is the suppliers, corporations such as Google [7].

- **Software-as-a-Service (SaaS):** SaaS applications are developed for end-users service over the network. The cloud users interact directly with this cloud software system and pays for usages. Popular examples of SaaS supplier is Google Apps [8][9].

II Cloud Content Delivery Network (CDN)

Recent advances in utility and cloud computing allow leasing resources, such as storage and bandwidth, to build Content delivery Networks (CDNs) in the cloud [10]. There is a growing trend to deploy cloud-based CDNs (CCDNs) or to complement traditional CDN infrastructure with cloud-based delivery, management and analytic services. Undoubtedly, there is a move to CCDNs – which is evident by the increase in traffic across datacenters, attributed largely to CCDNs [11].

III Cloud Data Center

With rapid growth in information and data flow, storage service is going to be the biggest business organization, along with the storage cost and enhancement are other key factors. This known approach of data storage service is the cloud data center (CDC). Cloud data centers is a group of heterogeneous computing with storage components clustered together to host various applications and to store data, using very fast communication links [12][13]. The customer pay charges to the cloud operator for cloud resource application usage based on the "pay-as-you-go" service model. This service model certifies that the customer pays only for the resources they have actually used in the specified time duration. Moreover, CDC offers a wide range of service types to be accessed via web links. However, cloud computational cost and application performance are profoundly affected due to internal resource fragmentation (low server utilization) and bandwidth constraints imposed by the CDC network architecture design [14][15].

Network architecture is a vital element in CDC network design since it significantly influences CDC throughput. Modern data center network architecture designs are based on the hierarchical tree-based Three-Tier architecture. In three-tier architecture, the core layer connects CDC to the Internet backbone, the aggregation layer implements assorted functionalities such as firewalls and content switching, and the access layer ensures inter-rack connectivity. According to network routing protocol design, CDC architecture is classified as switch-centric, server-centric, and hybrid models [16]. Figure 2 highlights a set of services offered by CDC to its consumers. As depicted in Figure 2, the cloud operator controls and manages the cloud services, including PaaS, SaaS and IaaS.

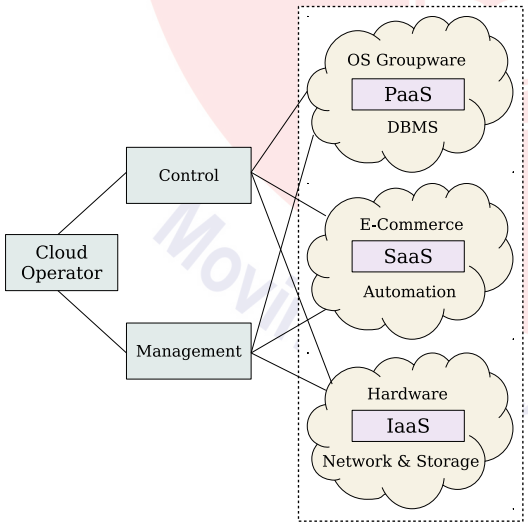

Figure 2: Cloud Service Models

Among others, the SaaS service model offers accounting applications, including e-commerce, office automation, and knowledge management services. Likewise, IaaS model encompasses hardware resources such as storage, processor, and network [17]. Alternatively, PaaS service model offers several cloud execution platforms, such as developer studio, database management software (DBMS), groupware, and operating systems (OS) to help the IT professionals to develop, deploy, test, debug, and host sophisticated web applications. Within a CDC, virtualization proficiently manages the cloud resources to effectively offer aforementioned services to the customer [18].

IV Surrogation Placement in Cloud Computing

Surrogation placement has become a fundamental element in today's CDC owing to the support of isolating, consolidating and migrating server workload. VM migration schemes migrate the state of virtual devices (e.g. memory, CPU, I/O) between physical hosts during the surrogation placement process [19][20]. The hypervisors migrate a surrogation placement either in non-live or live communication mode. The non-live surrogation placement mode does not service the migrant applications during surrogation placement but the live surrogation placement mode does.

Virtual machine migration schemes seek to upgrade manageability, fault tolerance, and application's performance within a CDC. Figure 3 illustrates a general overview of a distributed CDC framework, wherein a VM is migrated (live mode) from an underutilized server to a resource-rich server to power off the former for efficient resource utilization.

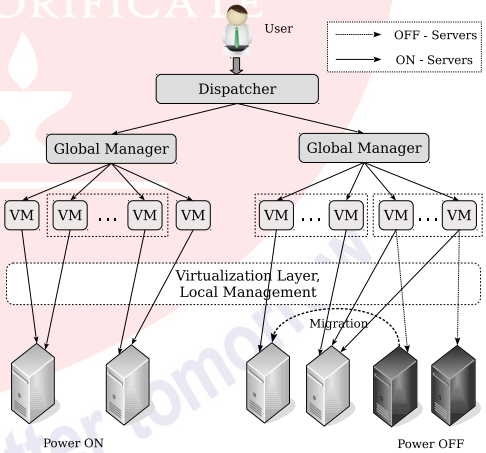

Figure 3: VM Migration Process on Distributed CDC

The global manager manages a set of physical servers (preferably a cluster) and creates surrogation placement for the new application in response to the request by the dispatcher module. However, the local manager examines the system components' health and recommends the global manager to migrate a few VMs if required. In response,

the global manager signals the VMM to trigger surrogation placement. Figure 3 presents surrogation placement between servers located under same administrative domain, however, surrogation placement can be triggered across the CDCs owned and managed by different providers.

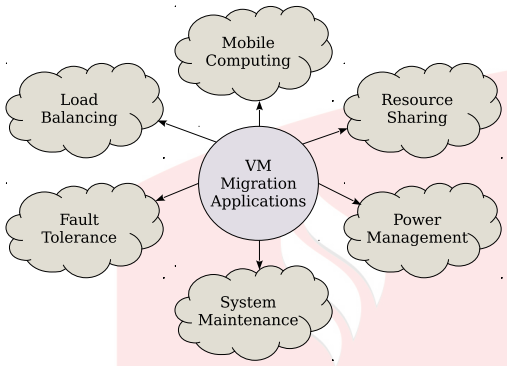

Figure 4: Applications under Surrogation Placement

The VM migration technology assists to achieve various resource management objectives as shown in Figure 4. A detailed discussion on the applications of VM migration is as explained below:

- Power management To attain power efficiency within a CDC, surrogation placement process shifts complete server workload from an under-loaded server (e.g. when resource usage below a threshold) to an underutilized server to switch off the former. Server consolidation and DVFS enabled VM migration methods aggressively co-locate the surrogation placements and decrease CPU clock rate, respectively, to achieve power efficiency within a CDC at the cost of application's performance degradation.

- Resource sharing The application performance degradation issue due to sharing limited system resources, such as system memory, cache, or CPU cycles, can be resolved by relocating resource hungry surrogation placement server to a resource-rich server. However, high system resource sharing reduces cloud operational cost as unnecessary servers can be switched off.

- Fault tolerance A fault-tolerant system triggers surrogation placement prior to fault happening. It migrates back VM to the original server after system maintenance endowment, if necessary. A fault-tolerant system vastly improves the system availability to enhance the CDC reliability feature.

- System maintenance Provisioning the periodic/dynamic maintenance extends system life time. Surrogation placement technology shifts running application to another host to continue servicing the application during system maintenance time.

- Load balancing Load balancing helps cloud operator to avoid single point of failure by distributing the server workload across several physical hosts within a CDC. Server's workload beyond its capacity degrades system performance; thus, load balancing approaches (using surrogation placement) reduce the possibility of application performance degradation by eliminating subhot-spots within CDC.

- Mobile computing Mobile computing exploits VM migration technology to augment portable computing capabilities. Nowadays, users do not prefer to work on desktop computers only. Rather, they prefer to work on smart phones while they are on the move. VM migration technology helps a user to migrate running applications along OS states from a desktop server to smart phone or vice versa

V Proposed Approach

Surrogation placement process, and thoroughly explained various surrogation placement schemes can improve application and network performance within Cloud Data Centers. It analyzed current surrogation placement schemes based on a thematic taxonomy and highlighted the commonalities and variances among VM migration schemes based on the selected performance parameters. It can solve the issues and challenges in existing VM migration schemes to design an optimal live surrogation placement scheme.

The challenges faced by surrogation placement schemes are very dynamic. Cloud resource heterogeneity, unpredictable workload nature, system workload, server memory size and resource-awareness are computationally inexpensive VM migration schemes. Large-sized VM memory can extend migration time and service downtime. Incorporating optimization methods such as memory contents compression, fine granular duplication, and dynamic write throttling instrumentation can improve application performance at the cost of required system resources. Finally, security is another major threat to the VM migration process and it can be safeguarded by preventing compromised entities' access to VM migration, isolating VM boundaries, and securing the network connections. Efficient placement of the VM can improved the overall performance of the system.

VI Conclusion

In this paper, a brief survey about the virtual machine migration and its importance, kinds of surrogation placement, working principle of pre copy and post copy migration are discussed. This survey explains about the various techniques used to reduce the downtime during virtual machine migration. Finally a proposed aspect and approaches are discussed for virtual machine environments for better effective-

ness. So it is a critical task to choose a technique that is suitable for both the cloud user and cloud provider.

REFERENCES

[1] H. Jin, S. Ibrahim, T. Bell, L. Qi, H. Cao, S. Wu, and X. Shi, *Tools and Technologies for Building Clouds*. London: Springer London, 2010, pp. 3–20. [Online]. Available: http://dx.doi.org/10.1007/978-1-84996-241-4_1

[2] E. M. Guerra and E. Oliveira, *Metadata-Based Frameworks in the Context of Cloud Computing*. London: Springer London, 2013, pp. 3–24. [Online]. Available: http://dx.doi.org/10.1007/978-1-4471-5107-4_1

[3] U. Gurav and R. Shaikh, "Virtualization: A key feature of cloud computing," in *Proceedings of the International Conference and Workshop on Emerging Trends in Technology*, ser. ICWET '10. New York, NY, USA: ACM, 2010, pp. 227–229. [Online]. Available: http://doi.acm.org/10.1145/1741906.1741957

[4] D. M. Leite, M. L. M. Peixoto, B. G. Batista, B. T. Kuehne, and C. H. G. Ferreira, "The influence of resource allocation on cloud computing performance," in *Proceedings of the Symposium on Applied Computing*, ser. SAC '17. New York, NY, USA: ACM, 2017, pp. 1516–1521. [Online]. Available: http://doi.acm.org/10.1145/3019612.3019736

[5] V. A. A. Quirita, G. A. O. P. da Costa, P. N. Happ, R. Q. Feitosa, R. d. S. Ferreira, D. A. B. Oliveira, and A. Plaza, "A new cloud computing architecture for the classification of remote sensing data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 2, pp. 409–416, Feb 2017.

[6] S. K. Panda and P. K. Jana, *An Efficient Resource Allocation Algorithm for IaaS Cloud*. Cham: Springer International Publishing, 2015, pp. 351–355. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-14977-6_37

[7] D. Beimborn, T. Miletzki, and S. Wenzel, "Platform as a service (paas)," *WIRTSCHAFTSINFORMATIK*, vol. 53, no. 6, pp. 371–375, 2011. [Online]. Available: http://dx.doi.org/10.1007/s11576-011-0294-y

[8] R. Sharma and M. Sood, *Cloud SaaS: Models and Transformation*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 305–314. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-24055-3_31

[9] A. Järvi, T. Mäkilä, J. Nissilä, and J. Karttunen, *Tutorial: SaaS Business – Theory and Practice*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 225–225. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-13633-7_26

[10] M. Pathan, R. Buyya, and A. Vakali, *Content Delivery Networks: State of the Art, Insights, and Imperatives*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 3–32. [Online]. Available: https://doi.org/10.1007/978-3-540-77887-5_1

[11] R. Buyya, A. K. Pathan, J. Broberg, and Z. Tari, "A case for peering of content delivery networks," *IEEE Distributed Systems Online*, vol. 7, no. 10, pp. 3–3, Oct 2006.

[12] B. Hu, S. Chen, J. Chen, and Z. Hu, "A mobility-oriented scheme for virtual machine migration in cloud data center network," *IEEE Access*, vol. 4, pp. 8327–8337, 2016.

[13] S. Mazumdar and M. Pranzo, "Power efficient server consolidation for cloud data center," *Future Generation Computer Systems*, vol. 70, pp. 4 – 16, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167739X16308093

[14] D. A. Maltz, "Challenges in cloud scale data centers," *SIGMETRICS Perform. Eval. Rev.*, vol. 41, no. 1, pp. 3–4, Jun. 2013. [Online]. Available: http://doi.acm.org/10.1145/2494232.2465767

[15] M. Dabbagh, B. Hamdaoui, M. Guizani, and A. Rayes, "Energy-efficient resource allocation and provisioning framework for cloud data centers," *IEEE Transactions on Network and Service Management*, vol. 12, no. 3, pp. 377–391, Sept 2015.

[16] X. Lin, Z. Liu, and W. Guo, *Energy-Efficient VM Placement Algorithms for Cloud Data Center*. Cham: Springer International Publishing, 2015, pp. 42–54. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-28430-9_4

[17] J. Bi, H. Yuan, W. Tan, M. Zhou, Y. Fan, J. Zhang, and J. Li, "Application-aware dynamic fine-grained resource provisioning in a virtualized cloud data center," *IEEE Transactions on Automation Science and Engineering*, vol. 14, no. 2, pp. 1172–1184, April 2017.

[18] Z. A. Mann, "Multicore-aware virtual machine placement in cloud data centers," *IEEE Transactions on Computers*, vol. 65, no. 11, pp. 3357–3369, Nov 2016.

[19] M. Bourguiba, K. Haddadou, I. E. Korbi, and G. Pujolle, "Improving network i/o virtualization for cloud computing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 3, pp. 673–681, March 2014.

[20] R. Jain and S. Paul, "Network virtualization and software defined networking for cloud computing: a survey," *IEEE Communications Magazine*, vol. 51, no. 11, pp. 24–31, November 2013.