INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS
ISSN 2320-7345

IMPROVING RESOURCE ALLOCATION FOR DATA CENTER OVERBOOKING

M.Ponmani Bharathi[1], C.Sindhuja[2], S.Vaishnavi[3], Ms.A.Judith Arockia Gladies[4]
[1]*BE (CSE), ULTRA college of Engineering, Madurai, India*
[2]*BE (CSE), ULTRA college of Engineering, Madurai, India*
[3]*BE (CSE), ULTRA college of Engineering, Madurai, India*
[4]*Assisstant Professor, ULTRA college of Engineering, Madurai, India*

Abstract: - Overbooking becomes feasible because user applications tend to overestimate their resources and requirements, that tends to utilize only a fraction of the allocated resources. Overbooking has to be carefully planned in order not to impact application performance. Resource utilization and Data centres utilization can be used in this overbooking scheduler. Data send can send from sources to destination via node. Resource utilization and allocated capacity can be increased by 50% with acceptable performance degradation. Fuzzy logic functions are used to check each overbooking decisions and estimate it. Changing the acceptable level of risk is depending on the current status of the cloud data centres. The suggested approach is extensively evaluated using a combination of simulations and experiments executing real cloud applications with real-life available workloads. Our results show a 50% increment at both resource utilization and capacity allocated with acceptable performance degradation and more stable resource utilization over time.

Keywords*:* Proportional Integral Derivative (PID), Mitigation algorithm, Greedy algorithm

1. Introduction

Authors Data set for overbooking levels. It is collections of some Services and work loading data. The data that represents the collection of fields that will be returned when the data set query runs on the data source. Dataset fields represent the data from a data connection. A field can represent either numeric or non-numeric data. Main features provided by cloud is elasticity, allows users to dynamically adjust resources allocations depending on their current needs. The objective is to make an efficient use of available resources, overestimating the required capacity results in poor resource utilization. Factors contributing to lower the Data Centre Utilization: cloud provides predefined VM Sizes, which have fixed amount of CPU, memory Disk Etc.
A set of distributed PID controllers are implemented to avoid performance degradation and to increase and keep the utilization evenly distributed among the servers. Overbooking addresses the utilization problems that cloud data centres face due to the elastic nature of cloud services. Overbooking has to be carefully planned in order not to impact application performance. It present an overbooking framework that performs admission control decisions based on fuzzy logic risk assessments of each incoming service deployment request. If delay beyond slack on critical path is initiated, then the completion time of the project may get delayed. Resource levelling is a method for smoothing a schedule that attempts to minimize the fluctuations in requirements for resources when the project completion time is fixed. Users are usually bad at estimating the requirements of their applications. This low resource utilization is a big concern for cloud data centred providers as data centres consume lot of energy and are being used in a rather inefficient way. Energy consumption does not decrease

linearly with resource usage. One way cloud providers can mitigate these resource utilization problems is by overbooking. The overbooking techniques always expose the infrastructure to a risk of resource congestion upon unexpected situations and consequently to SLA violations.
This leads to:

- Overestimating the required capacity results in poor resource utilization.
- Lower income from consumers.
- The contrary, underestimating may lead to performance degradation and/or crashes.

Overbooking is to address the utilization problems that cloud data centres face due to the elastic nature of cloud services. Overbooking has to be carefully planned in order not to impact application performance. It present an overbooking framework that performs admission control decisions based on fuzzy logic risk assessments of each incoming service deployment request. A set of distributed PID controllers are implemented to avoid performance degradation and to increase and keep the utilization evenly distributed among the servers.

- Overbooking within cloud data centres to increase resource utilization in a safe and balanced way.
- The cloud paradigm also introduces new obstacles for efficient resource management.
- The very large scale and multi-tenant nature of cloud infrastructures offers great potential for efficient multiplexing of different services.

Our initial work on this problem include scheduling for better server utilization and admission control for capacity planning, getting an initial understanding of the overbooking problem and the risk evaluation, respectively. Cloud applications do not use the same amount of hardware resources all the time. This low resource utilization is a big concern for cloud data centred providers as data centres consume lot of energy and are being used in a rather inefficient way. One way cloud providers can mitigate these resource utilization problems is by overbooking.

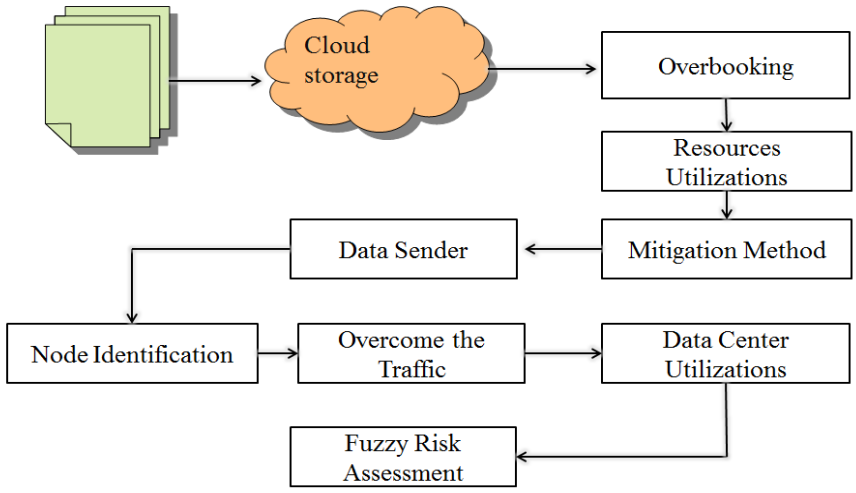

Figure 1 Overbooking Fuzzy Risk Assessment

2. Mitigation Algorithm for Reducing Service Level

Mitigation method is used to avoid sun expected misbehaviors, such as reducing the service level of some services to avoid performance degradation. This Algorithm is used to collocate, reducing the performance degradation when overbooking. This algorithm also clear traffics for data center overbooking utilization. Proportional Integral Derivative (PID) controller is a generic control loop feedback mechanism. PID calculates the differences between the measured and desired set points attempts to minimize it by reading the control input.PID involves three parameter, Present error(P), Accumulated error(I), Prediction of error may occur (D).

1. Data Collection

Dataset for overbooking levels is a collection of some Services and work loading data. The data that representing the collection of fields that will return when the dataset queries runs on the data source. Dataset fields represent the data from a data connection. A field can represent either numeric or non-numeric data.

2. Overbooking

Overbooking is techniques used as a solution to poor resource utilization in cloud data centres. Overbooking is mainly used to handle the data centred resource utilization problems and overbooking. An implemented an autonomic overbooking framework. An autonomic framework that provides better application performance, avoiding over passing total capacity at any of the dimensions will be provided.

3. Resources Utilizations

It determines the shortest project schedule with the limited resources available.

4. Schedule with collocation

It presents a greedy approach that perform traffic-aware VM placement to increase the rate of accepted requests. It avoid repeating poor performance and to increase the chances of good collocations. VMs are suitable to be collocated for improved utilization and stable performance.

5. Prediction method

The prediction step calculates a rough approximation of the desired quantity. The corrector step refines the initial approximation using another means. Overbooking system as well as admission control techniques when dealing with elastic services need insight in future resource usage. Service requirements to avoid performance degradation due to overloaded physical resources.

3. Distribution of PID Controller

PID controlled this fact motivates the use of feedback to adjust the level of risk that the overbooking system is willing to face over time. We also evaluate the distributed controller approach when the data centre size is reduced to 128 cores. Furthermore, choosing an acceptable risk threshold has an impact on data centred utilization and performance. High thresholds result in higher utilization but the expense of exposing the system to performance degradation, whilst using lower values leads to lower but safer resource utilization. When overbooking CPU and I/O capacity, and a more realistic approach for the memory. The rationale for this is that problems resulting from CPU or I/O congestion are less critical than the ones coming from running out of memory. Therefore, the different risk degrees presented can be combined according to the situation, considered capacity dimensions, knowledge about the incoming service, etc. The risk assessment module gets no feedback about the current status and behaviour of the system, the current workload mixture, the data centre size, etc. In order to address this issue, we propose here a control theory approach that dynamically (re)adjusts risk thresholds depending on the system behaviour and the desired utilization n levels, allowing the admission control to learn over time depending on current system behaviour. PID Controller works properly if the performance is measured at the data centre level, obtaining a smooth utilization fluctuations (close enough to the target one) for each congested capacity dimension. However, the utilization of each server may vary from the accumulated utilization – even after applying load balancing techniques. This effect cannot be totally avoided as load imbalance is also caused by the current workload characteristics. To reduce load imbalance we propose a distributed controller approach where each physical server has its own PID controllers, one for each capacity dimension.

4. Related Techniques

Mathematical models for SaaS providers to satisfy customers by leasing Cloud resources from multiple IaaS providers. It proposes three innovative admission control and scheduling algorithms for profit maximization by minimizing cost and maximizing customer satisfaction level. It demonstrates effectiveness of the proposed models and algorithms through an extended evaluation study by varying customer and provider side parameters to analyze which solution suits best in which scenario to maximize SaaS providers profit using actual IaaS data from Amazon and Go Grid. An extensive evaluation to study and analyze which solution suits best in which scenario to maximize SaaS provider's profit. In-house hosting can increase administration and maintenance costs whereas renting from an IaaS provider can impact the service quality due to its variable performance.

Dynamic consolidation of virtual machines (VMs) is an effective way to improve the utilization of resources and energy efficiency in cloud data centres. The problem of host overload detection by maximizing the mean inter migration time under the specified QoS goal based on a Markov chain model. Through simulations with

workload traces from more than a thousand Planet Lab VMs, we show that our approach outperforms the best benchmark algorithm and provides approximately 88 percent of the performance of the optimal offline algorithm. The data center efficiency is been improved and more enterprises are been to consolidate the existing system. All system resources and centralizing resource management allow increasing overall utilization and lowering management costs.

Server consolidation has emerged as a promising technique to reduce the data centre energy cost. We also present a distinguished analysis of an enterprise server workload from the perspective of consolidation and finding characteristics for it. Then observing a significant inherent for power savings if consolidation is performed using off-peak values for application demand. An implementation of the methodologies in a consolidation planning tool and provide a comprehensive evaluation study of the proposed methodologies. The size applications by an off-peak metric and place correlated applications together; there is a high risk of SLA capacity violation. If consolidation is performed by reserving the maximum utilization for each application, the application may require capacity equal to the size of its current entitlements.

As per the size of the cloud increases, the anticipation that all workloads paralyze scale up to their maximum demands. In this observation multiplexing is allowed to access cloud resources among multiple workloads, resource information have been improved. Hosting virtualized loads such that available physical capacity is smaller than the sum of maximal demands of the workloads is referred to as over-commit or over-subscription. It computationally and storage efficiently, while maintaining sufficient accuracy. It is simple method of estimating total effective nominal demand of a cloud and uses it for capacity sizing and placement reservation plan that is compliant with SLA.

5. Conclusion

Overbooking has to be carefully planned in order not to impact application performance. A set of distributed PID controllers are implemented to avoid performance degradation and to increase and keep the utilization evenly distributed among the servers. Feedback control is used to adapt the level of overbooking (risk threshold) that the cloud data centre has tolerating capacity. The utilization technique of data centre is not only increased in overall but also harmonized across hardware capacity dimensions and servers. A set of distributed PID controllers are implemented to avoid performance degradation and to increase and keep the utilization evenly distributed among the servers.

REFERENCES

[1] A. Ali-Eldin, J. Tordsson, and E. Elmroth, “An adaptive hybrid elasticity controller for cloud infrastructures,” in Proc. of Network Operations and Management Symposium (NOMS), 2012, pp. 204–212.

[2] A. Sulistio, K. H. Kim, and R. Buyya, “Managing cancellations and no-shows of reservations with overbooking to increase resource revenue,” in Proc. of Intl. Symposium on Cluster Computing and the Grid (CCGrid), 2008, pp. 267–276.

[3] X. Meng, C. Isci, J. Kephart, L. Zhang, E. Bouillet, and D. Pendarakis,“Efficient resource provisioning in compute clouds via VM multiplexing,” in Proc. Intl. Conference of Autonomic Computing (ICAC), 2010, pp. 11–20.

[4] Y. C. Lee and A. Y. Zomaya, “Energy efficient utilization of resources in cloud computing systems,” The Journal of Supercomputing, vol. 60, no. 2, pp. 268–280, 2012.

[5] L. A. Barroso and U. Holzle, “The case for energy-proportional computing,” Computer, vol. 40, no. 12 pp. 33–37, 2007.

[6] C. Mastroianni, M. Meo, and G. Papuzzo, “Probabilistic consolidationof virtual machines in self-organizing cloud data centers,” IEEE Transactions on Cloud Computing, vol. 1, no. 2, pp. 215–228, 2013.

[7] T. Wo, Q. Sun, B. Li, and C. Hu, “Overbooking-based resource allocation in virtualized data center,” in Proc of 15th IEEE International Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing Workshops (ISORCW), 2012, pp. 142–149.

[8] L. Larsson, D. Henriksson, and E. Elmroth, “Scheduling and monitoring of internally structured services in cloud federations,” in Proc. of IEEE Intl. Symposium on Computers and Communications (ISCC), 2011, pp. 173–178.

[9] D. Breitgand, Z. Dubitzky, A. Epstein, O. Feder, A. Glikson, I. Shapira, and G. Toffetti, “Pulsar: An adaptive utilization accelerator for iaas clouds,” in IEEE International Conference on Cloud Engineering (IC2E), 2014.

[10] M. Dobber, R. van der Mei, and G. Koole, “A prediction method for job runtimes on shared processors: Survey, statistical analysis and new avenues,” Performance Evaluation, vol. 64, no. 7-8, pp. 755–781, 2007.

M. Ponmani Bharathi, currently studying B.E. computer science and engineering in ultra college of Engineering and Technology for women at Madurai

C Sindhuja, currently studying B.E. computer science and engineering in ultra college of Engineering and Technology for women at Madurai

S.Vaishnavi, currently studying B.E. computer science and engineering in ultra college of Engineering and Technology for women at Madurai

Ms.A.Judith Arockia Gladies received her bachelor's degree (B.Tech-Bachelor of Information Technology) from Raja College of engineering and Technology, Madurai, and affiliated to Anna University, Chennai, and then did her Master Degree in computer science and engineering from Raja College Of Engg and Tech, Madurai. She is currently working as an Asst Prof in Ultra College of Engg & Tech for Women, Madurai.