ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, *Vol 8(23), DOI: 10.17485/ijst/2015/v8i23/79313 September 2015* ISSN (Online) : 0974-5645

The Online Control Framework on Computational Optimization of Resource Provisioning in Cloud Environment

N. Jayapandian[1*], A. M. J. Md. Zubair Rahman[2] and J. Gayathri[1]

[1]Knowledge Institute of Technology, Salem - 637504, Tamil Nadu, India;
njayapandian@gmail.com, gayathrikiot@gmail.com
[2]Al-Ameen Engineering College, Erode - 638104, Tamil Nadu, India;
mdzubairrahman@gmail.com

Abstract

The resource monitoring and management system for data center are capable of huge amount of activities. In terms of resource allocation and workload using a variety of K-means clustering algorithm. Workload and machine heterogeneity as well as an optimization problem that is described to see that considers reconfiguration costs. It will solve the problem by using control framework based on computational optimization algorithm. More efficient transfer of energy consumption and delay between each machine by using scheduling algorithm. Energy consumption characteristics of the various machines provide different capacities. Cloud data center are capable of sensing the machine and power consumption. In my proposed using the hybrid method for resource provisioning in data center. This method is used to allocate the resource at the working condition and also energy stored in the power consumption.

Keywords: Cloud Computing, Dynamic Capacity Provisioning, Machine Heterogeneity, Resource Monitoring

1. Introduction

Cloud computing as a service to the community and storage capacity to organize the final recipients and map includes the name of an abstraction of the system to a cloud-shaped symbol is used for critical infrastructure. Cloud computing is a network of a user's data, software and computing services entrusts. Proponents of cloud computing companies to meet the needs of fast-changing, unpredictable business and their applications faster with improved manageability and less maintenance, running with the right to information and allows you to get more evidence[1,3,4].

Data transmissions in the form of a large amount of work on the application of VM execution and the cloud infrastructure is a key one in the datacenter, cloud service provision functions, dealing with the growing demands of the network, there will be a tremendous pressure. Due to network virtualization and resource allocation, a number of leading, communication or data intensive applications, network congestion and peak bandwidth capacity of the network capacity is less problem in virtualized datacenters.

It Merge the attaches and other information system used to implement many of the relevant operators. Clouds are of different potential opportunities and challenges for efficient parallel data processing. Here a user planning and execution was given the task of resource allocation framework is estimated using the first data processing framework. The performance for comparison of the well established in the Hadoop data processing framework. It improves scalability and provides a rule based resource manager for the private cloud and reducing the cost. Also, at that time public cloud and private cloud services to fulfill the request in terms of resource utilization and cost in the hybrid cloud environment, it will evaluating the performance of the resource manager.

* *Author for correspondence*

2. Literature Survey

2.1 Energy and Operational Costs

In proposed system based on the feedback control theory, the system depends on the steady state analysis by using feedback control theory provides a way of addressing the dynamics of the system unstable. It is more responsive to changes in workload the system will allow to control the viscosity. The frequency has per application running on a server that will be able to meet the dynamic response time guarantees for each application to determine the frequency of an assessment and the objective of this sub-problem is to meet the response time at the effectiveness of assessment. In order to provide an estimate the frequency of use of each sub-problem to solve the allocation problems to determines the number of servers[2]. A high frequency of new servers and servers running cost of turning on the issue of the sub-objective is a balance between energy consumption.

2.2 Google Cluster Trace

Analysis of the data processed by the data, although some of the insights we have extracted Google's long-term there is insufficient information or generalizing our findings to other workloads. High-level job scheduler prevents from obtaining an understanding of the semantics of data anonymization. As a result, it is able to evaluate the process and disadvantages of alternative methods[3].

Task execution and allocation times for more information about the type of work the more fine-grained data with a range of items to be addressed. The advantages of the jobs have in common with the actual production of evidence and the needs of the design decisions can be based. However, these results validate the designs have a realistic and representative of the workload which is extractable trace. A qualified research to replicate the target system configuration and raw data without the overhead of a system generic that reproduces the behavior of the original workload and the workload is to create tools for the collection and it is an initial step in this direction of Hadoop Map Reduce workloads.

2.3 Virtual Machines Consolidation

To analyze the performance have immediate problem with this approach, the vector have packing problem is NP hard and therefore the optimal solution (deviation) is difficult. Instead, it is used a less known. Any dimension and capacity dimension of total demand is a lower bound on the rate of dissociation. In this ratio which indicates that all of the dimensions to the maximum. The capable of heuristics to select how many times a particular procedure compare percent overhead[5]. The best hosts normalize to vary widely in the relative trials and it worse experiences the best relative comparison.

The first of each have relevant experience for each set of VM. VM set of inputs that mix together multiple inputs and efficacy of heuristics. That contributes to the variation in the number of VM. VM's entry per set of the same number of tracks each has a set of small sets of reproducing. VM design integration and use of heuristics two related questions. It is how and when it was combined with an account of bottlenecks resource utilization and VM performance assessment to try to find a negative impact on integration. How to solve a variety of dimensions and lack of resource requests that are passed through to the VM for consolidation such as computation and storage[12].

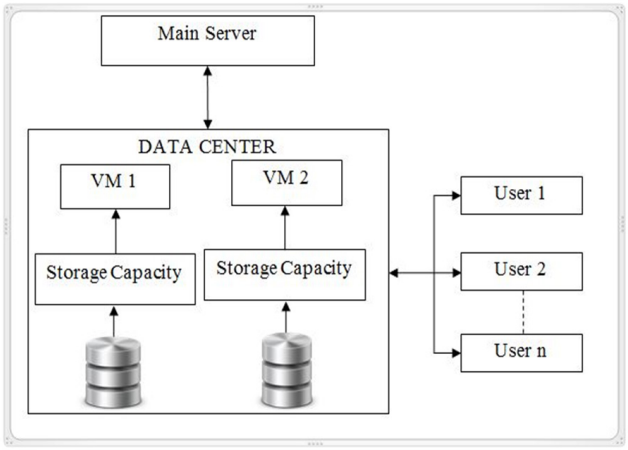

Figure 1. Dynamic capacity provisioning approach.

3. Existing System

The amount of energy in terms of power distribution and cooling, dynamic provisioning of capacity to adjust the number of turns is an approach for reducing energy consumption is a function of the match resource requests. The effectiveness of both the workload and the machine hardware heterogeneity considered the circumstances to provide complete solutions. The energy consumption characteristics of the engine at various levels and workload running in data centers have been typically different priorities and applications have different performance requirements. Due to the energy savings and long-term planning needs and provisioned sub-optimal delays

caused by incompatibility between the two machines will be offered a wide variety of resources as a result of the failure to consider the characteristics of the workload[6–8].

In this dynamic provisioning of a cloud computing environment is the ability to present a diversity-sensitive resource management system. It is used to divide the workload K clustering algorithm such as the characteristics of the various types of system resources performance to do the job[9,10]. It will work in groups of three priority needs and diversity of their work in order to study and understand the term rendered workload heterogeneity.

The Production capacity in the context of dynamic workload provisioning fully considered the heterogeneity of both hardware and machine. This work is an important challenge for the title and body of the workload heterogeneity is addressed. By using a variety of resources on the basis of the workload of the standard K-means clustering, divided into categories such as work shows that the number of assets and performance targets. Machine and workload heterogeneity as well as the optimization problem is described as a framework to consider the costs[11].

It has applied task classification to dynamic capacity provisioning problem in heterogeneous data centers. It is also the problem of heterogeneous data centers capable of changing the classification applied to the task. In production environments the workload and the machine hardware solutions are fully considered the both heterogeneity.

4. Proposed System

Dynamically adjust the number of active machines in the data center. To minimize the total energy consumption in terms of scheduling delay. Hybrid method for resource provisioning in data centers to allocate the resources at the working conditions and also energy stored for the power consumptions.

Dynamic capacity provisioning to create difference machine for stored based on the client demand that is production data centers often comprise heterogeneous machines with different capacities and energy consumption characteristics. Machine heterogeneity the production data centers often comprises several types of machines from multiple update. DCP are heterogeneous processor architectures, hardware features, and memory and disk capacities. Automatically, they have different runtime energy consumption rates.

Production data centers receive a huge number of heterogeneous resource requests such as resource loads, durations, priorities and performance. Both machine and workload in production cloud environments has profound implications on the design of DCP schemes.

Standard K-means clustering, we show that the heterogeneous workload can be divided into multiple task classes with similar characteristics in terms of resource and performance ideas. The workload suggestions consist of scheduling procedures, resource request and usage of records. The job is an application that consists of many tasks. Each task is scheduled on a single physical machine. When a job is submitted, the user can specify the maximum allowed resource demand for each task in terms of required CPU and memory. Suitable number of active machine in a data center for energy consumption while meeting the service level objectives of workloads. The figures correspond to a combination of CPU and memory requirements.

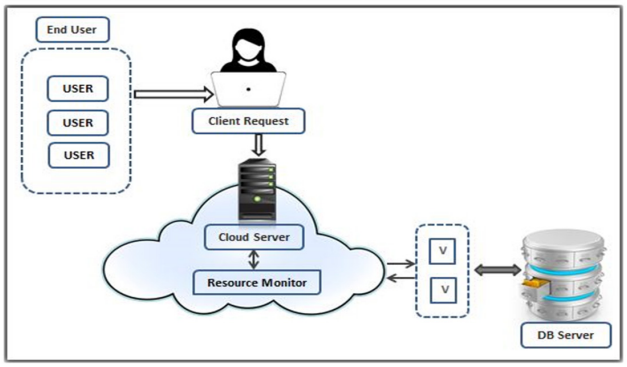

Figure 2. System architecture.

In our proposed algorithm to produce better quality compared to other algorithms. Here we take two normal algorithms such as Computational Optimization Algorithm (COA) and Scheduling Algorithm (SA). We compare our approach to a fixed number of dedicated machines for each cloud. In our proposed algorithm we achieve better energy consumption and low delay.

At this point same type of jobs will be assigned to different machines with different configuration and different algorithm implementation. This case our proposed algorithm will produce more than 30% better energy consumption.

Energy Consumption/Delay	COA	SA
High Energy Low Delay	5.0567	9.575
High Energy High Delay	4.6874	9.574
Low Energy Low Delay	3.8983	6.892
Low Energy High Delay	4.8983	5.321
Energy Consumption	4.635175	7.8405

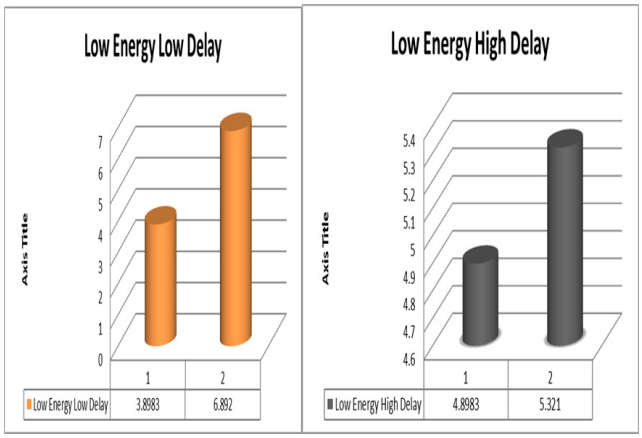

Figure 3. Energy comparison high to low.

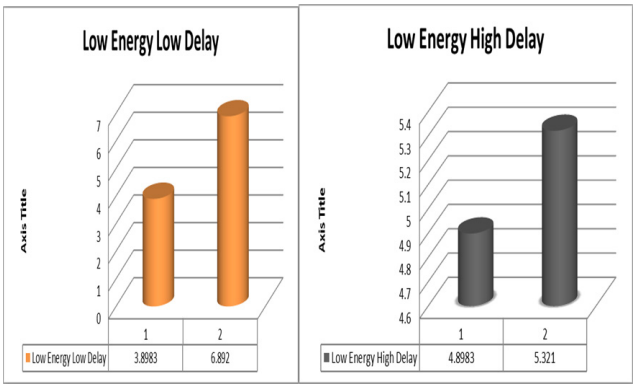

Figure 4. Energy comparison low to high.

5. System Methodology

5.1 Cloud Architecture

A local data center platform, cloud infrastructure and hosted application platform use terms of base load system and application processes running all the time. After removing sporadic spikes in volume do not vary as dramatically as the base load, the application of the resources of the local data center provisioning and service assurance even if the margin is expected to enable the use of a compact and portable. Transient provisioned on demand on the system is expected to be period. A big factor on system resources provisioned.

Hybrid cloud computing model is a result of dynamic workload management, application hosting, public cloud services that customers rely entirely on where the entire cloud computing model addresses many of the concerns.

5.2 Function

To assign all vertexes (data objects) to K (K=2 in our case) disjoint nonempty locations without the expected workload beyond their capacities and achieve minimal partition cost.

$$Min\left[\sum_{n_j \in N_E} C_j + \sum_{v_j \in V_f} S_i\right]$$

Where $\Sigma_{nj \in NE} C_j$ – Cost cut

$\Sigma_{vi \in vf} S_i$ - Sum of quantity of data in the system as opposed to the zone

ϒ–To substitute weights to the elements by a factor of two components.

6. Result and Discussion

The large numbers of progress in user's usage, the large number of problems are in the cloud. The hybrid method is used to allocate the resources with working conditions for overcoming the resource provisioning. It shows the energy consumption with good performance by overcoming the workload. In our proposed system show that the length for the conscious algorithmic rule will increase just about linearly with the quantity of machines, whereas generating cheap utilities. While generating affordable utilities. To estimate the optimality of our approach, we have a tendency to compare these utilities to the perfect utilities generated by the simulated results.

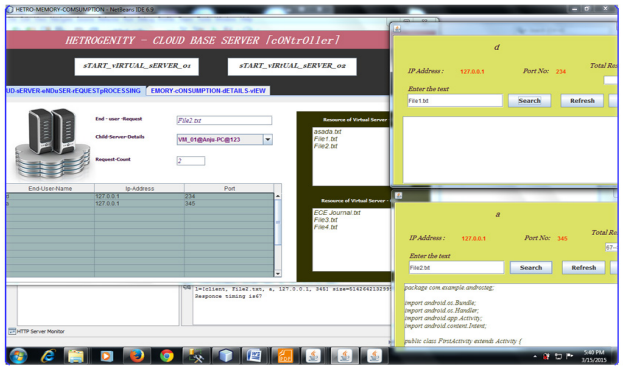

Figure 5.

7. Future Enhancement

The workload model and the failure correlations to redirect user's requests to the appropriate of cloud providers Resource provisioning policies to demonstrate their performance, cost as well as performance cost efficiency. By using hybrid based resource provisioning the cloud. To conduct the actual tests for this purpose to the implementation of the various test methods. The ability to allocate resources for integrating loosely coupled Multi Task Computing (MTC) is another type of application such as going to the local infrastructure, private and public clouds moving VM's resource needs to be another approach to deal with failure in between the machines.

8. References

1. Zhang Q, Zhani MF, Boutaba R, Hellerstein JL. Dynamic Heterogeneity-Aware Resource Provisioning in the Cloud. Proceedings of IEEE International Conference Distributed Computing Systems (ICDCS); 2013.
2. Chen Y, Das A, Qin W, Sivasubramaniam A, Wang Q, Gautam N. Managing Server Energy and Operational Costs in Hosting Centers. ACM SIGMETRICS Performance Evaluation Rev. 2005; 33:303–14.
3. Chen Y, et al. Analysis and Lessons from a Publicly Available Google Cluster Trace; 2010. Technical Report - UCB/EECS- 2010-95.
4. Ghodsi A, Zaharia M, Hindman B, Konwinski A, Shenker S, Stoica, I. Dominant resource fairness: fair allocation of multiple resource types. Proceedings of Eighth USENIX Conference Networked Systems Design and Implementation (USENIX NSDI); 2011.
5. Lee S, Panigrahy R, Prabhakaran V, Ramasubrahmanian V, Talwar K, Uyeda L, Wieder U. Validating heuristics for virtual machines consolidation. Microsoft Research; 2011. MSR-TR-2011-9.
6. Lin M, Wierman A, Andrew L, Thereska E. Dynamic right-sizing for power-proportional data centers. Proceedings of IEEE INFOCOM; 2011.
7. Mishra AK, Hellerstein JL, Cirne W, Das CR. Towards characterizing cloud backend workloads: insights from Google compute clusters. ACM SIGMETRICS Performance Evaluation Rev. 2010 Mar; 37:34–41.
8. Reiss C, Tumanov A, Ganger G, Katz R, Kozuch M. Heterogeneity and dynamicity of clouds at scale: Google Trace Analysis. Proceedings of ACM Symposium Cloud Computing; 2012.
9. Ren S, et al. Provably-Efficient job scheduling for energy and fairness in geographically distributed data centers. Proceedings of IEEE 32nd International Conference Distributed Computing Systems (ICDCS); 2012.
10. Sharma B, Chudnovsky V, Hellerstein JL, Rifaat R, Das CR. Modeling and synthesizing task placement constraints in Google Compute Clusters. Proceedings of Second ACM Symposium Cloud Computing (SOCC); 2011.
11. Verma A, et al. PMAPPER: Power and Migration Cost Aware Application Placement in Virtualized Systems. 2008 Proceedings of Ninth ACM/ IFIP/USENIX International Conference Middleware (Middleware); p. 24364.
12. Zhang Q, Hellerstein J, Boutaba R. Characterizing task usage shapes Google's Compute Clusters. Proceedings of LADIS Workshop Held in Conjunction with VLDB; 2011.