Middle-East Journal of Scientific Research
24 (Special Issue on Innovations in Information, Embedded and Communication Systems): 346-352, 2016
ISSN 1990-9233; © IDOSI Publications, 2016
DOI: 10.5829/idosi.mejsr.2016.24.IIECS.23183

Optimized Resource Provisioning for Dynamic Flow on Cloud Infrastructure Using Meta Heuristic Technique

R. Hari Prashanth and Mrs. S. Pushpalatha

CSE Department, PSNA College of Engg & Tech Dindigul, Tamil Nadu, India

Abstract: The more extensive of pervasive virtual and physical sensors, preeminent up to the Internet of Things, has increased the rate and quantity of data being engendered continuously. Application QoS is also crashed by changeability of resource performance exhibited through clouds and hence necessitates autonomous methods of providing flexible resources to support such applications on cloud infrastructure. Develop the concept of "dynamic dataflows" which handle equivalent alternate tasks as further control over the dataflow's cost and QoS. Developing the application model for dynamic dataflows as well as the infrastructure model for representation of IaaS cloud characteristics and propose an accuraral problem for resource providing that equalize the resource cost, application throughput and the domain value based on user-defined constraints,that Present through a Particle Swarm Optimization (PSO) - based heuristic for deployment and runtime adaptation of continuous dataflows to solve the flawless problem. Also propose efficient greedy heuristics that endure best solution over efficiency, which is critical for low latency streaming applications. elasticity to mitigate the effect on variability, both in input data rates and cloud resource performance, to encounter the correspond QoS of fast data applications.

Key words: Senors · Internet of Things · Cloud Computing · Optimization PSO · Virtual Machines

INTRODUCTION

Cloud computing is said as rendition of computing as a service in place of a product, whereby shared mutual resources, software and information have been afforded to computers and other devices as a quantify service over a network. In computer networking, cloud computing takesplace a large number of computers associated through a communication network such as the Internet, same as service provisioning computing. In science, cloud computing is a equivalent for distributed computing over a network and its aid the capability to run a program or application on many linked computers at the same time and Network based services, that pop up to be furnished through real server hardware and are in fact served up by virtual hardware counterfeited through software running on one or more real machines, is usually called cloud computing, Such virtual servers do not physically endure and can therefore be lifted around and scaled up or down on the speed without affecting the end user, cloud computing becoming larger or smaller without being a physical object.

Dataflow that bridging computational actors in stages that can execute parallely. Data-flow can also be called Stream processing and Receptive programming. The stream processing prototype reduces parallel software and hardware by prohibiting the parallel computation that can be implemented. Given a stream of data, a consequently of operations (*kernel functions*) is enforced to every element in the stream.

Some of the streaming application may consist of you tube, online advertising. These applications are described by the collection of input data streams, each with fluctuating data rates. Further,data appears at high pace and needs to be evaluated with assured low-latency even in the existences of data rate changes.

Approach of dynamic dataflow applications to empower extensible execution on clouds by affording cloud flexibility and also to locate the following problems:

- Datas in response to fluctuations in both input data rates and cloud resource performance.
- Providing applicable and finest trade-offs to balance monetary cost of cloud resources against the users anticipated application value.

Corresponding Author: R. Hari Prashanth, CSE Department, PSNA College of Engg & Tech Dindigul, Tamil nadu, India.

The virtue of data in cloud storage, however, is subject to skepticism and analysis, as data stored in the cloud can reliably be lost or depraved due to the necessary hardware/software defects and human flaw. To make this matter even worse, cloud service providers may be averse to inform users about these data bugs in order to manage the fame of their services and avoid losing profits. Therefore, the virtue of cloud data should be verified before any data usage, such as search or computation over cloud data.

Services Are Classified into Three Main Services Delivery Models: Infrastructure as service (IaaS), Platform as service (PaaS) and software as services (Saas). IaaS refers to the practice of distributing on demand IT infrastructure as a products to customers.PaaS provides a progress platform in which customers can create and execute their own applications. SaaS enrich the user with an unified Service Comprising hardware, development platforms and applications.

Related Works

Scheduling: The scheduling of multitask jobs on clouds is an Non deterministic Polynomial-hard problem have been stated in Figure 2.1. The difficulty lies in the large search space and high overhead of generating efficient schedules, especially for real-time applications with dynamic workloads.

The ordinal of best method is applied in each iteration to achieve sub-optimal schedules. IOO goal is to generate more efficient schedules from a overall perspective over a long period.

In cloud experiments on IBM RC2 cloud, we execute 20,000 tasks in Laser Interferometer Gravitational-wave Observatory verification workflow on 128 virtual machines [1].

The IOO efficient schedule results in a throughput of 1,100 tasks/sec with 7 GB memory demand, compared with 60 percent decrease in throughput and 70 percent increase in memory demand in using the Monte Carlo method [1].

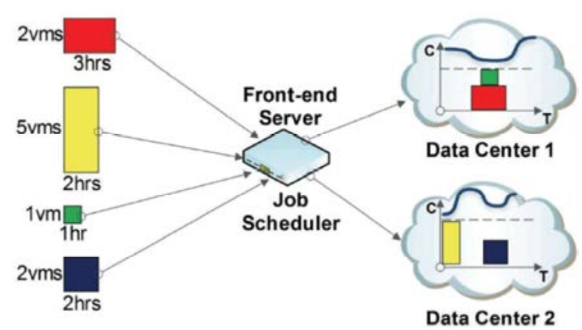

Fig. 2.1: Scheduling

The advantage of IOO method is to adapt sub optimal performance, Monte carlo simulation is used, for throughput and memory increament,adaptability to a scenario with fluctuating workloads [1].

Workflow: Electronic Science features large-scale, compute-intensive workflows of many computing modules that are typically executed in a distributed manner. analytical models are quantify to network performance of scientific workflows using cloud-based computing resources and evaluate a task scheduling problem to minimize the workflow end-to-end delay under a user-specified financial constraint. Workflow is minimized under the user specified financial constraint.This have been done through using the greedy algorithm [2].

Virtual machine (VM) scheduling is an important technique for the efficient operation of the computing resources in a data center have been shown in Figure 2.2. VM assignment and scheduling strategies in which we consider optimizing both the operational price of the data center and the performance degradation of the running applications.

Offline and online solutions for this problem by utilizing the spatial and temporal information of performance interference of Virtual Machine collocation, where Virtual Machine scheduling is performed by jointly considering the combinations and the life-cycle overlap of the VMs.This have been done through Maximum decreasing cost algorithm,Incorparating VM plan online algorithm,VM profit planning algorithm [3].

For Cloud systems,it ignores security requirements of workflow applications and only consider CPU time neglecting other resources like memory, storage capacity. These resource competition could noticeably affect the computation time and total cost of both submitted tasks and their required security services.For this purpose immovable dataset have been used in cloud through Runtime adaption and static assignment of SABA[4].

Isolation: isolation is considered as one of its important features.practical measurements that there exist two kinds of isolation problems in particular virtualized systems, due to cache interference in a multi-core processor.

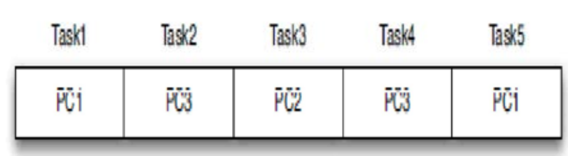

Fig. 2.2: Virtual Machine Scheduling

This problem have been overcome with the help of TSAC. It share some physical resources in a dynamical manner, in order to enforce isolation between the virtual machines without satisfying performance of the virtualized system which have been explained in Figure 2.3. TSAC could improve the performance of the damaged virtual machine by up to about 78% and perform well in blocking its cache-based load information leakage.

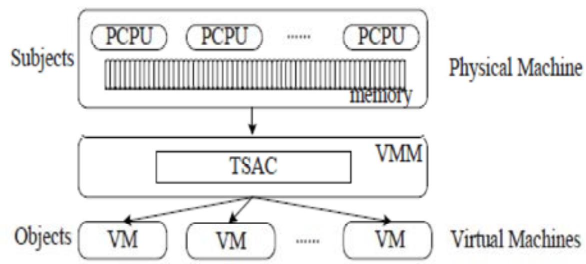

Fig. 2.3: Isolation Using Tsac

Iaas Cost: Infrastructure as a Service clouds provides diverse pricing options, A practical problem facing cloud users is how to reduce their costs by choosing among different pricing options based on their own demands have been stated in Figure 2.4. dynamic strategies for the broker to made instance reservations with the objective of minimizing its service cost.This have been done through heuristic algorithm like greedy and periodic decisions [5]

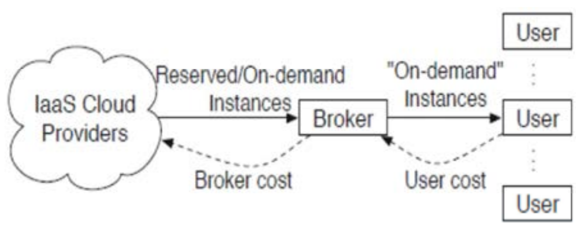

Fig. 2.3: Iaas Cost for Users

Energy Efficiency: Achieving an energy efficiency control and simultaneously satisfying a performance guarantee have become critical issues for cloud providers.The effect of energy-efficiency controls on response times, operating modes and incurred costs. Goal is to find the optimal service rate and mode-switching restriction, so as to minimize cost within a response time guarantee under varying arrival rates.This have been attained through the efficient green control algorithm [6].

Pso Algorithm: Particle Swarm Optimization is a swarm based intelligence algorithm influenced by the social behavior of animals such as a group of birds finding a food source or a school of fish protecting themselves from predator. The movement of each particle is coordinated by a velocity which has both magnitude and direction. Each particle position at any instance of time is influenced by its best position and the position of the best particle in a problem space. The performance of a particle is measured by a optimistic value, which is problem specific.The PSO algorithm is similar to other evolutionary algorithms.

In Particle Swarm Optimization, the population is the number of particles in a problem space. Particles are initialized randomly.

Each particle will have a optimistic value, which will be evaluated by a fitness function to be best in each generation. Each particle knows its best position pbest and the best position so far among the entire group of particles

Fig. 3.1: Particle Swarm Optimization

- Uses a number of particles that constitute a swarm moving around in the search space looking for the best result.
- Each particle in search space adjusts its “flying” according to its own flying experience as well as the flying experience of other particles

Concept: Each particle adjusts its travelling speed dynamically corresponding to the flying experiences of itself and its colleagues.

Each particle modifies its position according to:

- Its current position
- Its current velocity
- The distance between its current position and <u>*pbest*</u>
- The distance between its current position and <u>*gbest*</u>

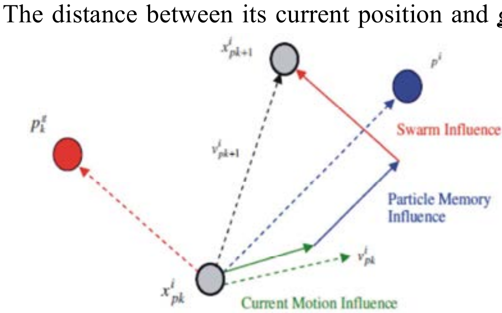

Fig. 3.2: Pso Process

Algorithm Parameters:

Algorithm parameters
A: Population of agents
P_i: Position of agent a_i in the solution space
f: Objective function
v_i: Velocity of agent's a_i
$V(a_i)$: Neighborhood of agent a_i (fixed)

```
[x*] = PSO()
P = Particle_Initialization();
For i=1 to it_max
    For each particle p in P do
        fp = f(p);
            If fp is better than f(pBest)
                pBest = p;
            end
        end
        gBest = best p in P;
        For each particle p in P do
        v = v + c1*rand*(pBest – p) + c2*rand*
        (gBest – p);
            p=p+v;
        end
    end
```

Steps Involved

- Set particle dimension as equal to the size of ready tasks in {ti} ? T
- Initialize particles position randomly from PC =1,..., j and velocity vi randomly.
- For each particle, calculate its fitness value
- If the fitness value is better than the previous best pbest,set the current fitness value as the new pbest.
- After Steps 3 and 4 for all particles, select the best particle as gbest.
- For all particles, calculate velocity and update their positions
- If 1)the stopping criteria or maximum iteration is not satisfied,repeat from Step 3.

Advantages

- Insensitive to scaling of design variables
- Simple implementation
- Easily parallelized for concurrent processing
- Derivative free
- Very few algorithm parameters
- Very efficient global search algorithm.

Proposed Work: Develop the application model for dynamic dataflows as well as the infrastructure model to represent IaaS cloud characteristics and propose an optimization problem for resource provisioning that balances the resource cost, application throughput and the domain value based on user-defined constraints [4].

Present a Particle Swarm Optimization (PSO) based heuristic for deployment and runtime adaptation of continuous dataflows to solve the optimization problem. Also propose efficient greedy heuristics (centralized and sharded variants) that sacrifice optimality over efficiency, which is critical for low latency streaming applications.

Benchmark: Benchmarking allows users and designers of databases to compare performance between different software and hardware systems. Comparisons are important both for researchers evaluating designs and for consumers deploying systems. A good benchmark will greatly simplify a consumers choice of database system and will help designers to build the kinds of systems that consumers desire [1].

It must measure the peak performance and price/performance of systems when performing typical operations within that problem domain.The purpose of the benchmark is to determine the maximum scale factor at which a stream data management system can run while still meeting the latency and accuracy requirements specified herein.

Linear Road Benchmark

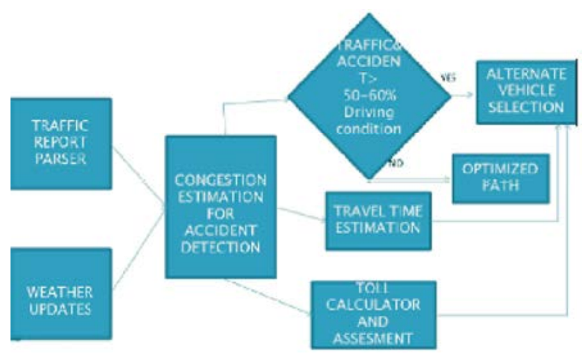

Fig. 4.1: Workflow for Lrb

LRB models a road toll network within a confined area,in which the toll depends on several factors including time of the day, current traffic congestion levels and proximity to accidents. It continuously takes "position reports" from different vehicles on the road and is responsible for

- Detecting average speed and traffic congestion for a section
- Detecting accidents
- Providing toll notifications to vehicles whenever they enter a newsection
- Answering account balance queries and toll assessments
- Estimating travel times between two sections.

The Aim is to support the highest number of expressways while satisfying the desired latency constraints,work flow for Lrb have been explained in Figure 4.1

Traffic Report Parser

Position report consist of
time =when report has been created
VID =Vehicle identification number
speed=Speed may be in random manner(0 to 100)
Lane=which lane the vehicle have been travelling.
This have been monitored using road side unit.

Weather Updates

Weather normal=40^0c +
Summer =40^0c
Winter =20^0c
Mist =10^0c

Weather have been checked with using accurateweather. com. dividing the task of weather into two condition as normal and not normal.

Sevral condition have been given for vehicles such as speed limit in which it needs to be travelled at each condition.if it fails,there may chance for occurring in congestion.at that time,it will check in traffic report parser.

Congestion Estimation: Congestion estimation is based on weather updates and Traffic report parser.while the condition fails in both the taffic parser and weather updates,there may occur in congestion.

If(NO.OF.VEHICLES>100&&SPEED LIMIT>50Kmph&&DISTANCE<=1000 Mts)congestion=true

In such case,checking of vehicle condition for alternate selection.Alternate selection is based on cost,Travel time and Damage of vehicle.By comparing the allocated and required resources alternates have been replaced.

Toll Calculator and Assesment: Toll calculator is used to store and retrieve vehicle information in case of alternate selection,it keeps the whole information about the each vehicle.scheduling may takes place at this place and also updates the vehicle information,whether the vehicle has balance in any tolls or already paid.

Toll management is done in the virtual machine.virtual machine acts as a centralized server and keep activates the run time data through data center.Each virtual machine is provided with specific bANDWIDTH,MIPS, RAM and CPU for completion of its efficient tasks.

ARCHITECHTURE:

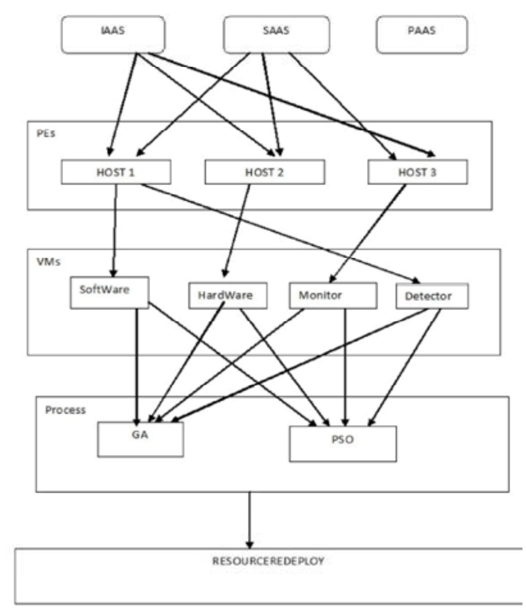

Fig. 4.2: Architechture for Lrb

Cloud Simulator: CloudSim is a framework for modeling and simulation of cloud computing infrastructures and services. CloudSim itself does not have a graphical user interface, extensions such as Cloud Reports offer a GUI for CloudSim.

This application uses the cloudsim as simulation tool for comparison of the genetic algorithm and particle swarm optimization algorithm.

RESULT

Comparision have been done based on the proposed Particle Swarm Optimization Algorithm and the Genetic Algorithm.We evaluate their overall monetary cost of execution(5.1),Time for execution(5.2), overall resource utilized over the optimization period(5.3).

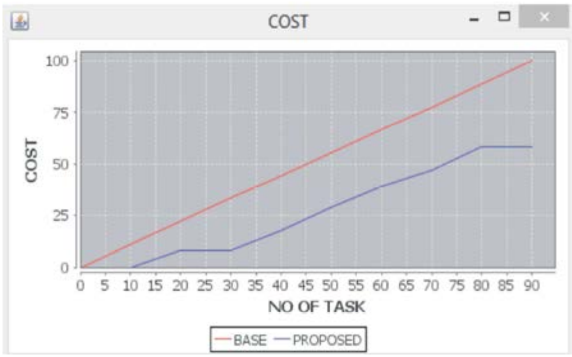

Fig. 5.1: Cost Comparision

The main advantage over this is population.while comparing population,Proposed algorithm can access random population while genetic algorithm cannot.genetic algorithm population is based on mutation.the application provides the availability over the streaming.the necessary constraint must be met but higher values beyond the constraint do not indicate a better algorithm.

Particle swarm optimization algorithm is better than Genetic --algorithm while it meets the necessary relative application throughput constraint and has a higher value for the objective function.

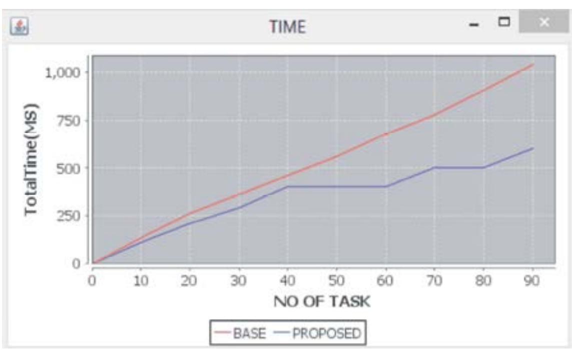

Fig. 5.2: Time Comparision

Similarly, an algorithm with a higher value is not better unless it also meets the constraint.The Figure for the streaming application stated in the paper.

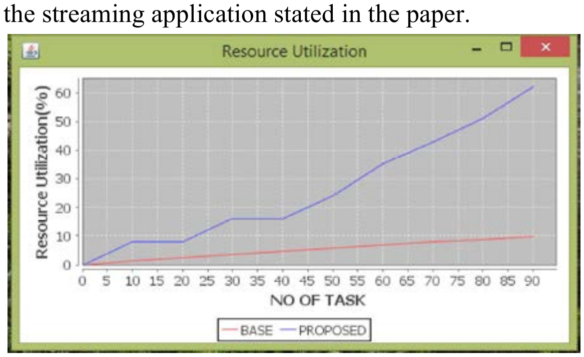

Fig. 5.3: Resourse Comparision

This Figures shows that output for streaming application through streaming application.In Genetic algorithm there is a chance for isolation problem but in PSO,optimized scheduling is done,which takes place major role in reducing cost,throughput and scalability,that may increase the quality of service parameters.

CONCLUSION

Fast dynamic multitask workload schedule in a cloud computing platform and dynamic dataflow,Idea behind that have been done by Particle Swarm Optimization algorithm.It is most famous algorithm for scheduling tasks,It is used to generate initial population in order to minimize makespan. The option for choosing of particle swarm optimization parameters can have a large impact on optimization performance. Selecting PSO parameters can yield good performance.Thus the survey has been done through workflow and scheduling of dynamic dataflows. Resource provisioning process when faced with drastically increasing workload and hence can ensure a high QoS.

REFERENCES

1. Fan Zhang, Junwei Cao, Kai Hwang, Keqin Li and Samee U. Khan, 2015. Adaptive Workflow Scheduling on CloudComputing Platforms with Iterative Ordinal Optimization, IEEE Transaction on cloud computing,APR-JUN.
2. Chase Qishi Wu, Xiangyu Lin, Dantong Yu, Wei Xu and Li Li, 2015. End-to-End Delay Minimization for ScientificWorkflows in Clouds under Budget Constraint" IEEE Transaction of cloud computing, APR-JUN.
3. Xibo Jin, Fa Zhang, Lin Wang, Songlin Hu, Biyu Zhou1 and Zhiyong Liu1, 0000. Joint Optimization of Operational Cost and Performance Interference in Cloud Data Centers, IEEE Transaction of cloud computing (JOURNAL ARTICLE)
4. Keqin Li, 2015. Improving Multicore Server Performance and Reducing Energy Consumption by Workload Dependent Dynamic Power Management, IEEE Transaction of cloud computing.
5. Jinzhao Liu, Yaoxue Zhang, Yuezhi Zhou, Di Zhang and Hao Liu, 2015. Aggressive Resource Provisioning for Ensuring QoS in Virtualized Environments, IEEE Transaction on CLOUD COMPUTING.

6. Lingfang Zenga, Bharadwaj Veeravalli a and Xiaorong Li b, 2014. SABA: A security-aware and budget-aware workflow scheduling strategy in clouds, Journal of parallel and Distributed computing.