IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 05, Issue 10 (October. 2015), ||V2|| PP 46-50

Combined Resource provisioning and Scheduling strategy for executing scientific Workflows on IaaS Clouds

Chetana Pradip Shravage, Dr. S.T. Singh
(Computer Department, P.K.Technical Campus/Savitribai Phule Pune University, India)
(Computer Department, P.K.Technical Campus/Savitribai Phule Pune University, India)

Abstract : Cloud computing is that the latest distributed computing model and it offers big opportunities to resolve large-scale scientific issues. However, it presents varied challenges that require to be addressed so as to be with efficiency utilized for progress applications. Although the advancement programing downside has been wide studied, there area unit only a few initiatives tailored for cloud environments. Furthermore, the present works fail to either meet the user's quality of service (QOS) needs or to include some basic principles of cloud computing like the physical property and no uniformity of the computing resources. This paper proposes a resource provisioning and programing strategy for scientific workflows on Infrastructure as a Service (IaaS) clouds. we tend to gift associate algorithm supported the meta-heuristic improvement technique, particle swarm improvement (PSO), that aims to reduce the general workflow execution value whereas meeting point in time constraints. Our heuristic is evaluated victimization CloudSim and numerous well-known scientific workflows of various sizes. The results show that our approach performs higher than the present progressive algorithms.

Keywords: *- Cloud computing, resource provisioning, scheduling, scheduling strategy, scientific workflow, Iaas*

I. INTRODUCTION

Cloud computing is a recently evolved computing terminology or metaphor based on utility and consumption of computing resources. Cloud computing involves deploying groups of remote servers and software networked that allow centralized data storage and online access to computer services or resources. Clouds can be classified as public, private or hybrid.The criticisms about it are mainly focused on its social implications. This happens when the owner of the remote servers is a person or organization other than the user, as their interests may point in different directions, for example, the user may wish that his or her information is kept private, but the owner of the remote servers may want to take advantage of it for their own business.

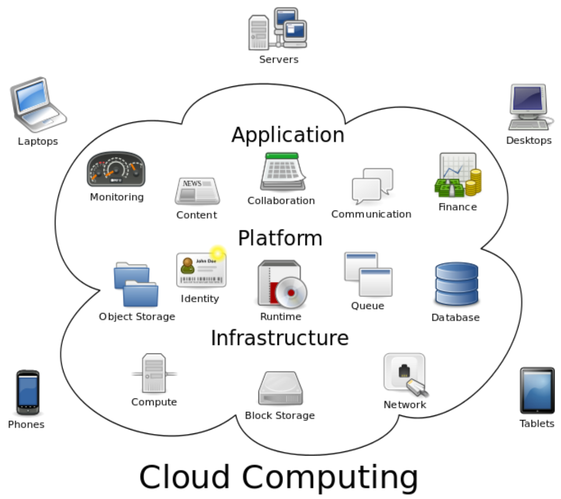

Fig 1: Cloud computing metaphor: For a user, the network elements representing the provider-rendered services are invisible, as if obscured by a cloud.

Service models-Cloud computing providers offer their services according to several fundamental models-Infrastructure as a service (IaaS)

In the most basic cloud-service model & according to the IETF (Internet Engineering Task Force), providers of IaaS offer computers – physical or (more often) virtual machines – and other resources. (A hypervisor, such as Xen, Oracle Virtual Box, KVM, VMware ESX/ESXi, or Hyper-V runs the virtual machines as guests. Pools of hypervisors within the cloud operational support-system can support large numbers of virtual machines and the ability to scale services up and down according to customers' varying requirements.) IaaS clouds often offer additional resources such as a virtual-machine disk image library, raw block storage, and file or object storage, firewalls, load balancers, IP addresses, virtual local area networks (VLANs), and software bundles.IaaS-cloud providers supply these resources on-demand from their large pools installed in data centers. For wide-area connectivity, customers can use either the Internet or carrier clouds(dedicated virtual private networks).

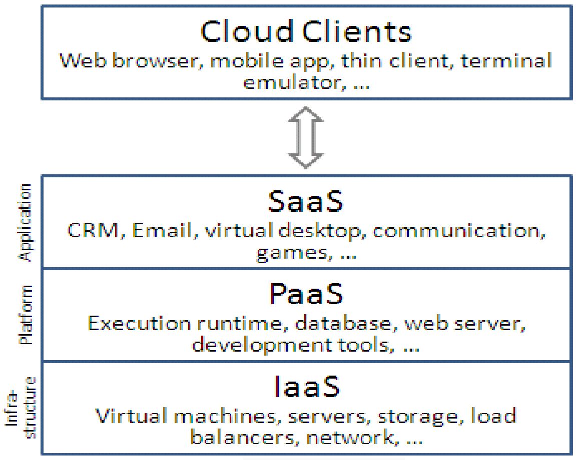

Fig 2: Service models

Workflows are often accustomed model large-scale scientific issues in areas like bioinformatics, astronomy, and physics. Such scientific workflows have ever-growing information and computing needs and therefore demand a superior computing surroundings in order to be dead during a affordable quantity of time. These workflows are unremarkably shapely as a collection of tasks interconnected via information or computing dependencies.

The orchestration of those tasks onto distributed resources has been studied extensively over the years, that specialize in environments like grids and clusters. However, with the emergence of latest paradigms like cloud computing, novel approaches that address the actual challenges and opportunities of those technologies ought to be developed.

As a years passes, distributed environments have evolved from shared community platforms to utility-based models; the latest of those being cloud computing. This technology enables the delivery of IT resources over the web, and follows a pay-as-you-go model wherever users ar charged based on their consumption. There ar numerous forms of cloud suppliers, every of that has completely different product offerings. they're classified into a hierarchy of as-a-service terms: package as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS). This paper focuses on IaaS clouds which provide the user a virtual pool of unlimited, heterogeneous resources that may be accessed on demand. Moreover, they provide the pliability of elastically acquiring or cathartic resources with varied configurations to best suit the necessities of associate degree application. Even though this empowers the users and provides them additional management over the resources, it additionally dictates the event of innovative programming techniques in order that the distributed resources ar with efficiency utilised.

There square measure 2 main stages once designing the execution of a work flow in a very cloud setting. the primary one is that the resource provisioning phase; throughout this stage, the computing resources which will be accustomed run the tasks square measure designated and provisioned. within the second stage, a schedule is generated and each task is mapped onto the best-suited resource.The selection of the resources and mapping of the tasks is done in order that completely different user outlined quality of service (QoS) requirements square measure met. Previous works during this space, especially those developed for Grids or Clusters, centered principally on the planning section. the rationale behind this can be that these environments offer a static pool of resources that square measure readily accessible to execute the tasks and whose configuration is known prior to. Since this can be not the case in cloud environments, each issues got to be addressed and combined so as to supply associate economical execution set up.

Another characteristic of previous works developed for clusters and grids is their specialize in meeting application deadlines or minimizing the make span (total execution time) of the work flow whereas ignoring the price of the utilized infrastructure. while this can be well matched for such environments, policies developed for clouds square measure duty-bound to contemplate the pay-per-use model of the infrastructure so as to avoid preventative and supererogatory prices.

II. PROBLEM DEFINATION

There are two main stages when planning the execution of a workflow in a cloud environment. The first one is the resource provisioning phase; during this stage, the computing resources that will be used to run the tasks are selected and provisioned. In the second stage, a schedule is generated and each task is mapped onto the best-suited resource. The selection of the resources and mapping of the tasks is done so that different user defined quality of service (QoS) requirements are met. The selection of initial resource pool as it has a significant impact on the performance of a cloud computing environment.

III. EXISTING SYSTEM

3.1 Particle swarm optimization-

The rule could be a random optimization technique within which the foremost basic conception is that of particle. To find out the position of an particle.

ALGORITHM 1
PARTICLE SWARM OPTIMIZATION

1. Set the dimension of the particles to d
2. Initialize the population of particles with random positions and velocities
3. For each particle, calculate its fitness value
   - **3.1** Compare the particle's fitness value with the particle's $pbest$. If the current value is better than $pbest$ then set $pbest$ to the current value and location
   - **3.2** Compare the particle's fitness value with the global best $gbest$. If the particle's current value is better than $gbest$ then set $gbest$ to the current value and location
   - **3.3** Update the position and velocity of the particle according to equations 7 and 8
4. Repeat from step 3 until the stopping criterion is met.

3.2 Schedule generation-

The pseudo-code to convert a particle's position into a Schedule. Initially, the set of resources to lease R and the set of task to resource mappings M are empty and the total execution cost TECand time TET are set to zero.

ALGORITHM 2
SCHEDULE GENERATION

Input: Set of workflow tasks T
Initial resource pool $\mathcal{R}_{initial}$
An array $pos[|T|]$ representing a particle's position
Output: A Schedule S

1. Initialize schedule components
- **1.1.** $\mathcal{R} = \emptyset, \mathcal{M} = \emptyset$
- **1.2.** $TEC = 0, TET = 0$

2. Calculate $ExeTime[|T| \times |R_{initial}|]$
3. Calculate $TransferTime[|T| \times |T|]$
4. **for** $i = 0$ to $i = |T| - 1$
- **4.1.** $t_i = T[i]$, $r_{pos[i]} = \mathcal{R}_{initial}[pos[i]]$
- **4.2.** **if** t_i has no parents
  $ST_{t_i} = LET_{r_{pos[i]}}$
  else
  $ST_{t_i} = max\left(max\left\{ET_{t_p} : t_p \in parents(t_i)\right\}, LET_{r_{pos[i]}}\right)$
  end if
- **4.3.** $exe = exeTime[i][pos[i]]$
- **4.4.** **for each** child t_c of t_i
  if t_c is mapped to a resource different to $r_{pos[i]}$
  $transfer += TransferTime[i][c]$
  end if
  end for each
- **4.5.** $PT_{t_i}^{r_{pos[j]}} = exe + transfer$
- **4.6.** $ET_{t_i} = PT_{t_i}^{r_{pos[j]}} - ST_{t_i}$
- **4.7.** $m_{t_i}^{r_{pos[i]}} = (t_i, r_{pos[i]}, ST_{t_i}, ET_{t_i})$
- **4.8.** $M = M \cup \{m_{t_i}^{r_{pos[i]}}\}$
- **4.9.** **if** $r_{pos[i]} \notin R$
  $LST_{r_{pos[i]}} = max(ST_{t_i}, bootTime)$
  $R = R \cup \{r_{pos[i]}\}$
  end if
- **4.10.** $LET_{r_{pos[i]}} = PT_{t_i}^{r_{pos[j]}} + LST_{r_{pos[i]}}$

5. Calculate *TEC* according to equation (4)
6. Calculate *TET* according to equation (5)
7. $S = (R, M, TEC, TET)$

IV. PROPOSED SYSTEM

4.1 Genetic Algorithm-

Genetic Algorithms (GAs) is adaptive heuristic search algorithmic program supported the organic process concepts of action and genetic science. Associate in Nursing of itself} they represent an intelligent exploitation of a random search accustomed solve optimization issues. Though irregular, GAs ar by no suggests that random, instead they exploit historical info to direct the search into the region of higher performance among the search house. The essential techniques of the GAs ar designed to simulate processes in natural systems necessary for evolution, particularly those follow the principles initial set down by Charles Robert Darwin of "survival of the fittest.". Since in nature, competition among people for scanty resources leads to the fittest people dominating over the weaker ones

Genetic algorithms belong to the larger category of organic process algorithms (EA), that generate solutions to optimisation issues victimisation techniques impressed by natural evolution, like inheritance, mutation, selection, and crossover. GAs simulate the survival of the fittest among people over consecutive generation for resolution a tangle. every generation consists of a population of character strings that area unit analogous to the body that we have a tendency to see in our deoxyribonucleic acid. every individual represents a degree during a search area and a potential resolution. The people within the population area unit then created to travel through a method of evolution.

GAs area unit supported AN analogy with the genetic structure and behavior of chromosomes inside a population of people exploitation the subsequent foundations:

- Individuals during a population contend for resources and mates.
- Those people most booming in every 'competition' can turn out additional offspring than those people that perform poorly.
- Genes from `good' people propagate throughout the population in order that 2 smart oldsters can generally turn out offspring that area unit higher than either parent.
- Thus every serial generation can become additional suited to their atmosphere. Based on action Implementation details-After associate initial population is indiscriminately generated, the algorithmic rule evolves the through 3 operators:
- selection that equates to survival of the fittest;
- crossover that represents pairing between individuals;
- mutation that introduces random modifications.

V. CONCLUSION

In this paper I presented a combined resource provisioning and scheduling strategy for executing scientific workflows on IaaS clouds. The scenario was modeled as an optimization problem which aims to minimize the overall execution cost while meeting a user defined deadline and was solved using the meta-heuristic optimization algorithm, PSO. The proposed approach incorporates basic IaaS cloud Principles such as a pay-as-you-go model, heterogeneity, elasticity, and dynamicity of the resources. Furthermore, our solution considers other characteristics typical of IaaS platforms such as performance variation and VM boot time. Also I will use genetic algorithm for optimization strategies.

VI. ACKNOWLEDGEMENTS

I would like to thank my Guide Prof. Dr. S.T. Singh who give me their valuable time for their support, guidance and encouragement.

REFERENCES

Journal Papers:

[1] G. Juve, A. Chervenak, E. Deelman, S. Bharathi, G. Mehta, and K. Vahi, “Characterizing and profiling scientific workflows,” *Future Generation Comput. Syst., vol. 29, no. 3*, pp. 682–692, 2012.

[2] P. Mell, T. Grance, “The NIST definition of cloud computing— recommendations of the National Institute of Standards and Technology” *Special Publication* 800-145, NIST, Gaithersburg, 2011.

[3] E. K. Byun, Y. S. Kee, J. S. Kim, and S. Maeng, “Cost optimized provisioning of elastic resources for application workflows,” *Future Generation Comput. Syst., vol. 27, no. 8*, pp. 1011–1026, 2011.

Books:

[4] A. Lazinica, Ed. Particle Swarm Optimization. Rijeka, Croatia: InTech, 2009.

[5] Maria Alejandra Rodriguez and Rajkumar Buyya , “Deadline Based Resource Provisioning and Scheduling”.

Proceedings Papers:

[6] Kennedy and R. Eberhart, "Particle swarm optimization*," in Proc. 6th IEEE Int. Conf. Neural Netw.*, 1995, pp. 1942–1948.
[7] M. Mao and M. Humphrey, "A performance study on the VM startup time in the cloud," *in Proc. 5th IEEE Int. Conf. Cloud Comput*., Jun. 2012, pp. 423–430.
[8] Maria Alejandra Rodriguez and Rajkumar Buyya, "Deadline based Resource provisioning and Scheduling Algorithm for scientific workflows on clouds*", IEEE Transaction on cloud computing, vol. 2, no. 2,* pp. 222-235, April-June 2014.
[9] S. Pandey, L. Wu, S. M. Guru, and R. Buyya, "A particle swarm optimization-based heuristic for scheduling workflow applications in cloud computing environments," *in Proc. IEEE Int. Conf. Adv. Inform. Netw. Appl., 2010*, pp. 400–407.
[10] Z. Wu, Z. Ni, L. Gu, and X. Liu, "A revised discrete particle swarm optimization for cloud workflow scheduling*," in Proc. IEEE Int. Conf. Comput. Intell. Security, 2010*, pp. 184–188.