ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

A Study Paper on Prediciting Workload with the Help of ARIMA Model

Rohini Mahajan[1], Prof. Rugraj Purohit[2]

ME Student, Dept. of Computer Engineering, Alard College of Engineering and Management.

Savitribai Phule Pune University, Pune, India[1]

Professor, Dept. of Computer Engineering and Management, Savitribai Phule Pune University, Pune, India [2]

ABSTRACT: Presently company mostly prefers Cloud services to store their data. There also competition between the Cloud Service Providers for offering similar services. To survive in such a competitive market, cloud-based firms shouldstrive to achieve good quality of service (QoS) for their users, without losing their customers. However, to realize the QoS with a less price is difficult as a result of workloads varies over time. This downside will be solved by victimization the proactive dynamic provisioning of resources. They forecast the long run want of applications in terms of resources and allot the required resources before and emotional them once they're not needed. It shows the conclusion of a cloud employment prediction module for SaaS suppliers supported the autoregressive integrated moving average (ARIMA) model. It inserts the prediction supported the ARIMA model and evaluates its accuracy of future employment prediction victimization real traces of requests to internet servers.

KEYWORDS: Cloud computing; Workload prediction; ARIMA

I. INTRODUCTION

Cloud computing had evolved from promising virtualization and data centre technologies to a consolidated paradigm to delivery of service to its end customers. This customer'spay for such services according to their use, for e.g utilities such as electricity, gas, and water.The technology is being adopted by enterprises is growing at faster rate, and hence the number of cloud-based companies offering cloud-based solutions for end users is also increasing. The shifting of desktop applications to public cloud hosted software as a service (SaaS) business model has intensified the competition for cloud providers. The reason behind this is the presence of multiplenumberof providers in the current cloud computing landscape which offer services under heterogeneous configurations. Selecting particular cloud service configuration (for e.g., Virtual Machine type, Virtual Machine cores, speed, cost, and their location) get translated to a certain level of QoS in terms of, acceptance rate, response time , reliability, etc.To survive in a competitive market, acceptable QoS must be provided by cloud providers to end-users of the hosted SaaS applications, or they might lose the customers. However, a issue that is aroused from the transition to a SaaS model is the fact that the pattern of access to the application varies according to time of the day, day of the week, and part of a year. It means that in some period of time, there are many users trying to use the service concurrently, whereas in others only a few users are concurrently accessing the servers. Thus static allocation of resources to the SaaS application becomes ineffective, when there is low demand there will be plenty of resources available, causing unnecessary cost for the application provider, in other case, during high utilization periods of the available resources, they may be insufficient, which will lead to poor Quality of Service and resulting in losing of costumers and money. For above problem Cloud provide dynamic provisioning which determines how many requests per second are expected in the near future will be combined with analytical models so that they the optimal number of resources can be determined in the presence of the predicted load. The need for workload prediction was recognized by proposed architecture; but it didn't propose a concrete method to predict workload. Thus, in it is presenting the design and realization of its workload prediction model using the ARIMA model that is autoregressive integrated moving average model. ARIMA is a method for non-stationary time-series prediction that is composed of an autoregressive and a moving average model, and was successfully utilized for time-series prediction in different domains such as finance.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

II. RELATED WORK

R. N. Calheiros, R. Ranjan, and R. Buyya [1] described a provisioning technique that will automatically adapt to workload changes and offered end users guaranteed Quality of Services (QoS) in large, autonomous.
To improve the efficiency of the system, authors had used queueing network system model and workload information to supplement intelligent input about system requirements to an application provisioner with limited information about the physical infrastructure.

N. Roy, A. Dubey, and A. Gokhale [2] described three contributions to overcome the general lack of effective techniques for workload forecasting and optimal resource allocation.
1) Discussed the challenges involved in auto scaling in the cloud.
2) Developed a model-predictive algorithm for workload forecasting that is used for resource auto scaling.
3) Empirical results were provided that showed resources can be allocated and deallocated by their algorithm.

V. G. Tran, V. Debusschere, and S. Bacha [3] A good server workload methodology for the server workload forecasting is being provided by seasonal ARIMA. A large set of their experiments confirmed that it had high performance, scalability and reliability and will be integrated in the system. They presented a general expression in development of forecast model.

V. Nae, A. Iosup, and R. Prodan[4] proposed a dynamic provisioning algorithm where the resource pool of a MMOG service can be resized to adapt to workload variability and maintain a response time below a given threshold. They had used a Queuing Network performance model to estimate the system response time for different configurations quickly. The performance model is used within a greedy algorithm.

S. Pacheco-Sanchez, G. Casale, B. Scotney, S. McClean, G. Parr, and S. Dawson [5] demonstrated the MAP/MAP/1 queue is a useful and versatile tool for performance prediction of web servers deployed in cloud, which is significantly faster solving it analytically rather than by simulation.

E. Caron, F. Desprez, and A. Muresan [6] proposed an approach to solve the problem of workload prediction by identifying similar past occurrences to the current short-tbbbbbbberm workload history. As well as presented the auto-scaling algorithm.

Q. Zhu and G. Agrawal [7] focused on the use of cloud resources for a class of adaptive applications. Framework was designed, implemented, and evaluated that supported dynamic adaptation for applications in a cloud computing environment. The key component of the framework was a multi-input-multi-output feedback control model-based dynamic resource provisioning algorithm.

III. EXISTING SYSTEM

In existing system, the resources were allocated statically to the SaaS application. But this was ineffective. The solution to this problem is dynamic allocating of resources by predicting workload which determines how many requests per second are expected in the near future combined with analytical models to determine the optimal number of resources in the presence of the predicted load. The need for workload prediction was recognized by proposed architecture; but it didn't propose a concrete method to predict workload.

A. DISADVANTAGES OF EXISTING SYSTEM

- During a low period demand there was extra number of resources available which incurring unnecessary cost.
- During a high utilization periods the availability of resources might be insufficient, so it provides poor service quality and losing of costumers.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

IV. SYSTEM ARCHITECTURE

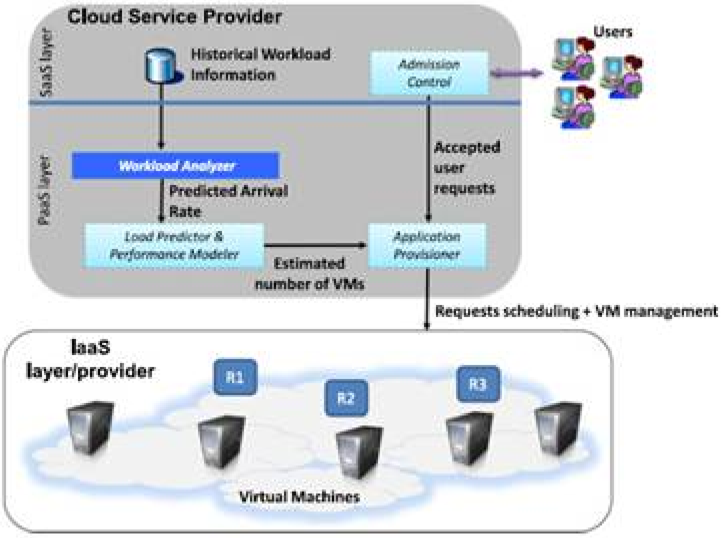

Fig.1 Architecture for adaptive cloud provisioning

1. Application provisioner: Receives accepted requests from the module Admission Control and forwards them to VMs that have enough capacity to process them. It also keeps track of the performance of the VMs. This information is then passed to the Performance Modeler and Load Predictor. The module Application Provisioner also receives from such module the expected number of Virtual Machines required by the application. If the expected number of Virtual Machines differs from the number of provisioned VMs, the number is adjusted accordingly.

2. Load Predictor and Performance Modeler: Decides the number of Virtual Machines to be allocated, will depend on the predicted demand by the Workload Analyzer module and on the observed performance of running VMs by the Application Provisioner. The performance is modeled via queueing networks, which, will depend on the predicted arrival rate of requests, return the minimum number of Virtual Machines that is able to fulfil the QoS metrics.

3. Workload Analyzer: Generates an estimation of future demand for the system or application. This information is then passed to the Load Predictor module and Performance Modeler module

V. PROPOSED SYSTEM

1. Proposed system shows ARIMA model which evaluates the future workload prediction accuracy by tracing the requests of users to the web servers.
2. It conjointly shows the impact of the achieved accuracy in terms of potency in resource utilization with marginal impact in latent period for users and smart QoS.

A. ADVANTAGES OF PROPOSED SYSTEM

1. Improve the cost
2. Provides good QoS
3. No loss of customers

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

VI. PROPOSED SYSTEM ARCHITECTURE

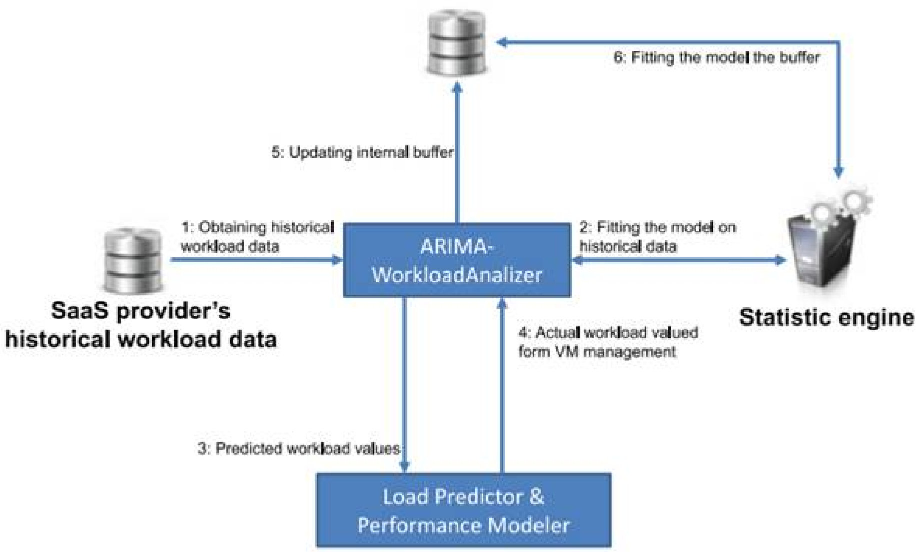

Fig.2 Workload prediction steps and its involving components

1. At the beginning of the execution, during a 1st most steps, the historical work information is fed into the work instrument, wherever the ARIMA model is work on them.
2. When the system starts operative, it delivers an estimation of the work with one time-interval earlier.
3. The length of the amount is adjusted to higher work the particular application.
4. The solely demand for economical system utilization is that the amount ought to be long enough to permit further Virtual Machines to be deployed.
5. Therefore, time windows as short as ten minutes can be appropriate counting on the chosen cloud supplier.
6. The request statistic contains the quantity of ascertained requests at whenever interval. it's enforced as a cyclic buffer in order that at successive prediction cycle, the particular variety of requests (obtained from the first dataset) is superimposed to the statistic employed in prediction whereas discarding the oldest worth. once constructing the request statistic, methodology} of fitting the ARIMA model is initiated supported the Box-Jenkins method
7. The statistic should be remodeled into a stationary statistic i.e. for every the time distinction (lag) between 2 information points, the mean and variance of the method should be constant and freelance of your time.
8. In this, historical information suggests that ascertained variety of requests that system receives in some hobby interval.
9. The autocorrelation plot is used to find how random a dataset is.

VII. CONCLUSION

Now-a-days largely firms prefers SaaS-based applications that square measure hosted on clouds, there square measure growing considerations concerning the QoS. Attributable to the raising fight within the SaaS market, application suppliers cannot afford to lose their customers to the competitors as results of meagre QoS. One amongst the key factors moving QoS is that the dynamicity within the employment, that ends up in variable resource demands. If at any given moment the employment exceeds resources' capability, QoS thereon explicit intervals are poor, moving customers' expertise with the applying.So the proposed system shows the prediction based on the ARIMA model. The workload analyzer in architecture enables the system to scale the resources without wasting them.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

REFERENCES

1. R. N. Calheiros, R. Ranjan, and R. Buyya, "Virtual machine provisioning based on analytical performance and QoS in cloud computing environments," in Proc. 40th Int. Conf. Parallel Process., Sep.2011, pp. 295–304.
2. N. Roy, A. Dubey, and A. Gokhale, "Efficient autoscaling in the cloud using predictive models for workload forecasting," in Proc. 4th Int. Conf. Cloud Comput., Jul. 2011, pp. 500–507.
3. V. G. Tran, V. Debusschere, and S. Bacha, "Hourly server workload forecasting up to 168 hours ahead using seasonal ARIMA model," in Proc. 13th Int. Conf. Ind. Technol., Mar. 2012, pp. 1127–1131.
4. V. Nae, A. Iosup, and R. Prodan, "Dynamic resource provisioning in massively multiplayer online games," IEEE Trans. Parallel Distrib. Syst., vol. 22, no. 3, pp. 380–395, Mar. 2011.
5. S. Pacheco-Sanchez, G. Casale, B. Scotney, S. McClean, G. Parr, and S. Dawson, "Markovian workload characterization for QoS prediction in the cloud," in Proc. 4th Int. Conf. Cloud Comput., Jul. 2011, pp. 147–154.
6. E. Caron, F. Desprez, and A. Muresan, "Forecasting for grid and cloud computing on-demand resources based on pattern matching," in Proc. 2nd IEEE Int. Conf. Cloud Comput. Technol. Sci., Dec. 2010, pp. 456–463.
7. Q. Zhu and G. Agrawal, "Resource provisioning with budget constraints for adaptive applications in cloud environments," in Proc. 19th ACM Int. Symp. High Perform. Distrib. Comput., Jun. 2010, pp. 304–307.
8. R. N. Calheiros, R. Ranjan, and R. Buyya, "Workload Prediction Using ARIMA Model and its Impact on Cloud Applications' QOS, Vol.3, NO.4, October-December. 2015