Volume 2, No. 12, December 2011
Journal of Global Research in Computer Science
REVIEW ARTICLE
Available Online at www.jgrcs.info
ISSN-2229-371X

STUDY AND PROVISING OF NETWORK TRAFFIC BASED ON MARKOV MODEL IN CLOUD ENVIRONMENT

Mousumi Paul*[1] and Goutam Sanyal[2]
*[1] Department of CSE, National Institute of Technology, Durgapur, West Bengal, India
mousumipual.88@gmail.com[1]
[2] Department of CSE, National Institute of Technology, Durgapur, West Bengal, India
nitgsanyal@gmail.com[2]

Abstract: The Cloud Computing is latest new computing paradigm where applications are provided across dynamic and geographically dispersed organization. The accurate modelling and simulation of flow of processes in cloud environment is an essential part of service level architecture. In this paper, an analytical model has been developed for identifying the pattern of inter-arrival processes and the service of processes. Empirical studies have shown that failure of processes after assignment to a suitable resource may affect the service pattern of the scheduling. Hence, the assumption that inter arrival, and service of processes are Poisson in nature does not always hold true, as it does not represent the real dynamics involved in service level cloud. The proposed paper has recognized the distribution pattern dynamically and scheduled the processes based on the distribution characteristics thus making the system more robust and dynamic. A fault tolerant mechanism has been built by rescheduling the failed processes.

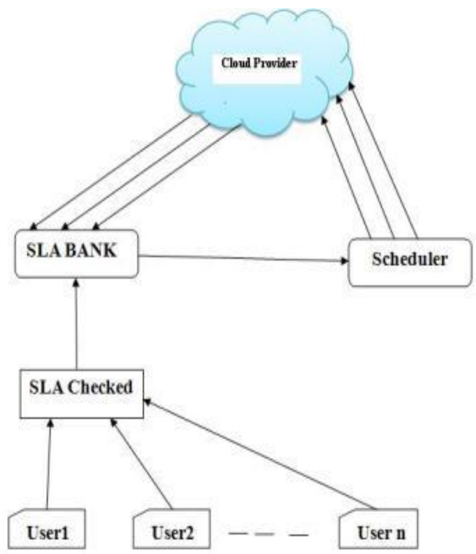
Keywords: cloud computing, network, Markov model

INTRODUCTION

Infrastructure-as-a-Service is one of the key element of cloud computing. Service providers offer their unused resources in the form of variable-priced virtual machines (VMs) for computing purposes. Therefore to lease spot instances at prices significantly lower than their standard fixed-priced resources. So that scheduling is the most important aspect of provisioning of these valuable resources. In the past few decades, the traffic modelling and dimensioning of networks developed by A.K. Erlang [1] has relied on two important user parameters namely the call arrival rate (λ) and average service time (1/μ) [2]. The Erlang theory has used the negative exponential distribution to model the call inter-arrival and service time [2] and was developed for traditional telephone networks. Since then, technology has grown many folds with additional features as well as types of services.
The proposed paper deals with the mobility prediction of the nodes to determine arrival pattern and service pattern of the processes coming into middleware to get resources for computing. Therefore absolute determination of the arrival and service pattern is necessary for efficient scheduling. This paper has used a different architecture that separates the resource discovery and scheduling by reducing load from middleware. The rest of the paper contains,

PROPOSED ARCHITECTURE

The two important aspect of cloud infrastructure-as-a-service are *resource providers* (we refer hereafter as CSPs—Cloud Service Providers) and *resource consumers* (we refer hereafter as SLAB—Cloud Resource Broker that acts as a consumer's software agent which deals on the basis of service level agreement). Whenever a user requires services, a variety of service access (those matched the users requirements) with different access cost is offered by the CSPs. If the users choose one of the options then the resource is provided to him by using the cloud scheduler.



PARAMETER ESTIMATION

Traffic Model

The distribution of processes at a cloud scheduler (say) can be assumed as the events i.e. arrival of processes occur randomly and independently with a constant probability of occurring in any small time interval Δt and that two events cannot occur exactly at the same time. Assuming the arrival of processes follows Poisson distribution, we may write

$$P_n(n) = e^{-\lambda} \lambda^n / n! \quad \text{------------------(1)}$$

Where λ is the mean arrival rate for a sample size of n. The relation between the mean length of process of Poisson's sampled data sets (as shown in Table 1) and the test data sets are studied. For our analysis, 3000 continuous processes are chosen as the test data set and few randomly selected samples are chosen from this data set. In the table, the entries are given according to their order they arrive. Assuming the traffic packets volume the mean packet length (X_n) of the sample set and its variance S^2 is calculated as :

$$X_n = \frac{1}{n}\sum_{i=1}^{n} X_i ----(2)$$

and

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - X_n)^2 \quad \text{------------(3)}$$

Table I Sampled data

Sample number	Λ	No of process(n)	Mean (X_n)	Variance (S^2)
1	1	2400	987.70	165.6
2	2	1534	970.50	316.8
3	4	895	949.94	439.72
5	16	286	1015.544	1564.6
6	32	151	940.37	2871.7

To check whether the samples represent the entire data set correctly or not we do a confidence test on them. Since the distribution is Normal, hence under Two Tail test:

- For 5% confidence, the limits for the mean of the data set is $\mu \pm 1.96\, \sigma / \sqrt{n}$
- For 1% confidence, the limits for the mean of the data set is $\mu \pm 2.58\, \sigma / \sqrt{n}$

Where σ the variance, n is is the sample size and μ is the mean of the entire data set.
The test values have been computed from the formula (4), shown below:

$$\frac{(X_n - \pi)}{\sigma / \sqrt{n}} \quad \text{------------------------(4)}$$

In this case μ = 950.7. The test result are listed below in table-2 for performing Null and Alternative Hypothesis test

Table II Test results

Sampling Set	Test Value	Results of Hypothesis Testing	
		95%	99%
1	1.86	Successful	Successful
2	-0.02	Successful	Successful
3	-1.23	Successful	Successful
4	-0.58	Successful	Successful
5	0.62	Successful	Successful
6	-0.08	Successful	Successful

It is observed that all the sample data set are within the significance level of 5% as well as 1%. Hence all the sample data set may be considered as true representative of all the bulk traffic collected for our analysis.

Scheduling using Markov model

The goal is to design a scheduler for each individual flows with parameters (s,λ,p) , where s is the packet size, λ is the mean rate of data arrival and p is the peak process arrival rate.
Suppose, there are k = 1,2, …….,K independent identically distributed (IID) resources. Processes from each flow, is stored in appropriate queue after classification. If overflow in queue length occurs, the process is dropped. Consider the queue buffer size for the k^{th} source is s_k, mean rate is λ_k , peak rate is p_k , and n_k, denotes the number of process of k^{th} class of traffic.
Two column vectors also defined as $s = [\, s_1, s_2, \ldots\ldots, s_k]^T$, buffer occupied at any time t by one set of process of each class, and $n = [\, n_1, n_2, \ \ldots n_k \ldots\ldots , n_k \,]^T$, total number of process at time t, where k denotes represents the type of process. Assuming a new process, if arrives, of class j should be entered in the queue if the buffer requirement of the process

$$S_j \le C - \sum_{k=1}^{k} n_k s_k --------(5)$$

So from equation (1) we get $s_j \le C - n^T.s$
The number of flows active and occupying the buffer is represented by the vector
$X(t) = [X_1(t), \ldots.. X_k(t), \ldots\ldots X_K(t)]^T , t \ge 0$ ----(6)

The processes are sent out through a number of queues. The time during which a packet stays in the system is IID exponential and the arrival process is from a stationary Poisson Process. One can define the process as
{X(t), t ≥0 }, a Continuous Time Markov Chain (CTMC), where the state $X(t) = [\, n_1, n_2, \ldots.. n_k, \ldots.. n_K \,]$.

Proposed Fault tolerance algorithm

The following algorithm describes the rescheduling of failed processes.

1. Start resource allocation control
2. Define queue size (B)
3. Allocate resource to a process and set time stamp
4. if a failed process arrives or time over
5. start timer
6. if a free resource available and matched the processes requirements
7. assign the process
8. stop and reset timer
9. else if Queue is not full
10. put the call in Queue
11. any resource node released, assign to a queued failed process on FCFS basis
12. if a queued failed process timed-out or actually leave the current node
13. drop the process from queue
14. end if
15. else

```
            drop the process
        end if
         end if
    end if
    if a new request arrives then
        if Queue is empty then
                  Allocate a free resource node  if available to
                  the new requestl
        else
              Drop the new request
        end if
    end if
    Stop
```

CONCLUSION

In this work, a new methodology has been established for performance evaluation in cloud environment by analysis and scheduling of user requests using Markov model. Since Markov is memory less process so that the queing of waited processes do not depend on storage capacity. A fault tolerance mechanism also make the system more robust and reliable which is ultimate goal of cloud computing Infrastructure-as-a –service (IAAS). Again designing a SLAB (Service Level Agreement Bank) separately which filtered out the matched resources thus reducing the load on middleware. The proposed methodology was effective.

REFERENCES

[1] Powell, C.J. and Leung, V.C.M. "Traffic engineering for integrated telephone and dispatch land mobile radio traffic", in Proc. Int. Conf. Selected Topic in Wireless Comm. pp. 168-171, 1992.

[2] Barcelo, F. and Sanchez, J.I. "Probability distribution of the inter-arrival time to cellular telephony channels", in Proc. 49th Veh. Tech. Conf., pp. 762-766, Houston, Texas, 1999.

[3] Mousumi Paul , Debabrata Samanta, and Goutam Sanyal," Dynamic job Scheduling in Cloud Computing based on horizontal load balancing", International Journal of Computer Technology and Applications (IJCTA) , Vol. 2 (5), pp.11552-1556, 2011,ISSN: 2229-6093.

[4] A. Adya, P. Bahl, and L. Qiu, "Analyzing Browse Patterns of Mobile Clients", Proceedings of ACM SIGCOMM Internet Measurement Workshop, San Francisco, CA, pp. 189-194, November 2001.

[5] A. Adya, P. Bahl, and L. Qiu, "Characterizing Alert and Browse Services for Mobile Clients", Proceedings of USENIX Technical Conference, pp. 343-356, 2002.

[6] D.Hang and S. Rapport, "Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and no prioritized handoff procedures", IEEE Trans. Veh. Technol, vol. VT-35, pp.77-92, Aug. 1986.

[7] D. Macmillan, "Traffic modeling and analysis for cellular mobile networks", Proceedings of 13th ITC, 1991, pp. 627-632.

[8] W. Jeon and D. Jeong, "Call Admission Control for CDMA Mobile Communications Systems Supporting Multimedia Services", IEEE Transactions on Wireless Communications, Vol. 1, No. 4, pp. 649-659, October 2002.