iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** IV **Month of publication:** April 2019

DOI: https://doi.org/10.22214/ijraset.2019.4162

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue IV, Apr 2019- Available at www.ijraset.com

N-Scalar

Anushka Kamerkar[1], Indraneel Baronia[2], Ashwini Bondge[3]
[1, 2, 3]*Pune Institute of Computer Technology*

Abstract: A system is said to be scalable if by adding more hardware, the application canlinearly take more requests than before. As a business grows, they have to scale their web application,common problems that arise when scaling a web application are :

1) Architectural Bottleneck
2) CPU bound applications
3) IO bound applications

Architectural Bottleneck refers to high latency components, that is a slow component in request pipeline puts a lower limit on the response time of the application.CPU bound and IO bound- An application is said to be CPU bound or IO bound if application throughput is limited by its CPU or IO operations respectively.

To solve the above problem, capacity planning,autoscaling and anomaly detection are required. There are tools present for the above mentioned problem, but they focus only on a single task. This project aims at building a tool that combines all the tasks required for scaling a web application.

I. INTRODUCTION

The motivation behind this project is using Newer generation technologies and techniques to enhance the performance of established operations and practices.

Using machine learning concepts to create a solution for the problem of auto-scaling of web application.

The project consists of four modules:

1) *Auto-Discovery:* This module will identify the applications running on various nodes, using neural networks and clustering.
2) *Capacity Planning:* Based on the applications running on the nodes,this module determines the stress or load on the system and creates a plan to reduce this load.
3) *Anomaly Detection:* This module will try to detect any differing behavior in terms of load.
4) *Auto Scaling:* This module will automatically scale cloud services like Virtual Machines and server capacities, up or down depending on the defined situations.

II. LITERATURE SURVEY

A. *On Demand Elastic Capacity Planning for Service Auto-Scaling*

Cloud computing allows on demand elastic service scaling. The capability of a service to predict resource requirements for the next operational period defines how well it will

exploit the elasticity of cloud computing in order to reduce operational costs. In this work, we consider a capacity planning process for service scale-out as an online pricing model. In particular, we study the impact of buffering service requests on revenues in various settings with allocation and maintenance costs. In addition, we analyze the incurred latency implied by buffering service requests. We believe that our insights will allow to significantly simplify predictions and mitigate the unknowns of future demands on resources.

B. *Efficient Load Balancing using Improved Central Load Balancing Technique*

Within a large-scale computing atmosphere the actual reasoning details centers and end users are usually geographically dispersed through the globe. The most significant for centers is how to handle and repair the actual countless needs that happens often by end users proficiently and correctly. Load leveling helps to have a great individual achievement and useful resource operation ratio by guaranteeing a simple yet effective and honest part of every computing resource. Appropriate Load Balancing aids in minimizing useful resource intake, employing fail-over, enabling, scalability, averting bottlenecks and over-provisioning etc. In this paper, “Central Load Balancer” a load leveling formula is overviewed to balance the burden among virtual products with reasoning data center. Final results exhibit that the formula can perform superior insert leveling in large-scale reasoning computing surroundings when in comparison to preceding balancing algorithms.

C. Auto Scaling Strategy for Amazon Web Services in Cloud Computing:

Auto scaling mechanisms have become a typical paradigm in cloud computing environments. Such mechanisms can increase or minimize the number of virtual machines according to user demands, consequently achieving pay-per-use objectives. However, auto scaling mechanisms provided by infrastructure-as-a-service providers must strictly follow user-defined thresholds; the drawback of such mechanisms is that they cannot respond to real-time Internet traffic loads by following user-defined thresholds. Therefore, we propose a dynamic threshold adjustment strategy that can expedite the creation of virtual machines according to workload demands. The proposed strategy can reduce the web application response time and error rate when the system is under a heavy workload.

III. PROBLEM FORMULATION AND PROPOSED WORK

A. Proposed System

This project aims at creatinga single tool which will combine all the processes of the various modules in single application. The modules are as follows

1) Auto-discovery
2) Capacity planning
3) Anomaly detection
4) Autoscaling

IV. THE DESIGN OF SYSTEM

A. Structure

Data Flow Diagram

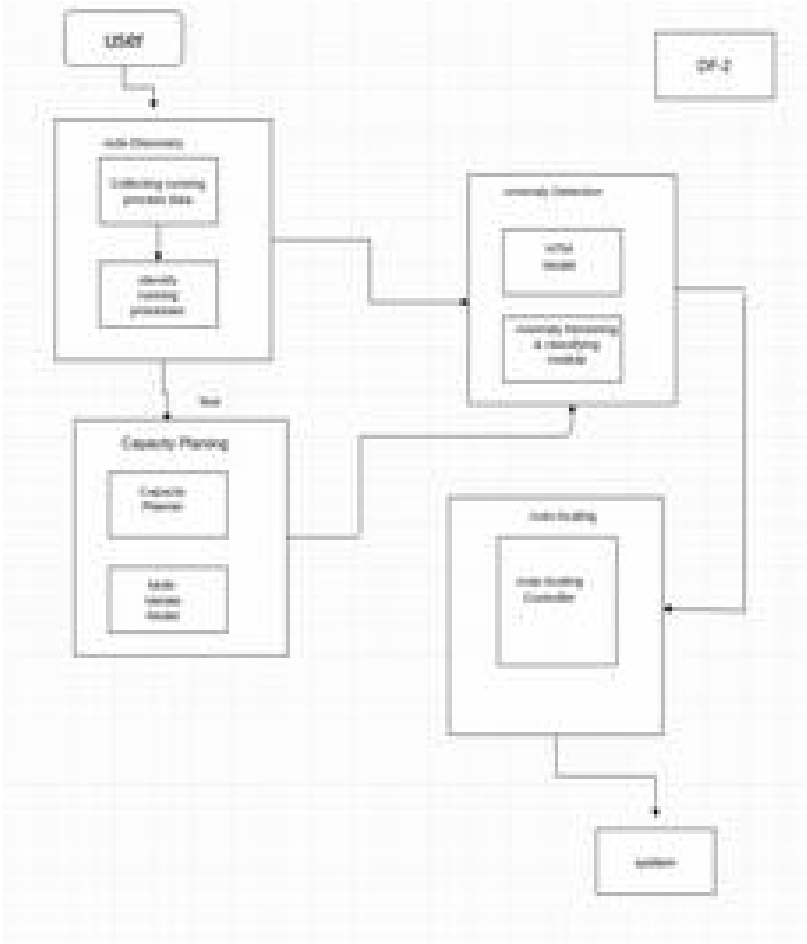


Fig. Data Flow Diagram

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue IV, Apr 2019- Available at www.ijraset.com

V. SYSTEM ARCHITECTURE ARCHITECTURE OF THE SYSTEM, WHICH IS SHOWN IN FIG.1

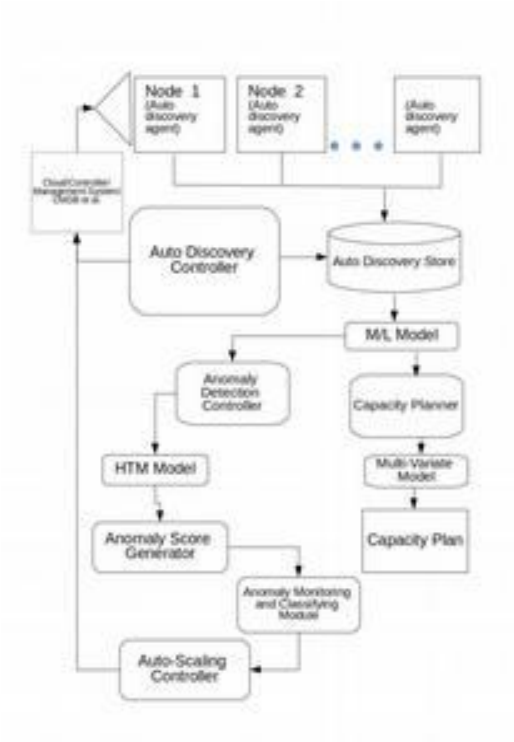


Fig 1. The structure of the system

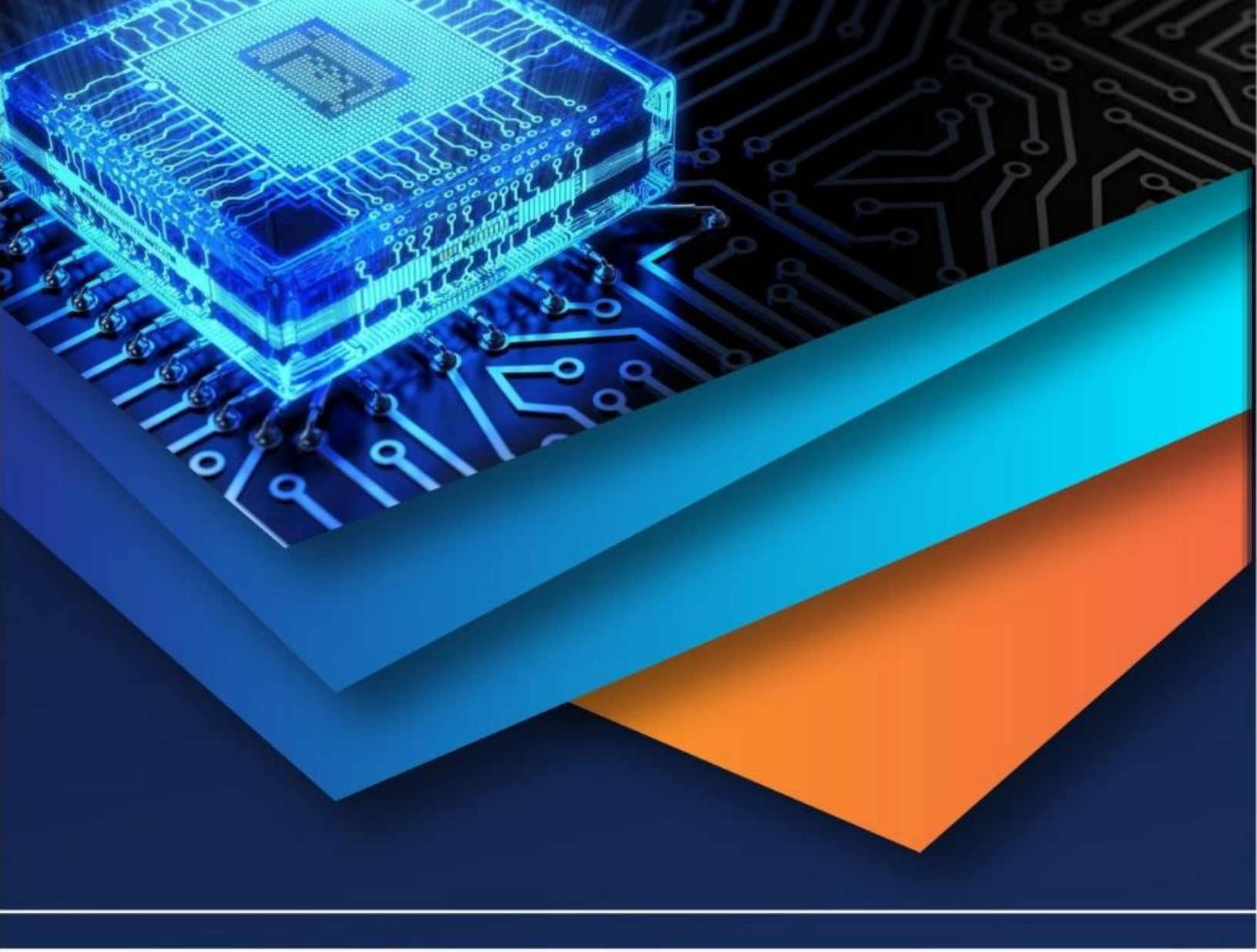
[4]The Duy Bui , Duy Khuong Nguyen, Tien Dat Ngo, “Supervising an Unsupervised

VI. CONCLUSION

This paper proposes the N-Scalar load balancing model which will try to reduce the load on web server by automation using auto-discovery, capacity planning ,anomaly detection and auto-scaling.

REFERENCES

[1] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, “Cloud Computing and Emerging It Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility,”Future Generation Computer Systems, Vol. 25, No. 6, pp. 599-616, 2009.
[2] Y. F. Li, W. Li, and C. F. Jiang, “A Survey of Virtual Machine System:Current Technology and FutureTrends,” IEEE international Symposium on Electronic Commerce and Security, 2010.
[3] Richard Dubes, AnilJain, “Clustering techniques:The user’s dilemma”,Department of Computer Science, Michigan State University, East Lansing, MI 48824, USA, October 24, 1976
[4] Neural Network” , First Asian Conference on Intelligent Information and Database Systems,Dong Hoi , Vietnam, 2009



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY