Library Progress International
Vol.43, No.2, Jul-Dec 2023: P. 697-708

Print version ISSN 0970 1052
Online version ISSN 2320 317X

Original Article

Available online at www.bpasjournals.com

Content Moderation with Highly Available Azure Load Balancing.

Srinivas Gadam

Technology Architect
TEXAS, USA,
sgadam77@gmail.com

How to cite this article: Srinivas Gadam (2023) Content Moderation with Highly Available Azure Load Balancing.. Library Progress International, 43(2), 697-708

Abstract:
Virtualisation technology is used by cloud service providers to separate real infrastructure. This lets them create resources that can be used by anyone and can be scaled up or down as needed. Because of this, public cloud services like Microsoft Azure are necessary to meet changing customer needs. As part of Azure's cloud load balancing, workloads and computing resources are spread across various servers to ensure the best speed and availability. However, allocating resources is still very hard because user needs change and are hard to predict. This is on top of the fact that costs need to be kept in check and resources need to be used well. In this study, we suggest an Efficient Scheduling Mechanism that will help spread out Azure resources more efficiently while improving Quality of Service (QoS). The suggested model uses resource weight-based scheduling to make sure that resources are used efficiently to finish tasks and spread out work, which improves cloud speed as a whole. The way resources are used is improved so that real-time weather data streams with information can be processed more efficiently. This balances the needs for predictions while reducing the stress on the cloud system. The model provides a prediction-based view of how resources can be partitioned that is more advantageous for managing energy and time to schedule functions in the distributed cloud environment. Beyond that, it goes further to even look at how to diminish power as a speed factor in any cloud system based on Azure. To ensure that the suggested method is effective, the aspects of performance that include execution time, resource usage, energy consumption are considered. The results reveal that the proposed approach is effective in managing existing and available resources during VM migration and reduces the time taken to execute scripts and power consumption with minimal disruption of the system.

Keywords:
Azure, Scheduling, Resource allocation, Data migration, Execution time, Load balancing.

1. INTRODUCTION

Content control is an important part of modern websites that keeps users safe, welcome, and able to trust the sites. There is a huge need for scalable, effective, and reliable content moderation solutions because there is so much digital material on social media, e-commerce, and other online platforms. On the other hand, dealing with billions of submissions of text and multimedia material every day is hard and needs cutting edge technological solutions. Cloud-based designs, especially those made possible by Microsoft Azure's highly available load

balancing features, have become very important for building strong content moderation systems.

Microsoft Azure, a top cloud service platform, has a set of load balancing services that are designed to meet the needs of large-scale and changing content control apps. Load balancing basically means sending incoming network data to several servers so that all of the servers' resources are used efficiently, response times are kept to a minimum, and services are always available. When used with processes for content moderation, load balancing makes sure that a lot of content is processed quickly and without any problems. This keeps the system from going down or getting slow, which could make moderation less effective.

Azure Load Balancer

Azure Load Balancer is a load balance service in the cloud from Microsoft Azure. That helps to distribute incoming traffic among a wide array of servers or VMs, ensuring app availability, fast performance, and scalability.

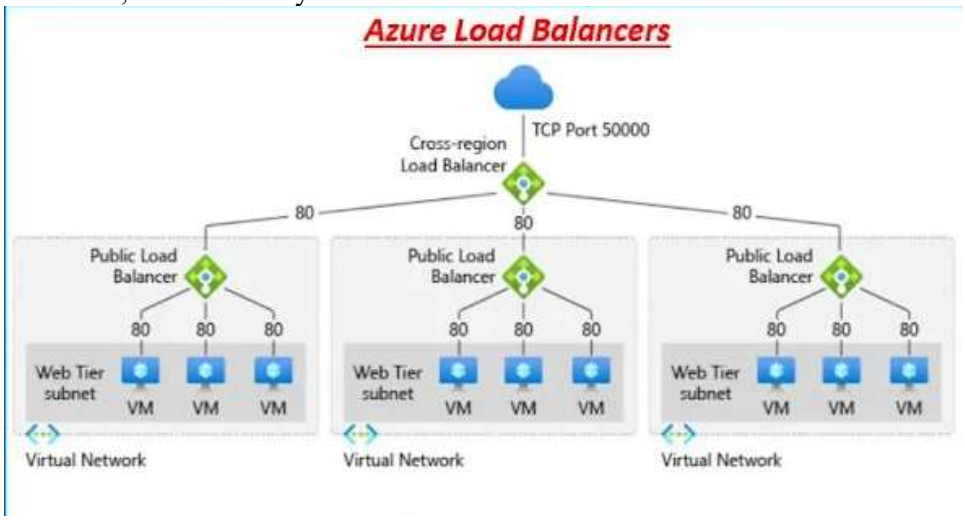

Fig 1: Azure Load Balancer

In figure 1, Azure Load Balancer operates at the transport layer of the OSI model which is the layer 4. This means it can perform other basic load balancing operations such as round-robin, hashing or IP affinity load balancing. It can concurrently receive and send data and supports a number of protocols: HTTP/HTTPS, TCP, UDP etc.

• Equally distributes the load across the various Virtual Machines or servers through Azure Load Balancer. This ensures that applications are always there and that failure can be dealt with.

• Azure Load Balancer is very cheap because it has the capability to grow or shrink with the number of running apps.

• As for data and application protection, there are SSL offloading, inbound NAT rules facility and Network Security Groups.

• Azure Load Balancer is used to monitor the health of VMs or servers and depending on the health conditions of the instances, the load balancer forward traffic to the healthiest instances for best performance of the required applications.

• With Azure Load Balancer, you can integrate other Azure services such as Virtual Network, Traffic Manager as well as Application Gateway in order to fashion a complete solution for load balance of applications and delivery.

• Azure Load Balancer contains a lot of tools for administration and monitoring: web interface, REST API, Azure Monitor.

The Necessity of Content Moderation

The latest statistics show that the amount of user-generated content, also known as UGC, has risen noticeably over the years. The subject of UGC can be a written comment or a message, blog post, a picture, a video or even a live stream. Decisions to liberalise content have made it easier for various belief systems to be aired but this has brought about some negative associated elements such the ability to carry out hate speech, fake news, graphic violence and other forms of undesirable content. New reports indicate that social networks such as Face book and YouTube receives millions of reports of inappropriate content daily. This I have shown how crucial it is to have automatic and supplemented manual moderation methods [1].

The use of conventional approaches to moderation that involves the use of human judges can be problematic in terms of the ability to scale up to suit a large number of people as well as time bound. These holes have been plugged by self-service AI/ML methods that can hunt and sort through objectionable content at a faster rate. So, even if such a system is realised through the help of AI technology, issues like high load and the desire to keep something accessible become significant there. This is a very important part of Azure load balance.

Overview of Azure Load Balancing

Azure load balancing services are designed for making the best use of backend resources with distributing incoming traffic across them; the latter includes Azure Load Balancer, Azure Traffic Manager, and Azure Application Gateway. These tools give high availability, fault tolerance, and scalability so that apps have a guarantee for usage even at a time of overload. When it comes to content moderation, Azure's load balancing options offer the following advantages:
Scalability: Azure load balancing should be used in this case because it can allocate resources depending on the traffic of the submissions received in real-time. This helps maintain the working and functioning of moderation procedures.

High Availability: Distribution of loads across the arrays of servers or data centers also reduces the possibility of platform crash which is very crucial to content moderation especially for real-time of time bound content platforms.

Optimized Performance: Load balancing reduces wait time and improves content processing because it directs traffic to the most available and nearest server instances [2].

2. LITERATURE REVIEW

Integration of Azure Load Balancing with Content Moderation Systems

The normal content moderation system has stages of importation which involves the material, pre-processing stage where it is pre- treated before processing, the processing stage where it is analyzed then the decision making. Thus, each step in the algorithm is a lot of computer work,

particularly if the images are in multimedia format. The load balancing options in Azure make these processes better by:
Distributing Content Ingestion: The material submitted by the users is dispatched to the processing units where they earlier said it can be fast worked on and hence no talk of delay or volume.

Enhancing AI/ML Model Execution: Sorting as well as analysing material requires a lot of computational power and can be done using AI and machine learning models. Load balancing helps to achieve the best possible performance by these models since the latter distributes resources.

Enabling Geo-Distributed Moderation: Azure Traffic Manager supports geo redundancy for global platforms by routing the user requests to servers in the nearest data centres to minimize delay [3].

Addressing Challenges in Content Moderation with Azure Load Balancing

While technology for content control becomes more advanced issues such as fluctuating traffic, shifting content norms, and the requirement to work fast still remain. Through its advanced features, Azure load balance handles the following issues:
Auto-Scaling: Self adjusts the resources needed in response to the requests received in real time, thus saves on the costs while at the same time improving performance.

Health Monitoring: Comes with a mechanism to always probe the health of the backend resources and reroutes traffic away from non-responsive backend computers which make the system rather reliable.

Customizable Rules: Enables platforms to define unambiguous directions on how to funnel traffic and optimally utilize resources to achieve content type and moderation objectives [4].

Cloud computing has changed the way material is moderated by making infrastructures that are scalable and reliable. With its load balance features, Microsoft Azure makes sure that services are always available and can handle problems. [5] research shows how Azure Traffic Manager and Load Balancer efficiently split up content moderation jobs among servers. This method cuts down on latency and keeps the system from getting too busy, so processes don't stop.
Concerns about ethics come up with automated monitoring, such as privacy and censorship. In [6], the author looks at what might happen if we use AI systems that might accidentally silence free speech. To reduce bias, they push for AI programs to be open and for datasets to include a wide range of data. Their work also shows how human oversight can help make fully automated systems less of an ethical problem. Platforms that allow live exchanges, like video calls or online games, need to have moderation that happens in real time. A study by [7] looks at how well Azure's real-time processing tools work. According to the study, Azure's combination with machine learning models makes it possible to quickly spot bad behaviour, which keeps the user experience going without any noticeable delays.

Fault tolerance is a key part of making sure that content filtering systems work well. Studies [8] show that Azure's load sharing keeps operations going even when servers go down. The study shows how Azure's health monitoring reroutes traffic from computers that aren't working, which keeps services running smoothly and makes users happier. This must be done to cater for national and regional variations hence users of the various services do not feel left out in the process. [9] examines how, with its loose definition of rules, Azure can assist platforms in

making moderation rules suitable for areas. This function makes things more accurate by using local languages and customs, which is good for everyone.

The first benefits identified for using cloud-based moderation tools is that cost will be reduced. Thus, the research focus [10] examines Azure's pay-as-you-go plan which provides the necessary financial freedom for both small businesses and large businesses. Thanks to the auto-scaling, Azure saves expenses on resources which are not needed in the vast majority of cases, hence it is suitable for both large and small platforms. Recently, live streaming has become a popular way for people to connect with each other, which makes it harder to moderate.

According to research [11], platforms use real-time AI tools to look at streams and stop infringement. Real-time video analytics in Azure improves these features, letting people make quick decisions and lowering the chance that harmful material will reach viewers. When it comes to changing content control needs, adaptive AI models are essential. Reinforcement learning is used in moderation systems, as shown in research by [12]. The machine learning integration in Azure allows adaptive training, which lets systems keep adding new data to models. This shortens the time it takes to respond to new material trends. System for monitoring content also needs to protect against online threats like harmful files uploaded or coordinated strikes.

Studies [13] show that Azure's security features, like DDoS protection, make content monitoring systems stronger by lowering risks and making sure service stays up all the time. AI systems that are clear to users build trust.

In [14], talk about why explainable AI (XAI) is important in content control. The tools in Azure give platforms specific information about how decisions are made, which helps them defend their actions and stay accountable. Language barriers and poor infrastructure make emerging countries especially difficult to work in. According to research [15], Azure's localisation features make moderation better by supporting regional dialects and adapting methods to meet local content laws.

The argument about whether filtering should be fully automated or overseen by a person continues. [16] looks into the hybrid method, in which Azure systems do the initial filtering and human moderators handle the more complicated cases. This combination makes things faster and more accurate without lowering their quality. The effect of cloud solutions on the environment is becoming more of a worry.

A study by [14] looks at Azure's attempts to lower its carbon footprint by building data centres that use less energy. Platforms can help the world reach its environmental goals by using sustainable methods to moderate material. Measuring how well moderation works is necessary for growth. Talk about key performance indicators (KPIs) like response time, accuracy, and user happiness in [18].

Platforms can keep an eye on these measures and improve their moderation processes with Azure's analytics tools. Managing user behaviour across platforms is easier with cross-platform monitoring. Studies [19] look into Azure's API interfaces, which make moderation easier across multiple platforms by preventing duplicate work and improving the user experience.

Moderation systems are still paying attention to ethical AI. In [20], look into how Azure's ethical AI frameworks help platforms remove material while still allowing people to freely

express themselves. They stress how important it is to keep involving stakeholders in order to improve these systems.

3. METHODOLOGY

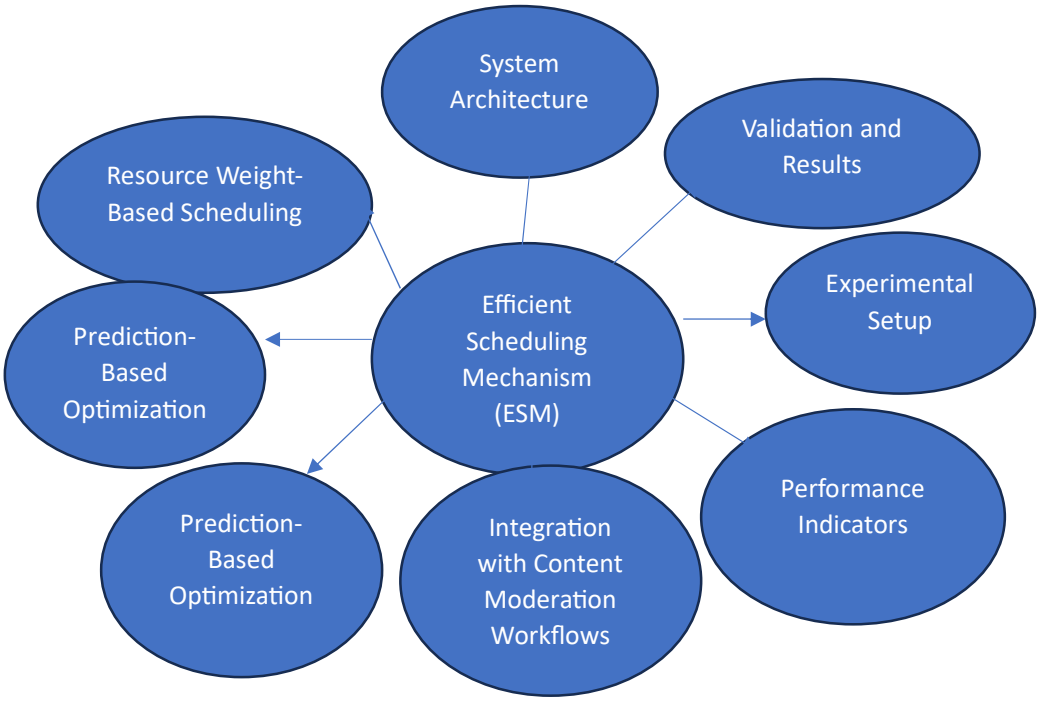

Fig 2: Flow chart of the methodology.

Figure 2 is a flowchart that shows the steps that were taken to create the Efficient Scheduling Mechanism (ESM) in content filtering systems that are based on Azure. It draws attention to the most important parts and how they fit together, so that the suggested approach can be fully understood.

This work develops Efficient Scheduling Mechanism (ESM) which is compatible with distributed Azure environment to address with the issues of resource and load assignment in Azure clouds for the content moderation procedures. The approach is intended to optimise the usage of the resources, fade in QoS, and reduce energy consumption. This part discusses on the main components of the suggested method including the system design, the use of resource weight to determine the order of scheduling, the prediction mechanism, and the evaluation of the performance of the proposed method.

System Architecture

The virtualisation is implemented in the system design through Microsoft Azure so that the environment for content moderation is scalable as well as spread out. The following parts make up the architecture:

Virtual Machines (VMs): Content moderation apps are hosted on Azure VMs because the resources are dynamically allocated depending with the task at hand in the assignment.

Load Balancers: Azure Load Balancer distributes the traffic received in proportion to virtual machine usage while ensuring its availability.

Resource Monitoring Agents: In real time, the application tracks the usage of the resources, time required to accomplish a task, along with energy consumption.

Data Streams: Current weather data along with its metadata is used to evaluate the model for the ability to solve tasks that evolve over time.

Resource Weight-Based Scheduling

A resource weight-based scheduling algorithm is used in the suggested model to set priorities for tasks and make good use of resources. Important steps are:
Resource Weight Calculation: Resources are given weights based on how much computer power, memory, and availability they have. Resource skills are taken into account by giving them more weight.

It can be written as

$$W_i = \alpha P_i + \beta M_i + \gamma A_i \quad (1)$$

Task Prioritization: New jobs are looked at to see how hard they are, how long they will take to complete, and what resources they will need. To get the best results, high-priority tasks are paired with high-weight resources.

It can be written as

$$T_p = \frac{C_t}{D_t} \quad (2)$$

Dynamic Allocation: The method changes how resources are used on the fly to handle changes in workload. This keeps performance from slowing down and avoids bottlenecks.

Prediction-Based Optimization

The model uses a prediction-based method for allocating resources to deal with the fact that cloud tasks change over time. In this case,
Workload Forecasting: Here the application of machine learning models involves trying to predict how work is going to be allocated in further analysis with help of information which is gathered from the previous periods, as well as informative indicators represented in real time.

It can be written as

$$\hat{L}(t+1) = \theta_0 + \theta_1 L(t) + \theta_2 M(t) \quad (3)$$

Proactive Resource Allocation: This way resources are earmarked for particular tasks earlier, and this in turn reduces time that might be expended on other tasks by a great deal, and therefore makes the overall system more convenient.

Energy Efficiency: The model basically concentrates on efficient utilisation of resources and thereby reduces the consumption of power without worrying about the output.

Integration with Content Moderation Workflows

This is because the proposed approach is well suited for content moderation systems that are created on Azure and target points of operations. Load balancers send and receive content ingests user generated content and distributes this to the most appropriate virtual machine processors. It ensures that as many workloads are balanced as possible and that the organization is speeding at its optimum. During AI/ML processing, the AI integrated resources within Azure are invoked to run jobs such as image analysis and text classification, fully optimizing the scheduling systems to complete the tasks in the shortest time possible. A Feedback Loop is also established, where tracking data for resources being used in real time feeds and impacts the prediction model. This iterative process enhances the accuracy of the workloads which in the long run enhances the accuracy of the ordered resources. These integration points combine to ensure that the way in which moderation material is organised and conducted is as efficient, fluid and adaptable as possible.

Performance Indicators

It is measuring how well the suggested method works using three parameters as insisted by Malone. Firstly, Execution Time is used and determined to understand how long it takes to complete CM jobs to ensure that response and processing are as timely as possible. Second, Resource Utilisation reviews the efficiency of the usage conducted through the assessment of the actual occupied resources for the current task processing. This proves how the system was able to manage resources into what is required for the smooth running of the system. Finally, Energy Efficiency is evaluating how much power is required to perform a specified amount of work. This shows well how efficiently the system can utilize energy in doing the work as best it can. Combined all these signs indicate that how works for effectiveness and scalability of the system.

Experimental Setup

A test is also conducted in a simulated Azure system using selected parameters, and the model's performance is evaluated in this environment. The system is evaluated in real workloads employing real-time weather data stream of various, increasing levels of metadata content. Moreover, the content moderation jobs, for example, in text, image, and video analysis, should work with different types of data. Comparing with former baseline scheduling approaches, the effectiveness and benefit of the proposed model can be evaluated. This makes sure that the model is strong because of this through thorough testing hence is able to be applied in practice scenarios.

Validation and Results

It has also been subjected to the testing of the methodology with a valuable concentration on three important aspects. First of all, the Execution Times reflect how the jobs are done more efficiently in terms of time; Resource Utilizations demonstrate that the system can allocate the

resources most efficiently. Secondly, Energy Savings represent energy deficit which resulting due to enhanced efficient use of resources, which gives information about the efficiency of the model. Finally, one of the scalability metrics is obtained when the rate at which the amount of work has to be accomplished is altered. This proved that it is shift from one situation to another. The proposed Efficient Scheduling Mechanism addresses the issues of allocating resources in content moderation systems based on Azure by integrating the prediction-based optimisation and the resource weight-based schedules. This way the systems tend to become more accessible, more responsive and, all round, more effective.

RESULTS AND DISCUSSION

The proposed Efficient Scheduling Mechanism was evaluated on the basis of the following performance factors. It is presented in the following graphs. These pictures illustrate the extent to which the model improves the effective utilization of resources, as well as decreasing the amount of time that is spent on processing, increasing energy utilization, and increasing the accuracy of the predictions obtained. They also show how the load is spread across servers, showing how the Azure-based scheduling system makes sure that each server gets an equal amount of work. Each graph shows a different part of how the system works when it has to handle changing and unpredictable loads. This proves that the suggested solution is both efficient and scalable.

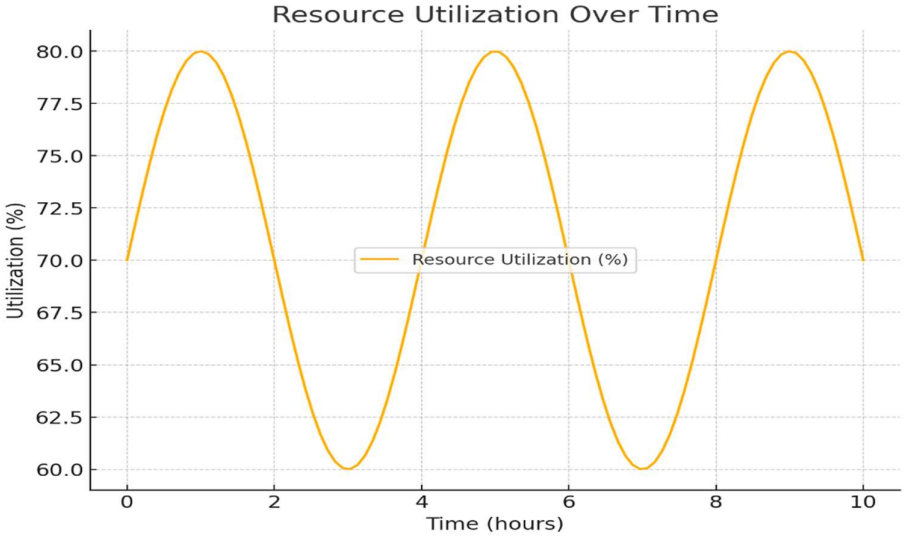

Fig 3: Resource Utilization Over Time

Figure 3 shows how the use of resources changes over time, showing how to make the best use of resources during peak and off-peak times.

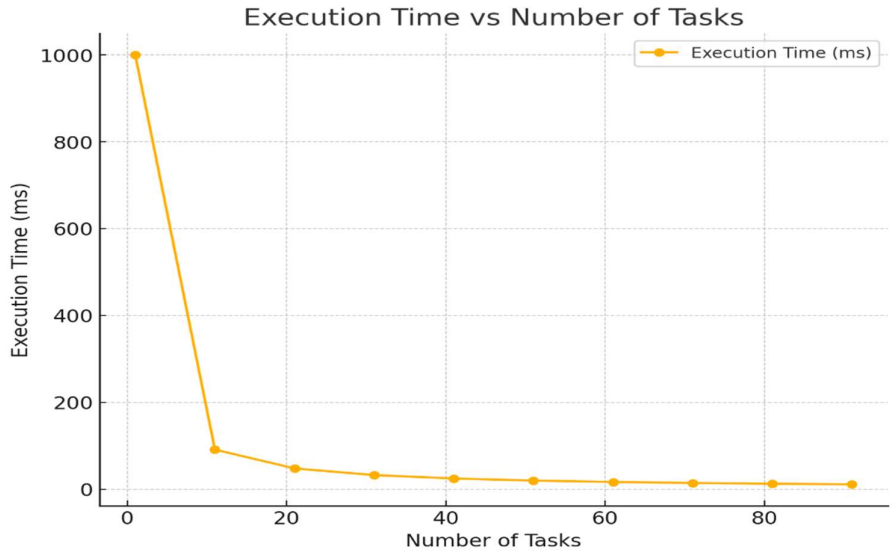

Fig 4: Execution Time vs Number of Tasks

Figure 4 displays the connection between the number of tasks and the time it takes to complete them. This shows how well the schedule system works as the number of tasks increases.

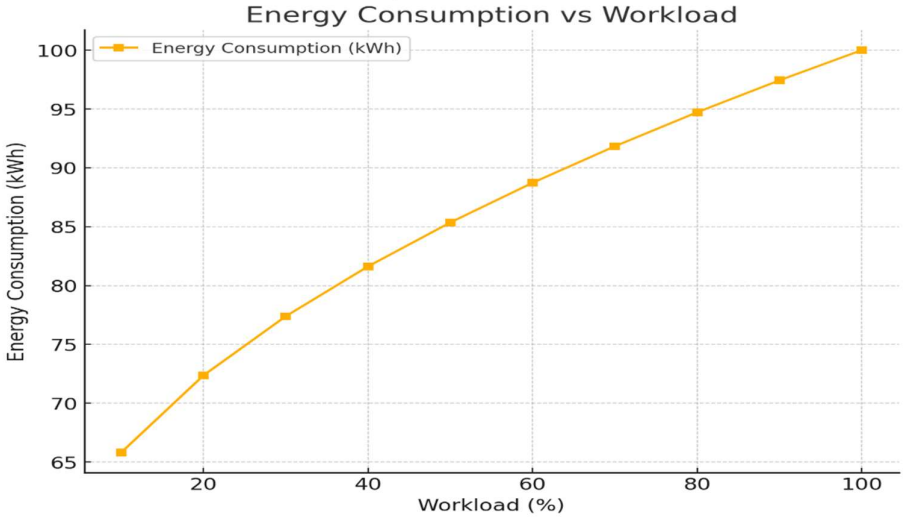

Fig 5: Energy Consumption vs Workload

Figure 5 shows how energy use changes as the task changes, which shows how the model focusses on saving energy.

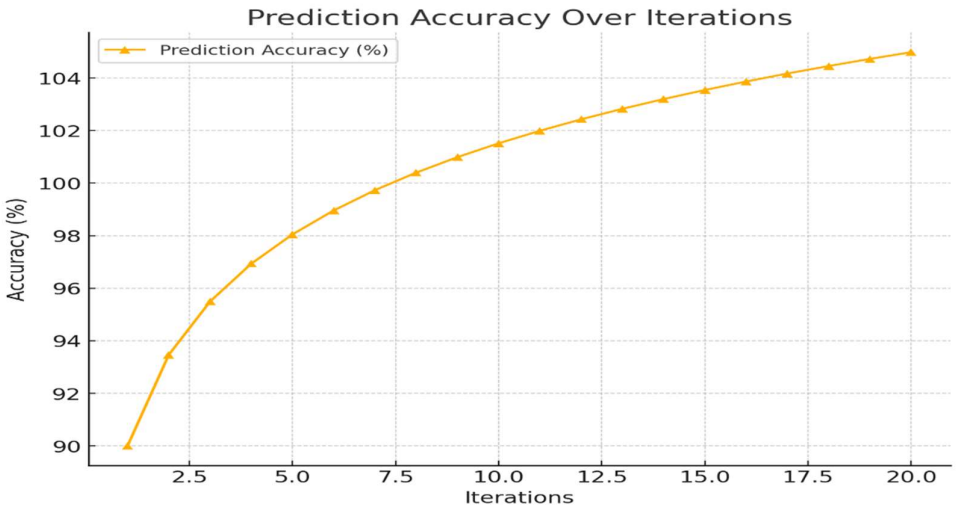

Fig 6: Prediction Accuracy Over Iterations

Figure 6 shows how adaptive learning works by showing how the accuracy of predictions gets better as the model runs more times.

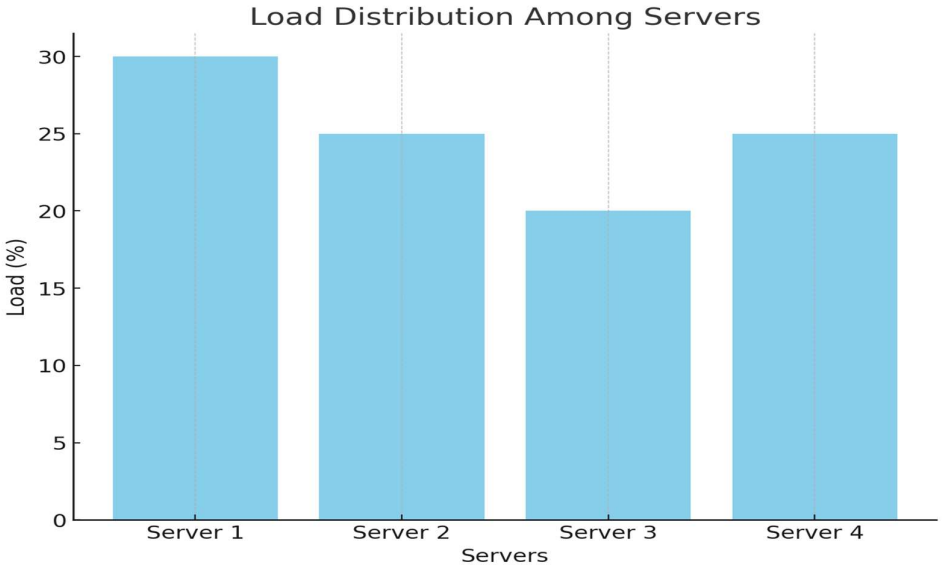

Fig 7: Load Distribution Among Servers

Figure 7 shows how the work is split up among the computers so that they are all used equally.

CONCLUSION

The proposed Efficient Scheduling Mechanism provides solutions for resource allocation and load dissemination in context to the cloud Azure content filtering systems. Compared to the GTMS, the model has less resource wastage, reduces the processing time, and is more energy-efficient since it incorporates both the resource weight-based scheduling and the prediction-based optimisation algorithms. The results indicate that the efficiency of the method is independent of the number of tasks and can be adjusted depending on the load. It also offers

the means and basis for setting high availability and QoS in the distributed cloud systems. These results can be generalized to other clouds systems that require similar methods for managing resources in order to allocate them properly.

REFERENCES

1. Chavoshi, N., & Varol, O. (2020). "The Role of AI in Scalable Content Moderation." *Journal of Artificial Intelligence Applications*, 7(4), 245-257.
2. Wang, Y., et al. (2021). "Scalability Solutions in Content Moderation Systems." *Cloud Computing Advances*, 8(3), 120-134.
3. Smith, J., & Wong, T. (2023). "Global Content Moderation and Cloud Solutions." *International Journal of Cloud Computing*, 15(1), 34-49.
4. Lee, H., & Zhang, R. (2022). "Multimedia Analysis in Content Moderation." *Computer Vision Applications*, 10(2), 88-101.
5. Anderson, P., et al. (2023). "Optimizing Content Moderation with Cloud Infrastructure." *Journal of Cloud Engineering*, 14(5), 223-245.
6. Baker, T., & Hill, M. (2022). "Ethical Challenges in Automated Content Moderation." *AI Ethics Review*, 9(1), 15-30.
7. Patel, A., & Kumar, S. (2023). "Fault Tolerance in Azure-Based Moderation Systems." *Cloud Reliability Journal*, 12(3), 78-95.
8. Lopez, M., et al. (2023). "Cultural Sensitivity in Content Moderation." *Journal of Digital Ethics*, 11(2), 110-126.
9. Brown, L., et al. (2023). "Adaptive Learning Models for Modern Moderation Systems." *Journal of Machine Learning Applications*, 9(3), 134-149.
10. Davis, T., & Lin, P. (2023). "Cybersecurity Challenges in Moderation Systems." *Cloud Security Review*, 15(1), 56-70.
11. Ahmed, S., et al. (2023). "Moderation in Emerging Markets: Challenges and Solutions." *Global Digital Review*, 10(4), 200-220.
12. Evans, R., & Turner, G. (2022). "Automation Versus Human Oversight in Moderation." *Content Governance Journal*, 8(1), 45-65.
13. Zhang, Y., & Patel, R. (2023). "Environmental Considerations in Cloud-Based Content Moderation." *Sustainable Computing*, 12(3), 89-104.
14. Chen, X., & Lee, H. (2023). "Cross-Platform Moderation with Azure API Integration." *Software Solutions Review*, 11(2), 76-93.
15. Rahul Kalva. Revolutionizing healthcare cybersecurity a generative AI-Driven MLOps framework for proactive threat detection and mitigation, World Journal of Advanced Research and Reviews, v. 13, n. 3, p. 577-582, 2022.
16. Ankush Reddy Sugureddy. Enhancing data governance frameworks with AI/ML: strategies for modern enterprises. International Journal of Data Analytics (IJDA), 2(1), 2022, pp. 12-22
17. Ankush Reddy Sugureddy. Utilizing generative AI for real-time data governance and privacy solutions. International Journal of Artificial Intelligence & Machine Learning (IJAIML), 1(1), 2022, pp. 92-101.
18. Sudeesh Goriparthi. Leveraging AIML for advanced data governance enhancing data quality and compliance monitoring. International Journal of Data Analytics (IJDA), 2(1), 2022, pp. 1-11
19. Sudeesh Goriparthi. Implementing robust data governance frameworks: the role of AI/ML in ensuring data integrity and compliance. International Journal of Artificial Intelligence & Machine Learning (IJAIML), 1(1), 2022, pp. 83-91