IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 5, May 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205061 |

Fault And Load Balancing Based on Cloud Secure Storage Mechanism

Kalyani Borkar[1], Prof. Pushpa Tandekar[2], Prof. Vijay Rakhade[3]

Dept. of Computer Science and Engineering, Shri Sai College of Engineering and Technology, Chandrapur, India[1]

Assistant Professor, Dept. of Computer Science and Engineering, Shri Sai College of Engineering and Technology, Chandrapur, India[2,3]

ABSTRACT: Cloud computing is the means of accessing a shared pool of configurable computing resources (including hardware, software, networks, servers, storage applications and services) that can be rapidly provided, used, and released with minimal effort on the part of users or service providers. But it has some of the main concerns like load management and fault tolerance. In this paper we are discussing load balancing approach in cloud computing. Load balancing is helped to distribute the workload across multiple nodes to ensure that no single node is overloaded. It helps in proper utilization of resources. It also improves the performance of the system This paper focuses on the load balancing algorithm which distributes the incoming jobs among VMs optimally in cloud data centers. In this paper, we have reviewed several existing load balancing mechanisms and we have tried to address the problems associated with them.

KEYWORDS: Cloud Computing, Load Balance, Security, Storage.

I. INTRODUCTION

In the world of cloud computing, Load Balancing is like a conductor, ensuring smooth and uninterrupted service. It works behind the scenes, so servers are not overloaded, keeping your apps and websites quick and reliable. Join us as we explore Load Balancing in Cloud Computing to understand how it maintains digital harmony by balancing the load. Load balancing in cloud computing is a crucial mechanism for optimizing the performance and reliability of digital services. Think of it as a traffic manager for data. It ensures that incoming network requests are efficiently distributed across a cluster of servers or resources. The primary goal is to prevent any single server from becoming overwhelmed, which could lead to slow response times or even system failures. Its main purpose is to monitor the health and capacity of each server and intelligently route traffic to the most suitable resource. This not only improves response times but also enhances fault tolerance. If a server becomes unavailable, the load balancer redirects traffic to operational servers, ensuring uninterrupted service.

For example, consider a popular online shopping website during a big sale. Many people visit the site at once. Load balancing ensures that all those visitors are evenly directed to different servers, preventing any one server from getting overloaded. This keeps the website running fast for everyone.

Goals Of Load Balancing

The goals of load balancing are:

- To improve the performance of the system.
- To have a backup of the load or entire server just in case the system fails or even partly fails.
- To maintain the system stability
- To accommodate future modification within the system

Types of Load Balancing

There are several types of load balancing methods used in cloud computing and web services, each designed to address specific needs and scenarios. Here are some common types:

- Round Robin Load Balancing Method
- Weighted Round Robin Load Balancing Method
- Random Load Balancing Algorithm
- Source IP Hash Load Balancing Algorithm
- URL Hash Load Balancing Algorithm
- Least Connection Method
- Weighted Least Connections Method
- Least Response Time Method

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205061 |

Working of Load Balancing in Cloud Computing

It works by efficiently distributing incoming network traffic or requests across multiple servers or resources. Here are the key functions of a Load balancer that explain how load balancing works in cloud computing:

- Traffic Distribution: When a client, such as a user's web browser, sends a request to access a website or application hosted in the cloud, the request first arrives at a load balancer. This load balancer is typically a specialized hardware appliance or software component configured to distribute traffic.
- Load Balancer Evaluation: The load balancer evaluates the incoming request and decides which server or resource in the server farm should handle it. The decision is made based on the specific load balancing algorithm or method chosen, which can be one of the types mentioned earlier, such as round robin, least connections, or content-based routing.
- Server Selection: Once the load balancer selects the server, it forwards the request to that server's IP address. The client remains unaware of this process, as it communicates only with the load balancer.
- Server Handling Request: The selected server receives the request and processes it. This can involve retrieving web pages, serving application data, or performing other necessary tasks.
- Response Sent to Client: After processing the request, the server sends the response back to the client, again through the load balancer. From the client's perspective, it appears as though the response came directly from the load balancer.
- Monitoring and Health Checks: Load balancers constantly monitor the health and performance of the servers in the server farm. They perform regular health checks to ensure that servers are online and responsive. If a server becomes unresponsive or experiences issues, the load balancer will automatically stop routing traffic to that server, ensuring that users aren't directed to malfunctioning resources.
- Scaling and Elasticity: In a cloud computing environment, load balancers play a crucial role in dynamic scaling and elasticity. As traffic fluctuates throughout the day or in response to events, the load balancer can automatically route traffic to additional servers as needed to maintain performance and availability. Conversely, if traffic decreases, it can scale down and reduce the number of active servers to save resources and costs.

Features of Load Balancing

- The client requests are distributed efficiently among several servers.
- It guarantees high reliability and scalability by transmitting requests only to those servers which are online.
- It offers flexibility to append or remove servers on demand.

II. LOAD BALANCING CLASSIFICATION

This is chiefly divided into 2 categories: static load balancing mechanism and dynamic load balancing mechanism:

1. Static Load Balancing: In the static load balancing algorithm the decision of shifting the load does not depend on the current state of the system. It requires knowledge about the applications and resources of the system. The performance of the virtual machines is determined at the time of job arrival. The master processor assigns the workload to other slave processors according to their performance. The assigned work is thus performed by the slave processors and the result is returned to the master processor. Static load balancing algorithms are not preemptive and therefore each machine has at least one task assigned for itself. Its aims in minimizing the execution time of the task and limit communication overhead and delays. This algorithm has a drawback that the task is assigned to the processors or machines only after it is created and that task cannot be shifted during its execution to any other machine for balancing the load. The four different types of Static load balancing techniques are Round Robin algorithm, Central Manager Algorithm, Threshold algorithm and randomized algorithm.

2. Dynamic Load Balancing: In this type of load balancing algorithms the current state of the system is used to make any decision for load balancing, thus the shifting of the load is depend on the current state of the system. It allows for processes to move from an over utilized machine to an underutilized machine dynamically for faster execution. This means that it allows for process preemption which is not supported in Static load balancing approach. An important advantage of this approach is that its decision for balancing the load is based on the current state of the system which helps in improving the overall performance of the system by migrating the load dynamically.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205061 |

a) Centralized approach: - In centralized approach, solely one node is liable for managing and distribution among the complete cloud system model. Alternative all nodes aren't liable for handling the requests and providing the response.

b) Distributed approach: - In distributed approach, every node severally builds its own load vector. The work is divided among all the nodes of the server. They aggregate the load information of alternative nodes. Distributed approach is additional appropriate for complicated and very large systems inside the cloud computing

.

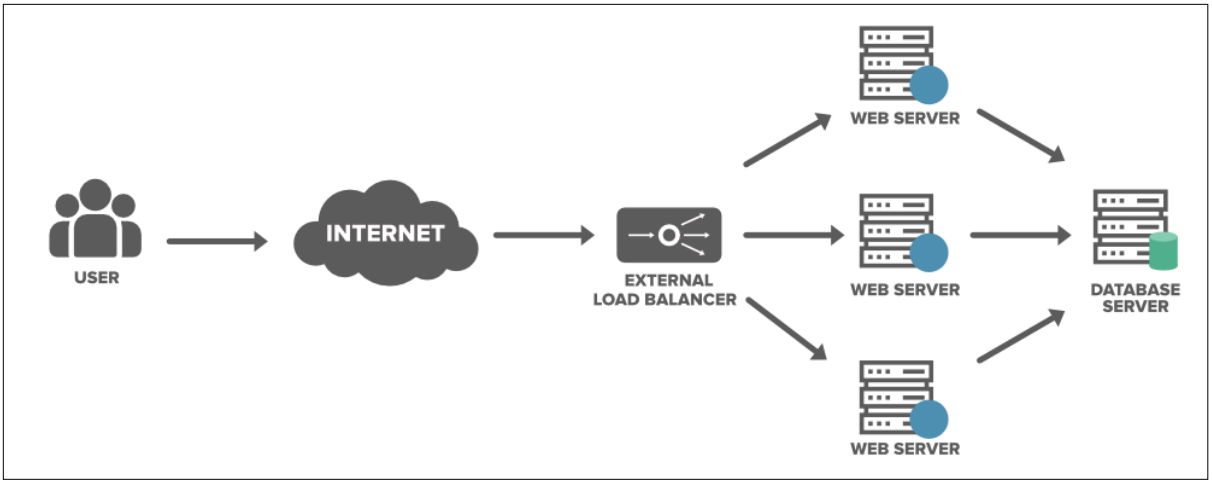


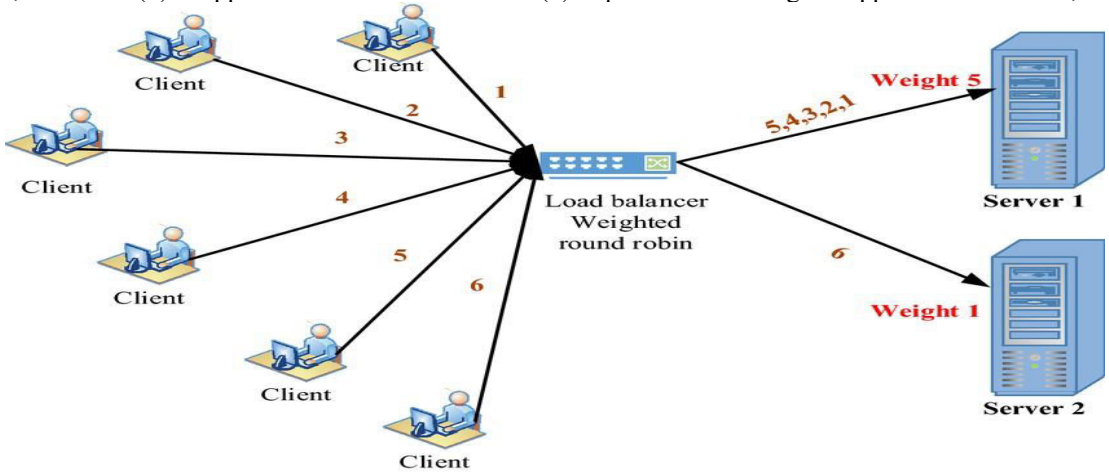
Fig 1: System Architecture

III. LOAD BALANCING TECHNIQUES

1) Weighted Round Robin load balancing method

Weighted round robin is similar to the round-robin load balancing algorithm, adding the ability to spread the incoming client requests across the server farm according to the relative capacity of each server. It is most appropriate for spreading incoming client requests across a set of servers that have varying capabilities or available resources. The administrator assigns a weight to each application server based on criteria of their choosing that indicates the relative traffic-handling capability of each server in the farm.

So, for example: if application server #1 is twice as powerful as application server #2 (and application server #3), application server #1 is provisioned with a higher weight and application server #2 and #3 get the same, lower, weight. If there are five (5) sequential client requests, the first two (2) go to application server #1, the third (3) goes to application server #2, the fourth (4) to application server #3. The fifth (5) request would then go to application server #1, and so on.



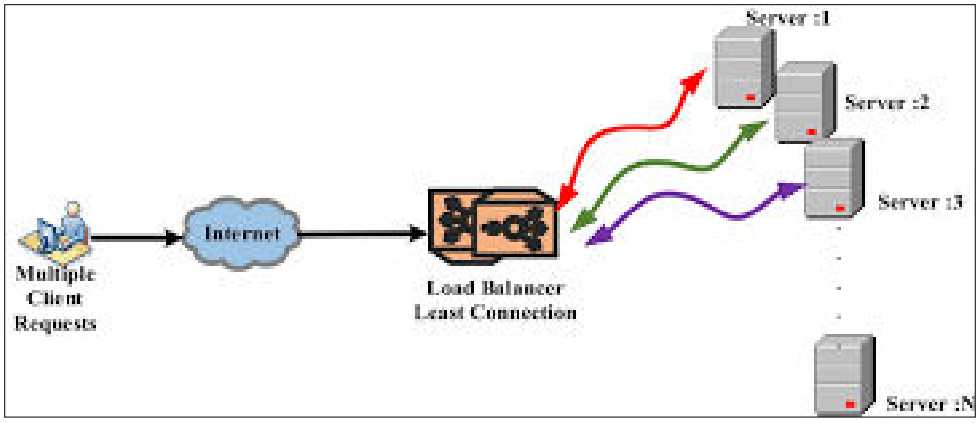
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205061 |

2) Least Connection load balancing method

Least connection load balancing is a dynamic load balancing algorithm where client requests are distributed to the application server with the least number of active connections at the time the client request is received. In cases where application servers have similar specifications, one server may be overloaded due to longer lived connections; this algorithm takes the active connection load into consideration. This technique is most appropriate for incoming requests that have varying connection times and a set of servers that are relatively similar in terms of processing power and available resources.



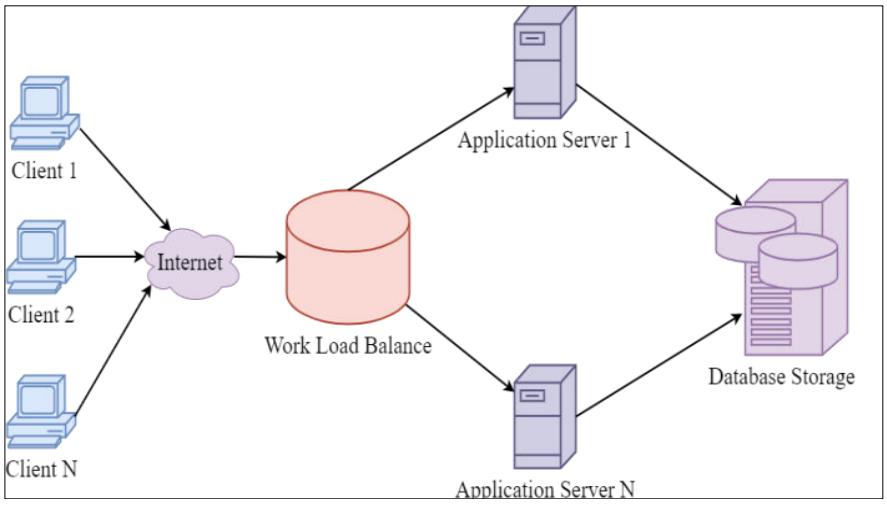
3) Weighted Least Connection load balancing method

Weighted least connection builds on the least connection load balancing algorithm to account for differing application server characteristics. The administrator assigns a weight to each application server based on the relative processing power and available resources of each server in the farm. The Load Master makes load balancing decisions based on active connections and the assigned server weights (e.g., if there are two servers with the lowest number of connections, the server with the highest weight is chosen).

4) Resource Based (Adaptive) load balancing method

Resource based (or adaptive) load balancing makes decisions based on status indicators retrieved by Load Master from the back-end servers. The status indicator is determined by a custom program (an “agent”) running on each server. Load Master queries each server regularly for this status information and then sets the dynamic weight of the real server appropriately.

In this fashion, the load balancing method is essentially performing a detailed “health check” on the real server. This method is appropriate in any situation where detailed health check information from each server is required to make load balancing decisions. For example: this method would be useful for any application where the workload is varied and detailed application performance and status is required to assess server health. This method can also be used to provide application-aware health checking for Layer 4 (UDP) services via the load balancing method.



| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205061 |

IV. KEY CHARACTERISTICS

1. On-demand self-service: A consumer can unilaterally provision computing capabilities, such as server time and net work storage, as needed automatically without requiring human interaction with each service's provider.

2. Broad network access: Cloud computing provide the users with various capabilities over the network which are acc essed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phon es, laptops etc.)

3. Resource pooling: The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer deman d. Examples of resources include storage, processing, memory, network bandwidth, and virtual machines

4. Rapid elasticity: Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scal e out, and rapidly released to quickly scale in. To the consumer, the capabilities available for provisioning often appear t o be unlimited and can be purchased in any quantity at any time**.**

5. Measured Service: Cloud systems automatically control and optimize resource use by leveraging a metering capabil ity1 at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

V. METHODOLOGY

1. Login / Logout: In this module the user can login by using their unique username and graphical password. The login module verify the user given username and password with the stored username and password in the cloud. If the userna me and password is matched the user can access the resources. If it does not match the user does not allowed to access t he resource.

2. Upload / Download: In this module the buyer and seller can post the ads and also able to download the data's that ar e post by other sellers. This module is mainly used to upload and also download the big data files.

3. Data Dispersion: Two data sets can have the same mean but they can be entirely different. Thus to describe data, on e needs to know the extent of variability. This is given by the measures of dispersion. Range, interquartile range, and sta ndard deviation are the three commonly used measures of dispersion.

4. Encryption: There are two basic encryption algorithms for cloud-based data: Symmetric encryption: The encryption and decryption keys are the same. This method is most commonly used for bulk data encryption.

Benefits of the Cloud mechanism

Cloud computing offers diverse shared pool of resources to provide better services to the user. It helps to reduce the ma kes span of the running applications, lower deployment cost and decreasing the human efforts to manage the cloud reso urces. Some of the important features of the cloud computing are:

1) Increased Throughput: Vast amount of the VMs deployed in large number of data centers reduces the computational t ime required to complete a cloud task.

2) Minimize risk of Infrastructure: Self- deployment of VMs helps to reduce the cost of installation of various infrastruc ture components such as severs installation, topology control management etc.

3) Low cost of entry: the rapid application development helps to reduce the time to market. This helps the decision mak es to make better decision to increase its reliability in the market place.

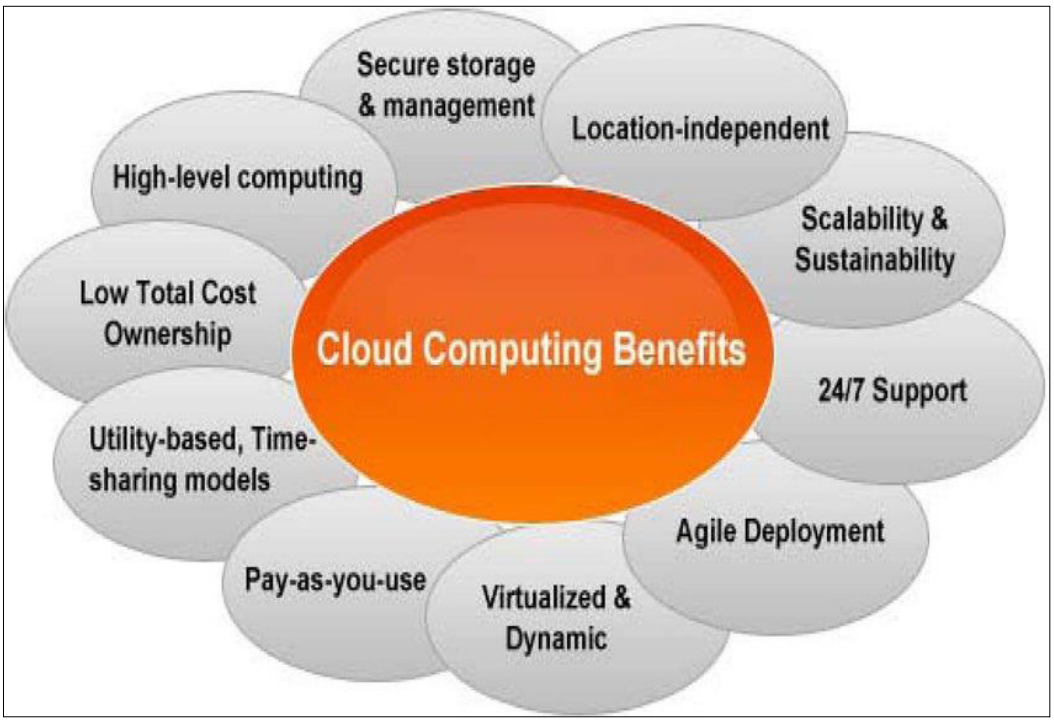
4) Scalability: the performance capabilities of the cloud enable the user of sundry areas and commercial areas to use its services.

5) Availability: fail- over mechanism helps to process the tasks even in the presence of the failures such as, check pointi ng mechanism, replication strategies, and so on.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205061 |



VI. CONCLUSION AND FUTURE WORK

Load Balancing in cloud computing is a vital mechanism for optimizing resource utilization, enhancing performance, and ensuring high availability. Across various applications and business sectors, load balancers play a crucial role in maintaining a balance and reliability in the dynamic landscape of cloud computing. Their significant benefits contribute to the smooth and efficient operation of applications, making them indispensable in modern IT infrastructure.

REFERENCES

1. R. Panwar and D. B. Mallick, "Load Balancing in Cloud Computing Using Dynamic Load Management Algorithm,"IEEE, pp. 773-778, 2015.
2. M. Belkhouraf, A. Kartit, H. Ouahmane, H. K. Idrissi, Z. Kartit and M. E. Marraki, "A secured load balancing architecture for cloud computing based on multiple clusters," IEEE, 2015.
3. S. S. MOHARANA, R. D. RAMESH and D. POWAR, "ANALYSIS OF LOAD BALANCERS IN CLOUD COMPUTING,"
4. International Journal of Computer Science and Engineering (IJCSE) , pp. 102-107, 2013. [4]. Nitish (2015) "Comparative study on load balancing techniques in cloud computing." Open Journal of Mobile Computing and Cloud Computing 1.1 (2014).
5. M. Randle's (2010) "An in-depth analysis and study of Load balancing techniques in the cloud computing environment." Procedia Computer Science 50 (2015).
6. Chunling (2018)" Load balancing and resource monitoring in cloud", Proceedings of the CUBE International Information Technology Conference. ACM, 2018.
7. Brian Chang (2023) "Leading Load Balancing" ICDCN '23: Proceedings of the 24th International Conference on Distributed Computing and Networking January 2023.
8. Steven Hofmeyr (2010) "Load Balancing on Speed" PPoPP '10: Proceedings of the 15th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming January 2010.
9. Bogdan Alexe, Thomas Deselaers, and Vittorio Ferrari, "What is an object?",. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2010.
10. Lowlesh Yadav and Asha Ambhaikar, "IOHT based Tele-Healthcare Support System for Feasibility and performance analysis," Journal of Electrical Systems, vol. 20, no. 3s, pp. 844–850, Apr. 2024, doi: 10.52783/jes.1382.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205061 |

11. L. Yadav and A. Ambhaikar, "Feasibility and Deployment Challenges of Data Analysis in Tele-Healthcare System," 2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI), Raipur, India, 2023, pp. 1-5, doi: 10.1109/ICAIIHI57871.2023.10489389.
12. L. Yadav and A. Ambhaikar, "Approach Towards Development of Portable Multi-Model Tele-Healthcare System," 2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI), Raipur, India, 2023, pp. 1-6, doi: 10.1109/ICAIIHI57871.2023.10489468.
13. Lowlesh Yadav and Asha Ambhaikar, Exploring Portable Multi-Modal Telehealth Solutions: A Development Approach. International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC), vol. 11, no. 10, pp. 873–879, Mar. 2024.11(10), 873–879, DOI: 10.13140/RG.2.2.15400.99846.
14. Lowlesh Yadav, Predictive Acknowledgement using TRE System to reduce cost and Bandwidth, March 2019. International Journal of Research in Electronics and Computer Engineering (IJRECE), VOL. 7 ISSUE 1 (JANUARY- MARCH 2019) ISSN: 2393-9028 (PRINT) | ISSN: 2348-2281 (ONLINE).

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details