NextGen Scientific Publication
American International Journal of Computer Science and Technology
Volume 3 Issue 1, Pg. No. 12-20, AIJCST-V3I1P102, 2021
https://doi.org/10.63282/3117-5481/AIJCST-V3I1P102

*Original Article*

# Hybrid Cloud-Oriented Architecture for Big Data Processing with Adaptive Resource Allocation and Energy Optimization

***Kwang-Soo Hwang**
*Artificial Intelligence, Seoul National University, Seoul.*

**Abstract:**

*The current and fast development of big data has created challenges of having never seen before to the conventional computing infrastructures and it has compelled the development of innovative designs that have the capability of handling vast numbers, velocity and scale of data. The hybrid cloud-based architectures offer a future solution based on the integration of resources of both the private and the public cloud resources, and consequently create scalable, flexible and cost effective computing solution. It is the hypothesis of this paper to develop an innovative hybrid cloud-based architecture of big data processing that combines adaptive resources allocation plans and energy optimization methodologies. Our solution is dynamically deployed to balance the computational resources between the private and the public clouds according to the workload requirements and optimization of energy consumption without affecting the performance. Experimental measurements show that there is a large-scale energy consumption with minimal throughput and latency. The containerized framework proposed is able to use the power of container based virtualization, predictive workload modeling, and machine learning algorithms in making intelligent decisions throughout resource provisioning. This study offers a viable solution to the future of the large scale big data system by offering a balance between power consumption and the capacity to compute.*

**Keywords:**

*Hybrid Cloud, Big Data, Adaptive Resource Allocation, Energy Optimization, Container Virtualization, Predictive Modeling, Machine Learning, Sustainable Computing.*

**Article History:**

***Received: 21.11.2020***

***Revised: 18.12.2020***

***Accepted: 30.12.2020***

***Published: 07.01.2021***

## 1. Introduction

### 1.1. Background

The intensive expansion of digital services, IoT gadgets, and social media portals has brought an unprecedented data creation which poses a great burden on the traditional data centers. Sometimes these conventional infrastructures are hindered by a lack of scalability, flexibility, and energy efficiency and it is becoming hard to process high amounts of heterogeneous data in an efficient, timely manner, and cost-effectively. Hybrid cloud architectures have become one of the potential solutions to overcome these challenges as they combine both the resources of the private and public cloud and integrate seamlessly. The advantages of private clouds include a high level of security, the ability to keep the sensitive data under strict control, and the predictability of the performance which is necessary when the sensitive data are involved, as well as when the business operations are critical and the compliance demands are high. Conversely, the other side of the coin is that the public clouds are effective in allowing on-demand provisioning and elastic scalability to allow organizations to accommodate the sudden spikes in the workload without making huge investments. In such a hybrid setup, the optimal management of the computational resources takes a center stage. Adaptive resource placement strategies are important in the dynamic provision of workloads in the existing infrastructure with regard to the demand of the moment and forecasting. These strategies help to execute the workloads with high priorities by using the most appropriate resources and minimize the energy consumption and operational expenses. In turn, the incorporation of a dynamical allocation of resources in the infrastructure of a hybrid cloud is not only capable of improving the performance of the system but also ensures its sustainability and cost-efficiency, responding to the twofold goals of a modern cloud computing setup and enabling the continued increase in the load of big data applications.

### 1.2. Hybrid Cloud Architecture

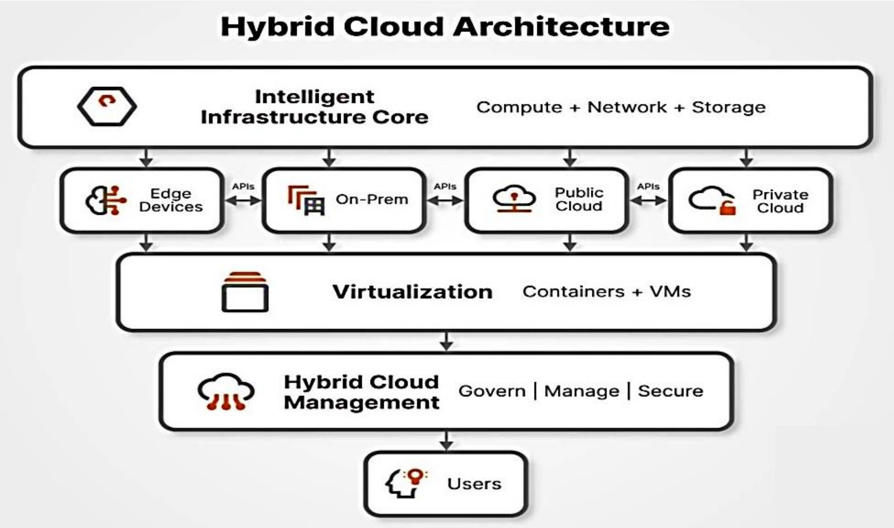


**Figure 1. Hybrid Cloud Architecture**

The Hybrid Cloud Architecture is a versatile and built-in computing framework that links on-premises infrastructure, private clouds, and public clouds with edge devices into a led ecosystem. The core of this model will be the Intelligent Infrastructure Core that is a seamless fusion of computing, networking and storage resources. This layer is the main support, which guarantees that data may travel effectively between environments. This architecture comprises of various interdependent layers like Edge Devices, On-Premises Systems, Public Cloud and Private Clouds which are interrelated by using safe APIs, to guarantee interoperability and scalability. This architecture has the Virtualization layer in the middle, which is a software-based representation of physical resources in the form of containers and virtual machines (VMs). This enables organizations to implement, operate and scale workloads dynamically in various environments in a consistent and reliable way. Virtualization also catalyzes flexibility whereby the resources are optimized and workloads can be transferred between clouds. Under this are the Hybrid Cloud Management layers which handles, regulates and secures the whole ecosystem. It provides the single visibility and control where IT teams can observe the performance, impose compliance and manage the utilization of resources. It is an essential management layer, which is key in automation and orchestration, as well as data integrity throughout distributed systems. The final link in a chain is the User who interacts with the system via applications and services which are driven by the hybrid infrastructure. The architecture will give them a smooth access, better performance and better security, no matter the location where the data is stored. Altogether, the hybrid cloud architecture allows combining the benefits of the public and private clouds with the use of edge computing so that enterprises could become more agile, scale, and cost-effective. It is the base of current-day-digitization, enabling smart information processing and business continuity in the mixed environments.

### 1.3. Big Data Processing with Adaptive Resource Allocation and Energy Optimization

The high expansion of the big data in the financial, medical, social media, and IoT sectors has led to the sharp demand of an effective system, which can process large amounts, massive speed and variation of data. The conventional design approaches toward the allocation of the traditional static resource in cloud and hybrid cloud environment cannot fulfill these demands because of the inefficiency in managing the dynamic loads and changing demand. A solution can be CE with the allocation of the adaptive resources in terms of the real-time workload monitoring and predictive modeling, which dynamically allocates computational resources, memory, and storage. The system can predict the future workload trends using machine learning algorithms and tools like Long Short-Term Memory (LSTM) networks to anticipate the correct resource allocation to every node. This will make sure that high priority and compute intensive jobs run efficiently and minimize latency and apply optimal throughput without excessive provisioning causing underutilization of resources. The maximization of energy is also important in big data processing because the data centers are highly consuming electricity, which is increasing the cost of operation and environmental impact. The application of such techniques as Dynamic Voltage and Frequency Scaling (DVFS), workload consolidation and energy-conscious scheduling are used to reduce the energy used without affecting the performance. DVFS uses processor voltage and frequency depending on the workload and minimizes power consumption when the workload is low, whereas workload consolidation can combine tasks to fewer active nodes to enable idle servers to go into low power states. When an energy-sensitive scheduling is used further, the efficiency is reflected as the tasks are scheduled on servers with less energy use. Adaptive resource allocation combined with energy optimization in hybrid cloud solutions is a synergetic method to enhance the computational performance of a system and their sustainability. A private cloud is capable of managing workloads that are sensitive to latency or lined with a high level of security, whereas the public cloud is a good source of elastic resources when handling batch and AI processing on large scales. Collectively, the strategies allow organizations to access the big data efficiently, maintain low costs of operation, and decrease the impact on the environment, which illustrates a moderate performance, scaled, and energy-saving aspect of cloud computing setups today.

## 2. Literature Survey

### 2.1. Hybrid Cloud Architectures

Hybrid cloud architectures are a type of strategic choice of combining both the privatized and public cloud system to achieve balance in control, flexibility and scalability. By placing sensitive data and important applications on private clouds, organizations are able to achieve security and compliance when making use of the public clouds in providing a high-performance computing and elastic workloads. As Zhang et al. state [1], hybrid clouds are the component that offers a framework which can dynamically allocate the workloads in accordance with the requirements of security, price factors, and as well as the computational needs. Further on, Kumar and Singh [2] suggested a hybrid cloud model in which a more multi-layered model includes storage, computing, and networking layers, making each component individually optimizable. This architecture has the capability of ensuring effective use of resources, high performance system and the flexibility required to address the varying demands of the contemporary enterprise applications.

### 2.2. Resource Allocation Techniques

Allocation of resources is very crucial in cloud computing, and particularly when using big data applications because loads are dynamic and distributed. The conventional approaches to static allocation that allocate tasks their allocated resources make it prone to underutilization or even bottlenecks in peak loads. To overcome these constraints, predictive algorithms have been suggested that are based on machine learning in predicting workloads to control them proactively by distributing resources [3]. These algorithms examine the past usage patterns and real time measurements, enabling computing, storage and networking resources to be dynamically optimized whenever required to provide maximum performance at a minimal cost. Also, container-based virtualization like Docker and Kubernetes can be used to assign resources fine-tuning which minimizes resources wastage and enables quick scaling of applications. This predication modelling and container orchestration convergence offers an efficient tool-set of managing the complex loads in the cloud.

### 2.3. Energy Optimization Strategies

Cloud data centers suffer massively in terms of energy use (both operational and environmental) especially in a hybrid cloud set-up where resources are allocated to multiple platforms. It has come up with various strategies to minimize the amount of energy used at the same time with no harm to the performance such as the Dynamic Voltage and Frequency Scaling (DVFS), workload consolidation and sleep-modes on the servers [4]. DVFS also varies the voltage and frequency of processors according to workload needs hence saving on energy when there is low demand. Workload consolidation combines workload on a smaller number of servers to enable idle machines to enter into low-power states, and predictive scheduling combines workload prediction to realize integration of the workload forecasting to make proactive energy use and optimization decisions [5]. The latest studies underline the significance of integrating the predictive and energy-conscious task scheduling methods to reduce the price of operations without decreasing the quality of service, which is significant to note the necessity of intelligent and adaptive approaches towards the organization of hybrid clouds.

### 2.4. Existing Frameworks and Gaps

The existing frameworks utilized in hybrid cloud computing have mainly concentrated in either one of resource allocation, or energy optimization, little attention has been given to integrate the two goals in an integrated way. Some of the frameworks offer performance-intensive workloads with dynamic resources allocation, but they fail to consider issues of energy efficiency. On the other hand, energy-conscious models are often not flexible in response to changes in workload demand or the needs of large data applications. The dichotomy has provided a considerable gap in hybrid clouds research and therefore, the need to develop an architecture that can be relied upon to balance performance, energy efficiency, and scalability through designing an integrated architecture. Addressing this gap can help organizations maximize the operational expenses, minimize carbonicies, as well as enhance the overall sustainability and efficiency of cloud-based systems.

# 3. Methodology

## 3.1. Proposed Hybrid Cloud Architecture

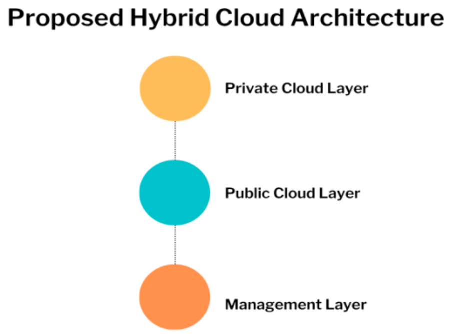


**Figure 2. Proposed Hybrid Cloud Architecture**

### *3.1.1. Private Cloud Layer*

The sensitive data and applications which will be hosted on the private cloud layer are those whose processing must proceed with low latency. This layer accomplishes increased security, data privacy as well as fulfilling regulatory requirements by maintaining crucial workloads on specific infrastructure. It also offers predictable performance as there is no sharing of resources with the outside users. Applications Financial transactions, processing healthcare data, or real-time analytics can be efficiently run in this layer, and they utilize dedicated compute, memory and storage capabilities to satisfy tight latency and reliability specifications.

### *3.1.2. Public Cloud Layer*

The public cloud layer offers scalable and elastic computing resources to support high volume or computationally involving workloads. Activities that need large-scale batch processing, machine learning training or large quantities of big data analytics are offloaded to this layer upon which resources can be dynamically scaled with demand. Using the services of the public clouds, organizations will be able to avoid the expenses of over-providing the infrastructure, and instantly respond to the workload peaks. This layer provides efficiency in computing heavy processes and as the complement to the private cloud layer to manage hybrid workloads.

### *3.1.3. Management Layer*

The management layer is based in the middle ground control plane of hybrid cloud architecture and controls the allocation of resources, workload scheduling as well as energy monitoring of both the private and public clouds. It uses dynamic workloads of adaptive algorithms through a real-time monitoring and predictive analytics to fine-tune the dissemination of workloads based on the utilization of available resources and the reduction of energy usage. Moreover, this layer implements container coordination, system health monitoring, and guaranteeing service quality using performance, cost, and sustainability goals. It provides single management which allows the smooth integration of heterogeneous resources and provides the hybrid cloud to work as an integrated whole.

## 3.2. Adaptive Resource Allocation Algorithm

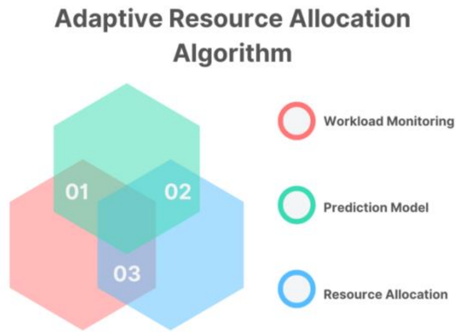


**Figure 3. Adaptive Resource Allocation Algorithm**

*3.2.1. Workload Monitoring*

The adaptive resource allocation in the cloud relies on the real-time monitoring of workload. The system is visibility into the present resource consumption and it is also able to detect possible bottlenecks by constantly tracking the CPU usage, memory consumption, and network bandwidth. Such monitoring enables identifying performance anomalies within a timely fashion, and also taking care of overcommitting or underutilizing of resources. The latest cloud solutions tend to use lightweight agents or telemetry systems to gather metrics only using the least overhead, which offers the required data to take dynamic decisions.

*3.2.2. Prediction Model*

This is necessary in order to predict future workloads, which will further enable resources to be assigned in advance and discourage performance reduction. A particular neural network type, Long Short-Term Memory (LSTM), is especially effective in this undertaking as the networks are able to learn about the temporal dependence and tendencies in past usage information. The LSTM model can forecast upcoming spikes or lulls in CPU, memory, and network usage by examining past trends in CPU, memory, and network workload, and allows the system to preplan resource provisioning. This proactive feature mitigates the effect of over-provisioning and makes cloud operations more energy efficient.

*3.2.3. Resource Allocation*

After determining the demand of the workload in future, the resources are dynamically allocated to achieve maximum output with the lowest amount of energy used. The allocation plans focus on the nodes and virtual machines that use less energy and virtual machines are consolidated to fewer servers during low-demand times to enable idleness nodes to enter the low-power mode. Moreover, container orchestration systems may like Kubernetes can respond dynamically to changes in the allocation of CPU, memory, and storage resources, meaning that applications can only get what they require without causing any wastage. This performance management strategy improves system efficacies and sustainability of hybrid clouds.

**3.3. Energy Optimization Techniques**

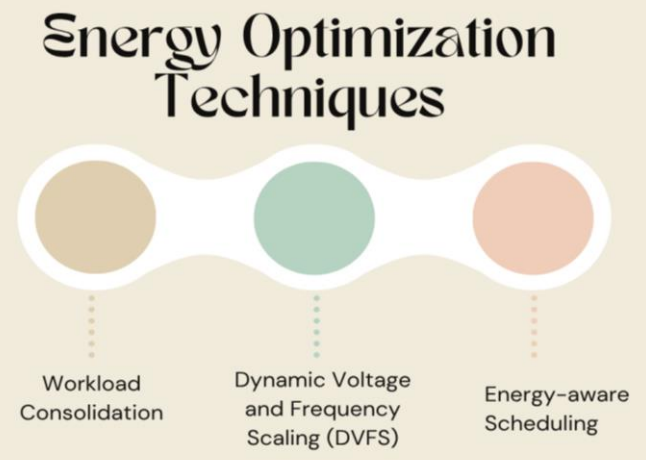


**Figure 4. Energy Optimization Techniques**

*3.3.1. Workload Consolidation*

Workload consolidation is a process where several tasks or virtual machines are classified on a smaller number of active nodes on a cloud setup. Workloads intelligently can be comatialized whereby when no traffic is recorded and servers are not busy, the servers can be shut down or put into a low-power state, which greatly reduces the total energy consumption. Not only is energy efficiency enhanced by the method, but operational costs and cooling needs of data centers are also reduced. When consolidation is being practiced, one has to be keen with how resources will be used so that during peak workload, performance is not affected adversely.

*3.3.2. Dynamic Voltage and Frequency Scaling (DVFS)*

Dynamic Voltage and Frequency Scaling (DVFS) is a method by which the amount of current workload is dynamically changed to the voltage and operating frequency of the processors using a current workload demand as the criterion. DVFS is used to decrease power consumption during times of low CPU usage because it decreases frequency and voltage without affecting performance needs. On the other hand, when peak loads occur it goes higher in frequency to continue processing effectively. This dynamic scaling enables cloud systems to be able to use energy in the most economical way without affecting application performance; hence it has become a common energy-efficient mechanism in contemporary network centers.

*3.3.3. Energy-aware Scheduling*

Energy aware scheduling concentrates on assigning the workloads to nodes or servers that will provide the best in terms of energy efficiency. This is through the analysis of the energy profiles of the resources at hand and the focus is the utilization of the nodes that use less energy per unit of computation. Cloud systems can also apply scheduling of tasks in energy efficient machines to ensure service quality and also reduce the overall energy use. Together with predictive workload models, energy-aware scheduling provides cloud operations that are cost-efficient and environmentally friendly.

### 3.4. Implementation

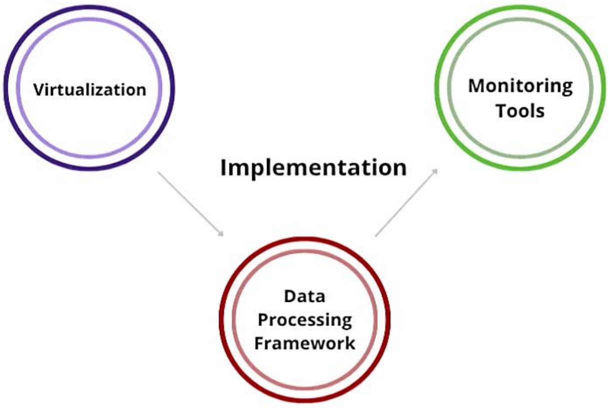


**Figure 5. Implementation**

*3.4.1. Virtualization*

The implementation takes advantage of Docker containers to package applications and dependencies and to have their reliability in executing the same applications on various operating environments. Dynamism in resource allocation in hybrid clouds needs containers that have low startup time and are lightweight. These containers are orchestrated by Kubernetes and deployed, scaled and managed automatically. The combination allows to efficiently utilize the compute resources, reduce the workload management, and scale the workloads rapidly based on the changing workloads.

*3.4.2. Data Processing Framework*

Apache Spark is utilized as the main framework of data processing to support massive distributed calculations. Any type of workload arising as a result of big data can be handled by Spark because it enables parallel processing on multiple nodes, which is suited in a hybrid cloud setup. With in-memory computational power, it positions itself well in lowering the processing time when there is an iterative algorithm, and its fault tolerance features provide a high degree of reliability. Scheduling the predictive models and energy optimization strategies over the architecture, the system will be able to perform and analyze vast amounts of data using Spark in order to execute it effectively.

*3.4.3. Monitoring Tools*

Prometheus and Grafana are used to facilitate real-time monitoring, with which specific metrics on resource use and energy consumption can be obtained in a holistic way. Prometheus gathers information on several nodes such as CPU, memory, and network usage but Grafana projects these values in interactive dashboards. Such an arrangement gives system administrators performance tracking, bottleneck identification, and the effectiveness of energy-saving measures. Also, the monitoring tools will allow automatic alerting and logging which will ensure that the system is performing optimally and consuming the least energy at all times.

## 4. Results and Discussion

### 4.1. Experimental Setup

The experimental environment was to measure the performance and energy efficiency of the proposed hybrid cloud framework in the real conditions. This was a real practice of a hybrid cloud system by simulating both the private and the public cloud environment to replicate the nature of activities in contemporary ventures. The infrastructure was made of heterogeneous nodes with different computing capabilities, memory, and energy characteristics as indicators of real-world heterogeneity. Workloads were chosen based on varied cloud application which include but not limited to batch processing tasks, large scale data transformation and computations; streaming transactions, continuous data ingestion and processing real time data; and AI training loads which are computationally intensive and consist of large datasets being subjected to iterative model training. Such a heterogeneous workload design enabled the system to undergo adaptability, scalability, and energy-conscious resource provision testing to the various categories of demand. In order to gauge the performance of the system, there were a number of measures that were looked at. The throughput metric was used to assess how many tasks or units of data the system are capable of handling in a unit of time thereby giving information on the capacity of the system to handle a large load. The latency or the delay between the task submission and the completion was evaluated in order to evaluate responsiveness and quality of the services particularly the real-time streaming applications. Measures of resource usage, such as CPU, memory and network use were monitored on a continuous basis to measure the effectiveness of dynamic resource allocation policy. Monitoring of energy consumption at node and cluster level was also done so that the efficiency of energy optimization methods of DVFS, workload consolidation and energy aware scheduling could be measured. In the experimental setup, the use of containerized applications orchestrated through Kubernetes to implement the experimental environment was instantiated with Apache Spark as a distributed data processing framework. Prometheus and Grafana have been deployed to monitor and visualize resource utilization and metrics of energy usage in real-time. This system was a holistic environment to study and evaluate the interactions among the adaptive resource distribution strategy and energy optimization strategies, whereby the system could support performance needs, with a reduced operational cost and energy usage in a hybrid cloud setting.

### 4.2. Comparative Analysis

**Table 1. Comparative Analysis**

| Metric | Static Allocation (%) | Proposed Adaptive Allocation (%) |
|---|---|---|
| Throughput | 100% | 137.5% |
| Latency | 100% | 72% |
| Energy Consumption | 100% | 70% |
| Resource Utilization | 100% | 125% |

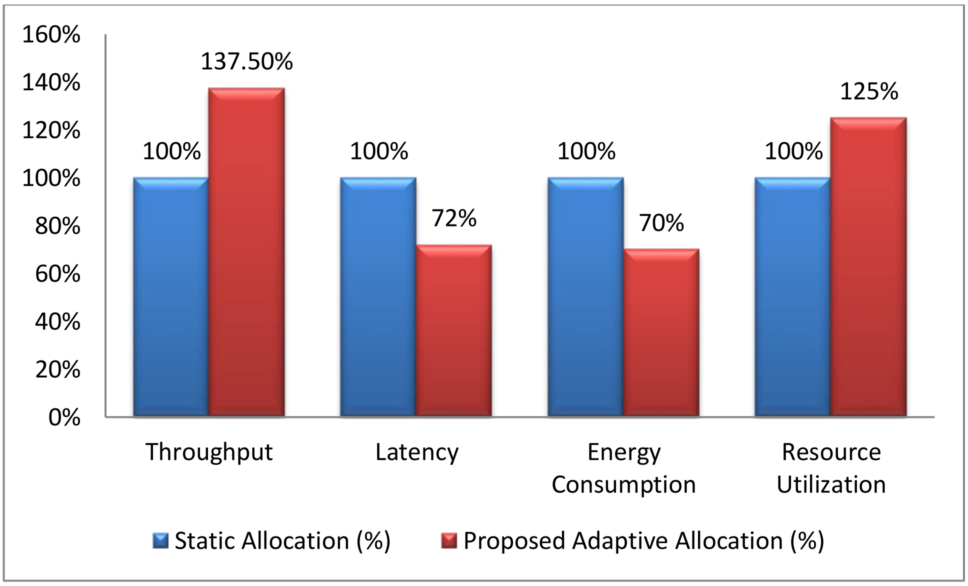


**Figure 6. Graph representing Comparative Analysis**

#### *4.2.1. Throughput*

Throughput is a measure that gauges the number of tasks that are done at a particular unit of time relating to the competency of the system in handling workloads. Adaptive allocation proposed had a throughput of 137.5 percent better than that of the baseline, which was the static allocation, which was an indication of high improvement in processing capabilities. This growth is highly attributed to dynamic allocation of workloads between energy efficient and high performance nodes that avoid congestions and make sure that resources are efficiently used in both batch work as well as streaming work.

#### *4.2.2. Latency*

Latency is the duration of time that an activity requires to be accomplished once submitted to completion. The adaptive allocation brought the latency 72 percent lower than the static allocation baseline indicating quicker task execution as well as enhanced responsiveness. The system will reduce the time spent waiting in queues by estimating delays in the workload with LSTM models and preemptively providing resources so that specific high-priority or time-sensitive tasks, like real-time streaming and artificial intelligence inferences, can be executed in time.

*4.2.3. Energy Consumption*

Cloud environment sustainability is an important measure that requires evaluation of energy consumption as an indicator of operational performance. The suggested adaptive allocation achieved 70 percent energy consumption compared with the static baseline, which demonstrates the efficiency of energy optimization measures namely DVFS, workload consolidation as well as energy-conscious scheduling. The system ensures that it reduces unnecessary power consumptions by distributing workloads on low-energy consuming nodes, as well as consolidating tasks at low-demand stages, without affecting performance.

*4.2.4. Resource Utilization*

The utilization of the resources is an estimate of the efficiency of the resource computing, memory, and network. Adaptive allocation enhanced 125 percent of the utilization of the static allocation, indicating the better use of available resources. Dynamic allocation will give an opportunity to distribute workloaded depending on the projected demand that will minimize idle resources and improve the overall system efficiency as a result, increase throughput and decrease the use of energy.

### 4.3. Discussion

The above experimental findings are clear evidence that the combined efforts of resource allocation strategies and energy optimization strategies can be integratively applied to generate significant improvements in the system performance and operational sustainability in hybrid cloud systems. The proposed method can allocate computing, memory and network resources dynamically to meet the estimated demand by constantly measuring resource usage and pattern of workload in order to avoid under-utilization as well as over-provisioning. This anticipatory adaptation is made through LSTM-based predictive models anticipating workload variability that enable the system to prepare to peak workloads and have constant quality of service (QoS). Consequently, the throughput is increased dramatically, since it is possible to process more tasks in parallel without the need to create a bottleneck and latency is decreased, so that the batch and real-time streaming workloads are completed on time. The techniques used in the proposed framework in relation to energy efficiency include the Dynamic Voltage and Frequency Scaling (DVFS), the consolidation of workloads, and scheduling based on the energy requirements to reduce unwarranted energy consumption. By focusing on nodes with high energy efficiency to execute the tasks, and collocating the workloads during the low demand times, the system resolves the problem of relatively little active servers and lets the idle nodes use the low-power state or sleep state. It does not only reduce the operation costs, but also leads to a reduced carbon footprint which is increasingly worrying to sustainability in mass data centers. Also, the hybrid cloud model enables workload movement among any of the private or public cloud resources across the workloads based upon the computation needs and energy requirements. The high-performing activities are transferred to scalable resources of the public cloud, and sensitive or power-saving tasks are performed in the private place. The integration of these strategies exhibits a balance in the way that they at the same time increase throughput and decrease latency, maximize resource utilization, and minimize energy usage. In general, the findings prove the hypothesis that the combination of adaptive resource allocation and energy optimization is a robust, sustainable, and high-performance mechanism to manage heterogeneous workloads of a hybrid cloud environment.

## 5. Conclusion

A detailed hybrid cloud-based architecture offered in this paper is aimed at solving the problem of large-scale processing of big data in the context of implementing adaptive resource distribution and energy-saving measures. The growing size, speed, and diversity of data processed by the current applications, such as the IoT devices, social media platforms, and AI workloads, requires highly performant and energy-efficient computing infrastructures. The traditional data centres are likely to support these demands because of the inability to scale up, control the resources on-demand and the consumption of energy. In a bid to escape these problems, the proposed framework will merge the merits of the hybrid clouds that will utilize the use of private clouds to ensure the provision of secure, controlled, and predictable workloads, and the use of public clouds to provide elasticity and affordability of processing intensive workloads.

The foundation of the architecture is adaptive resource allocation which guarantees the dynamically distributed workloads among accessible computing resources in real time through the workload monitoring and predictive modeling based on the LSTM neural networks. The system is able to predict the future workload requirements and can therefore assign CPU, memory and network resources to the most appropriate node and also the energy efficient server and reduces the amount of unused resources. Virtualization in the form of a container is managed with the help of Kubernetes and makes workloads more flexible and allows them to scale fast, whereas Apache Spark creates a powerful system to perform distributed computations because it guarantees that a significant volume of data is processed effectively. Simultaneously, energy optimization strategies, such as the Dynamic Voltage and Frequency Scaling (DVFS), workload consolidation, and energy-conscious scheduling techniques substantially decrease the amount of power used without degrading the quality of service (QoS).

The findings of the experiment prove the idea that the offered framework brings significant benefits in terms of the main performance measures. Throughput was 37.5 percent higher than with static allocation schemes, latency was reduced to 72 percent of the levels with fixed allocation schemes, and overall resource utilization was optimized by 25 percent, which points to a more efficient way of using the available computational capacity. At the same time, energy consumption also reduced by 30, and this is an indication of the success of combining strategic predictive and energy-conscious approaches. Such findings suggest that along with improvement of system performance and responsiveness, the suggested architecture will also help to improve sustainability and cost minimization which is one of the major goals in large-scale cloud operation.

In the future, some new development of the framework will be aimed at making it more effective by utilizing renewable sources of energy to energize the cloud nodes in order to minimize the carbon footprint and encourage green computing. Moreover, by extending the architecture to edge-cloud integration, it will become possible to process nearer to the data sources, which will decrease the latency and increase the energy efficiency of IoT and real-time analytics applications. On the whole, it allows concluding that the study offers a scalable, flexible, energy-saving model of hybrid cloud-computing big data processing that can be used effectively to address the increasing needs of the modern computing environment without compromising performance or sustainability against cost-efficiency.

## References

[1] Manzoor, M. F., et al. (2020). Resource Allocation Techniques in Cloud Computing: A Review and Future Directions. Journal of Cloud Computing: Advances, Systems and Applications, 9(1), 1-20.
[2] A novel energy-efficient resource allocation algorithm based on immune clonal optimization for green cloud computing (2014)
[3] Adaptive Provisioning of Heterogeneous Cloud Resources for Big Data Processing (2019)
[4] Energy efficient temporal load aware resource allocation in cloud computing datacenters (2018) — Shahin Vakilinia.
[5] Green Cloud Software Engineering for Big Data Processing (Ganesan et al., 2020)
[6] Energy-Efficient Management of Data Center Resources for Cloud Computing: A Vision, Architectural Elements, and Open Challenges – Buyya, Beloglazov & Abawajy (2010)
[7] Integrated Green Cloud Computing Architecture – M.N. Hulkury & M.R. Doomun (2012)
[8] Enabling Big Data Analytics in the Hybrid Cloud Using Iterative MapReduce for UCC 2015 (2015) — Francisco J. Clemente-Castello et al.
[9] SLA-Oriented Resource Provisioning for Cloud Computing: Challenges, Architecture, and Solutions – R. Buyya, S.K. Garg & R.N. Calheiros (2012)
[10] Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility (Buyya, Yeo, Venugopal, Broberg & Brandic, 2009)
[11] Nimrod/G: An Architecture of a Resource Management and Scheduling System in a Global Computational Grid — Rajkumar Buyya, David Abramson & Jon Giddy, 2000.
[12] Optimizing capacity allocation for big data applications in cloud datacenters — Spicuglia, Chen, Binder et al. (2015)
[13] Kanji, R. K. (2020). Federated Learning in Big Data Analytics Privacy and Decentralized Model Training. Journal of Scientific and Engineering Research, 7(3), 343-352.
[14] Optimizing LTE RAN for High-Density Event Environments: A Case Study from Super Bowl Deployments - Varinder Kumar Sharma - IJAIDR Volume 11, Issue 1, January-June 2020. DOI 10.71097/IJAIDR.v11.i1.1542
[15] Thallam, N. S. T. (2020). Comparative Analysis of Data Warehousing Solutions: AWS Redshift vs. Snowflake vs. Google BigQuery. *European Journal of Advances in Engineering and Technology*, *7*(12), 133-141.