Optimizing Multimedia Processing over Cloud Cluster using modern Container and Container orchestration Technology

Lukman Ibrahim Isa[1], Sohit Agarwal[2], Ravi Jangid[3]

[1]*Suresh Gyan Vihar University, Jagatpura 302025, Jaipur, Nigeria*
[2]*Suresh Gyan Vihar University, Jagatpura 302025, Jaipur, India*
[3]*Suresh Gyan Vihar University, Jagatpura 302025, Jaipur, India*

---***---

Abstract - *In this literature, we present a slightly different approach to multimedia processing while leveraging the modern container and container orchestration technology to cluster computing using AWS known cloud platform. This approach optimizes multimedia processing over cloud clusters using modern container and container orchestration technology. In this approach, we first outlined the multimedia components and developed the architecture for our multimedia processing technology using modern container and container orchestration technology over cloud clusters, then we integrated and deployed the framework over AWS platform. The result of our work provided a new technique for optimizing large scale multimedia processing solutions which demonstrated the power of modern container and container orchestration technology to cluster computing using known cloud platform AWS. We simulated our multimedia processing model and generated optimized results.*

***Key Words*: AWS cloud platform, Container, Container orchestration technology, Cluster computing, multimedia processing.**

1. INTRODUCTION

Multimedia is a compound word of multi- and media. The prefix multi- originated from Latin word multus meaning "numerous". The root media is the plural form of the Latin word medium; medium is a noun and means "middle, centre". The meaning of Media varies with the context in which it is used [1]. As a matter of fact, multimedia had been in existence since the era of slide projectors and tape recorders, for instance, 50 years ago series of photographic images were projected onto a screen or wall while simultaneously attempting to synchronized an audio tape that's being played, that was way back when there's no means of combining different kind of media to create a more whole unison multimedia, in fact, the technology (known as muxing) to do that does not yet exist at that time, and computers were not as rampant as they are today, and one's that did exist were large, and costly, and are geared toward researchers [2]. These days, the phrase Multimedia is linked (almost exclusively) with Computer, and it's constituent that constitute multimedia program are digital. Various forms of media are merged together to perform (in unison, on a computer) as a single unit, and are programmed using authoring software or programming languages. Various forms of communication are combined with multimedia to allow for a myriad of outcomes [1].

The concept of "cloud" and later on "cloud computing" has its inception with utility computing, an idea that sprouted out and popularized in 1961 [3]. John McCarthy stated; "*If computers of the kind I have advocated become the computers of the future, then computing may someday be organized as a public utility just as the telephone system is a public utility. ... The computer utility could become the basis of a new and important industry*". In 1969, a chief scientist of ARPANET project prophesied the internet stated that: "*As of now, computer networks are still in their infancy, but as they grow up and become sophisticated, we will probably see the spread of "**computer utilities**'"*".

Furthermore, according to NIST [4], Cloud computing is a model that enabled ubiquitous, convenient, on-demand access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be provisioned and discarded with minimal to no management effort or provider interaction. Cloud models constitute five essential characteristics, three service models, and four deployment models. Cloud Computing a latest concept to become popular in the computer industry with an ideal sharing of computing resources among a community of users. At present, cloud computing emerged as a web based computing technology that provides a freedom in the establishment of IT infrastructure, it's basically representing internet and web based applications, and works on user interactive software as simple as a web browser. The various cloud vendors do not require their own infrastructure; rather, they can rent or use third party providers [5].

Cloud computing has revolutionized the way computation and resources access is being done and accessed, hence the need for implementing the idea of utility computing whose advantages are undeniable for every business [6]. But despite all its hype and usage, the concept is pretty elusive and its definition quite vague. In simplistic terms, the cloud provides remote computing and storage services from a pool of shared resources to its consumers [7]. This underpins the idea of Internet of Things (IoT), which is rapidly changing our society to a world where "everything" is connected to the Internet, making computing pervasive like never before. This tsunami of connectivity and data collection relies more and more on the Cloud, where data analytics and intelligence actually reside. Cloud applications share the design objectives of Distributed applications. These objectives abbreviated as IDEAS refer to the Interoperability, Distributed scale out, Extensibility, Adaptivity and Simplicity aspects of distributed applications. Acting as a hybrid between the desktop applications and traditional web applications, a variation in the traditional application life cycle phases is required to architect cloud applications [8]. In fact, cloud computing is based on earlier models such as cluster computing, distributed computing, utility computing and grid computing in general [9]. It is worth pointing that; Cloud infrastructures consists of heterogeneous resources that are increasingly being utilized for hosting large scale distributed applications for diverse users with separate needs, the multitudinous cloud applications imposed varied requirements for computational resources along with multitude of performance implications. Hence, successful hosting of cloud applications obliges service providers to take into account the multiplicity existing in the behaviour of users, applications and system resources while respecting the user's agreed Quality of Service (QoS) criteria [10].

Cloud computing, with big data, is the biggest buzz these days, just like Internet and Web took us by surprise in 1990s and early 2000s, and smartphones shaping the new world in communication for the last decade or so, cloud computing is also expected to further shape a new way in which services would be provided to potential consumers, and businesses. It is a fact that almost all major tech giants like Microsoft, Google, Amazon and Apple provide cloud services to their consumers and even to other major businesses - for instance Netflix utilizes Amazon Web Services for hosting their streaming services [7]. As an emerging technology, Cloud Computing focused to isolate the delivery of computing services from its underlying technology. This new technology intends to render everything as a Service (XaaS). The cloud service model is categorized into three layers eminently Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). Cloud Consumers (CC) provision computing capabilities provided by the Cloud Provider (CP) and in turn pay whatever they use and as per the duration of usage. The resources in the cloud are abstracted and can be accessed using easy web interfaces. The interfaces reduce complexities by exposing a minimum set of capabilities based on target use cases. Cloud services shifts businesses cost from Capital Expenditure (CAPEX) to Operational Expenditure (OPEX) and are often associated with Utility Computing [11]. The potential of cloud computing and its convergence with technologies such as wireless networks, sensor technologies and mobile computing, allows for creation and delivery of newer types of cloud services [12].

When data sizes are constantly escalating and traffic patterns have becoming very unpredictable; Cloud is the viable option for enterprises to serve their customers with quicker response times, hence the businesses class can rely on cloud provider to host their applications, and thus they can focus on their core business goals. As number of Cloud Computing players keeps growing(due to spurring up of large number of Cloud providers with their offerings), the run time services provided to the cloud customers act as key differentiators [11].

In 2009, Google introduced "Google Apps" as plugins with its browser which enabled web and app developers alike to make their product and host it on Google servers as web application. At about same time, Microsoft together with Apple launched their version cloud storage products-OneDrive and ICloud respectively. Microsoft also introduced Microsoft Azure which lets consumers use it for variety of use cases from online storage to databases, APIs to full fledged web apps to fully hosted Linux and Windows virtual machines. Over a period of time, more players are entering this arena, cloud technology is expected to get cheaper and much more ubiquitous and useful, particularly for tech start ups and entrepreneurs [7].

2. RELATED WORK

Imran Bajwa et al [13] developed a middle-ware app that tackle existing problems of processing huge data in distributed environments in cluster computing, this provided a solution that can provide a quick way of processing at law cost.

X. Liu, X. Qiu, X. Xie, B. Chen and K. Huang (2012) presented a design and implementation of a small virtualised manager known as (LWIV Manager), the developed LWIV Manager provides an easy use and integration way to allocate computing resources of CPU, memory and network in the cloud. This was necessitated by the fact that; Cloud computing is changing the way IT resources of different kinds is delivered and consumed. Private cloud computing is more appealing than public cloud computing, and on the other hand, currently widely used open source virtualised managers cannot meet the demand of private cloud, especially the small scale private cloud, due to their bloated functionality and complexity, also there's a significant amount of fees involved when looking to integrate applications that's already running in the cloud.

J. Wu and T. Wang (2014) put forward a framework of combining SOA and cloud computing model, this was necessitated by respective advantages and interrelationship of SOA and the cloud framework. Furthermore, as a form of a model, SOA fused the development and the integration process which further diminishes the differences of various practical systems and realizes the Enterprise Application Integration (EAI).

The framework they presented has the characteristics of application flexibility, the features of easy extension, easy maintenance as well as low cost.

Pankaj Deep Kaur, Inderveer Chana (2014) proposed a QoS-Aware Resource Elasticity (QRE) system that permits specialist co-ops to make an appraisal of utilization conduct, and create mechanism that empower dynamic adaptability of cloud assets hosting the application components. This development was triggered by the fact that; cloud infrastructures comprising of heterogeneous assets are progressively being used for facilitating enormous scale distributed applications from different clients with discrete needs, and the varied demands imposed by multifarious cloud applications for computational assets alongside large number of execution suggestions. Their Experimental outcomes directed on Amazon EC2 cloud unmistakably exhibited the adequacy of their methodology while consenting to the concurred QoS traits of clients. Hence, Effective facilitating of cloud applications requires service providers to consider the heterogeneity existing in the conduct of clients, applications and framework assets while regarding clients concurred Quality of Service (QoS) models.

W. Zhu, C. Luo, J. Wang and S. Li (2011) Introduced the principal concept of multimedia cloud computing and presents a novel structure. They tended to multimedia cloud computing from multimedia-aware cloud (media cloud) and cloud-aware multimedia (cloud media) perspectives. To start with, they introduced a multimedia-aware cloud, which tends to how a cloud can perform distributed multimedia processing and storage and provide quality of service (QoS) provisioning for multimedia services. To accomplish a high QoS for multimedia services, their work proposed a media edge cloud (MEC) design, in which storage, processing(CPU), and graphics processing unit (GPU) clusters are introduced at the edge to give distributed parallel processing and QoS variation for different sorts of gadgets.

Selvaraj kesavan, Jerom Anand, Dr. J. Jayakumar (2012) considered the issue of multimedia management and processing, and created private controlled cloud architecture for the media which stores, process and convey the media substance to the validated clouders on the go. This advance-ment was necessitated by cloud clients having interest to access and share the media within the community by utilizing the gadgets with limited ability, and the inconceivability of prohibitive media content access and processing within the cloud community with the current cloud architecture. They likewise caught a portion of the critical points of interest of their architecture over the current technique.

Yunchang Liu, Chunlin Li, Youlong Luo, Yanling Shao and Jing Zhang (2018) proposed a scheduling scheme for multimedia services in multi clouds, and introduced a scheduling algorithm coordinating trust of entities thinking about the deadline, cost and trust necessities of multimedia services. In building up their scheme, they originally introduced a novel scheduling architecture and fabricated a trust model including both abstract and target trust to assess the trust degree of multimedia service providers. This advancement was set off by actuality that; security is a basic factor for multimedia services running in the distributed computing environment, and that trust can improve security level and mitigate attacks within distributed cloud computing environment, and further identified that current scheduling strategy for multimedia services in the distributed cloud computing environment don't incorporate trust mechanism when making scheduling choices. The consequence of their itemized reproduced experiments exhibited the viability and possibility of the proposed trust scheduling scheme.

Ashish Seth, Himanshu Agarwal, Ashim Raj Singla (2012) proposed a design that coordinate SOA and distributed cloud computing by utilizing SOA standards to make a generally essential plan and focus on how architectural support the use of distributed cloud computing. This was necessitated by the need to understand the existing and future state architecture before you begin selecting platforms and technology, and the fact that; SOA and cloud together can offer a total service based solutions for SMEs, and that both SOA and cloud manages conveying services to business improved agility, increase speed and reduced cost that can lead to greater innovation and effective returns on investment.

X. Liu, X. Qiu, B. Chen and K. Huang (2012) proposed a technique that covers the software involved with the entire cycle of Modelling as a Service (MaaS), Analysis as a Service (AaaS), and Execution as a Service (EaaS). They at that point introduced four scheduling algorithm. This improvement was set off by the way that; cloud computing paradigm draws in an increasing amount of Modeling and Simulation (M&S) experts to carry out their simulation in the cloud. Also, that, two issues, to be specific; the architecture of the Cloud-based Simulation (CSim) and the parallel simulation job scheduling in the CSim, should be addressed first to make the CSim down to earth. Their broad examinations on notable follows demonstrated that all the four calculations essentially beat their rivals.

Pankaj Deep Kaur et al (2014) proposed and designed infrastructure level mechanisms to provide dynamic resource elasticity for Cloud Based Intelligent Health Care Service (CBIHCS), that performs ongoing real time monitoring of client well-being data for diagnosis of chronic ailment, for example, diabetes. They utilized advanced body sensor components to gather user specific health data and store in cloud based storage repositories for subsequent analysis and classification Their Experimental results demonstrated that classification exactness of 92.59% is accomplished with their model framework and the anticipated examples of CPU utilization offer better opportunities for versatile resource flexibility.

[Claus Pahl, Antonio Brogi, Jacopo Soldam and Pooyan Jamshidi (2017)] carried out a state of the art review of Cloud Container Technologies, their study concluded that; the pattern towards cluster-based orchestration, joined with the interoperability of successful container technologies, additionally permits the management of profoundly distributed topologies of smaller virtualized devices beyond centralized clouds as in the edge cloud domain.

[Uchechukwu Awada (2018)] carried a survey study of Container Orchestration Tools and Platform-as-a service clouds. They opined that Container technologies like Docker and Rocket are examples of application containers; they were designed to packaged, isolate and run applications. They further stated that, this technology provides faster and better means for application deployment and execution, and simplify the complexity of achieving such state of the art deployment and execution, several container service platform have been recently introduced.

Cristian Ramon-Cortes et al (2018) introduced a framework, which is a mix of the COMP Superscalar (COMPSs) programming model and runtime, to effortlessly build and execute parallel applications in container based distributed processing platforms in a client transparent way with a view to provide a direct method to create task based parallel applications from sequential codes, and container management platforms that facilitate the deployment of applications in computing environments (for example, Docker, Mesos or Singularity). They build a prototype which combined COMPSs with different containers engines in the following scenarios:

1. A Docker Cluster

2. A Mesos cluster and Singularity in an HPC cluster.

The system provides researchers and developers with a simple method to actualize parallel distributed applications and deploy them easily.

From the literature review above, no study that optimizes multimedia processing over cloud cluster using modern container and container orchestration technology was carried out, this informs my drive to develop a technique that optimizes multimedia processing over cloud cluster so as to exploit modern container and container orchestration technology together with cloud platform AWS to provide or implement a slightly, albeit, optimized solution to multimedia processing over cloud cluster.

3. METHODOLOGY

Multimedia processing can be daunting at times considering the amount of horsepower needed for crunching complex numbers that make up either video or audio frames. Hence our approach to optimizing media processing over cloud clusters with the help of modern container and container orchestration technology involves coming up with an architecture that will let us achieve the desired result. Thus, in this segment we are going to outline the approach used as in the how to our work basically.

3.1. Method

3.1.1 Platform

Amazon Web Services (AWS) cloud platform was used to base our proposed technique and the architecture used. The Chosen cloud platform was used to run our proposed technique.

3.1.2 Components used

Within the AWS platform the following components were used;

1. Elastic Kubernetes Services (EKS)

EKS is a service within AWS platform that provides customers with a managed Kubernetes service basically EKS gives us flexibility to start, run, and scale Kubernetes applications on AWS cloud.

2. Elastic Compute Cloud (EC2)

EC2 is a service within AWS platform that allows users to rent virtual machines to run software either on-premise or the cloud. We've leveraged EC2 instances for running Kubernetes nodes.

3. Elastic Container Service (ECS)

ECS, a service within AWS platform that allows us to easily run, scale and manage containers on a managed elastic Kubernetes service.

3.1.3 Container

Containers were used to help package our deployable cluster application. We've used Docker as our container of choice.

3.1.4 Container orchestration

Kubernetes was also used as our container orchestration technology of choice to manage and coordinate pods deployment.

3.1.5 Architecture

The architecture that we've come up with and based our work on works is presented as follows;

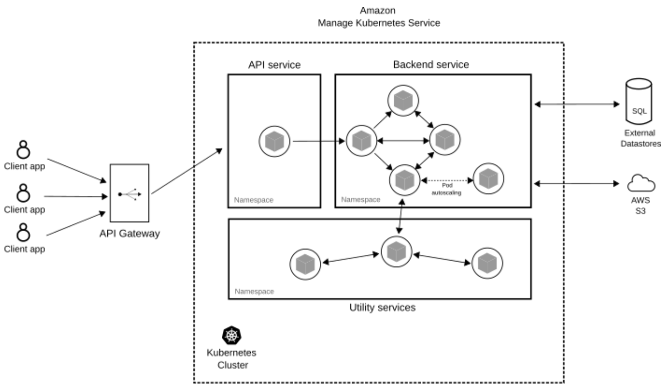

Fig. 1. Edge cluster architecture

3.1.6 Description

As seen from the above figure 1, client request first goes through API gateway which then routes the request to a desired edge cluster that is in close proximity. Edge clusters are replicas that are dispersed across different regions. Video and audio streams are retrieved from the request, sanitized and passed down to media processing service. Kubernetes manages and coordinates nodes within edge clusters. Software that does the actual multimedia processing execute and runs processing tasks.

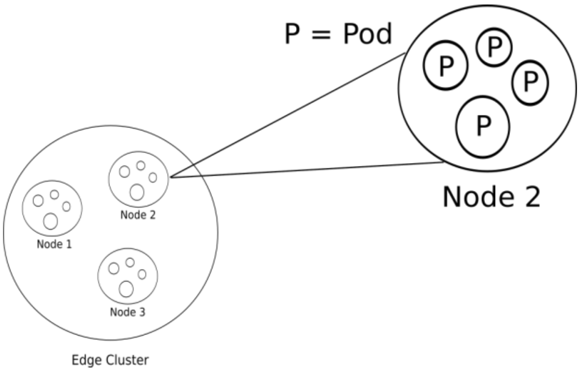

Fig. 2. Edge cluster

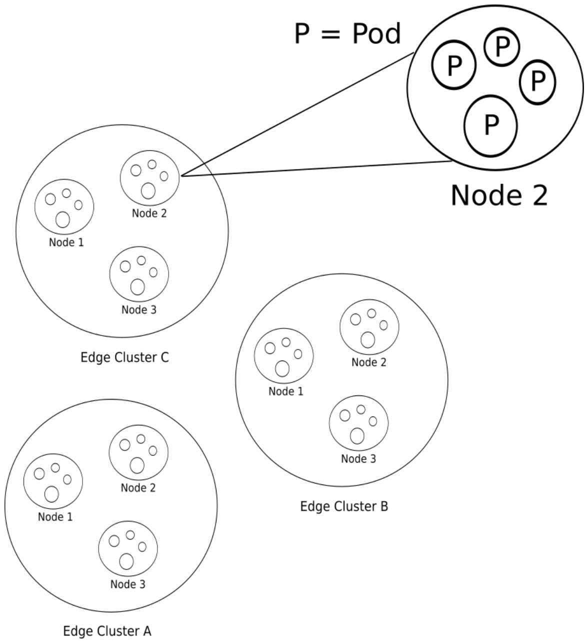

Fig. 3. Edge cluster group dispersed across different region

4. RESULT

The following is a graphical presentation of the result of the simulation of our technique to media processing over cloud clusters by taking advantage of AWS platform. We've discovered that there is a significant performance increase that occurred which enables us to achieve the optimization goals. Specific areas of concern are increase in processing speed to amount of workload before and after employing our technique.

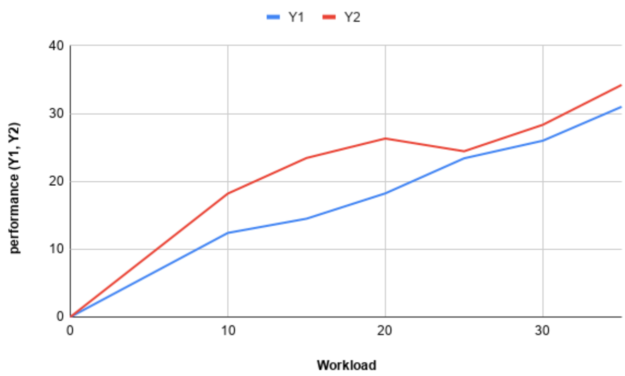

Fig. 4. Performance / workload

From above graph, Y1 denotes performance / workload before optimization, Y2 denotes performance / workload after optimization

5. CONCLUSION

Media processing has always been a tough, time consuming task needing huge processing power. Hence, an optimized method of processing media is of utmost importance, that's the reason we embark in this study to come up with optimized techniques to media processing while at the same taking advantage of an already available AWS infrastructure. During the course of our study we are able to come up with a design that allows us to achieve an incredible performance gain by bolting together AWS services to achieve our desired result.

6. REFERENCES

[1] Steinmetz, R., and Nahrstedt, K., 2011. Multimedia Fundamentals: Media coding and content processing. Multimedia Fundamentals. PHI Learning. 1

[2] Dowling, J. C., 2011. Multimedia Demystified. McGraw-Hill Osborne Media. 1

[3] Garfinkel, S., and Abelson, H., 1999. Architects of the information society: thirty-five years of the Laboratory for Computer Science at MIT. MIT Press. 1

[4] Mell, P. M., and Grance, T., 2011. Sp 800-145. the nist definition of cloud computing. Tech. rep., Gaithersburg, MD, USA. 2

[5] Reddy, E., 2014. Information Security in Cloud Computing., Vol. 3. International Journal of Computer Applications Technology and Research. 2

[6] De Donno, M., Giaretta, A., Dragoni, N., Bucchiarone, A., and Mazzara, M., 2019. "Cyber-storms come from clouds: Security of cloud computing in the iot era". Future Internet, 11(6), p. 127. 2

[7] Alzahrani, H., 2016. "A brief survey of cloud computing". Global Journal of Computer Science and Technology. 2

[8] Kaur, P. D., and Chana, I., 2011. "Evaluating cloud platforms- an application perspective". In Information Technology and Mobile Communication, V. V. Das, G. Thomas, and F. Lumban Gaol, eds., Springer Berlin Heidelberg, pp. 449–453. 2

[9] Sadashiv, N., and Kumar, S. D., 2011. "Cluster, grid and cloud computing: A detailed comparison". In 2011 6th International Conference on Computer Science & Education (ICCSE), IEEE, pp. 477–482. 2

[10] Kaur, P. D., and Chana, I., 2014. "A resource elasticity framework for qos-aware execution of cloud applications". Future Generation Computer Systems, 37, pp. 14 – 25. Special Section: Innovative Methods and Algorithms for Advanced Data-Intensive Computing Special Section: Semantics, Intelligent processing and services for big data Special Section: Advances in Data-Intensive Modelling and Simulation Special Section: Hybrid Intelligence for Growing Internet and its Applications. 2

[11] Kaur, P. D., and Chana, I., 2011. "Evaluating cloud platforms-an application perspective". In International Conference on Advances in Information Technology and Mobile Communication, Springer, pp. 449–453. 2

[12] Kaur, P. D., and Chana, I., 2014. "Cloud based intelligent system for delivering health care as a service". Computer Methods and Programs in Biomedicine, 113(1), pp. 346 – 359. 2

[13] Bajwa, I., Ali, A., and Naeem, M. A., 2011. "Processing large data sets using a cluster computing framework". Australian Journal of Basic and Applied Science (AJBAS), 5, 06. 2