A STUDY OF SCHEDULING AND MANAGING DATA BASED CLOUD COMPUTING SYSTEM

Hira Lal[1] and Dr. Arpana Bharani[2]
Department of Computer Science, Dr. A. P. J. Abdul Kalam University, Indore, Madhya Pradesh

ABSTRACT

Companies now have even another reason to think about shifting their data management systems to the cloud, owing to the modern age of "cloud computing." Based on previous successful cloud deployments, data management applications are a good match for the cloud. Cloud-based information management has a number of benefits, but it also has a number of disadvantages. This article provides an overview of cloud data management systems as well as a look at where researchers are currently focusing their efforts. Because cloud-based data management is always evolving, keeping an open attitude is essential. As the largest consumer of data management services, business is a prime candidate for further investigation. Research on cloud computing management information systems is a primary topic of this investigation. An extensive evaluation of management and business process journals is done in order to achieve the study's objectives. Because of this, it is necessary to get a deeper grasp of how information is sent. Systems function and how cloud-based information systems are evolving, results are examined. Based on the results of the evaluation, we've come to an agreement to use cloud-based MIS in our current investigation. Cloud-based management information systems (MIS) are becoming more popular with both small and large businesses (both public and private). That's why we're writing about cloud-based management information systems in this piece. Cloud-based MIS may be used in outsourcing and management because of its many benefits. When it comes to cloud computing, it's all about flexibility and scale not a new technology but one that is only now gaining traction, is being used by the majority of businesses as a means of storing and running their critical data. A focus on the most efficient method of data management is required in this case. When it comes to cloud data storage, Big Table and Data Store are two options that work in distinct ways. In order to achieve your goals, you must make use of the appropriate technologies.

KEYWORDS: Data Based Cloud Computing System, Cloud-based management information systems, cloud-based data management, cloud computing.

INTRODUCTION

Fog computing, data deduplication, and privacy protection are all highlighted in this chapter's cloud computing overview. This chapter also includes an outline of the research's overarching objectives, ambitions and contributions to SePeCloud's security and performance.

When it comes to moving IT ahead, the cloud's dispersed pool of flexible system resources and higher-level services is at the core of the equation. Availability and cost-effectiveness rely on how resources are allocated. This includes hardware virtualization with service-oriented architectures as well as autonomous utility computing" high-capacity networks and low cost computers and storage devices clearly, the cloud is an Internet platform for spreading computing resources and providing services, as shown by this graph. By using pay-per-use technology, cloud computing aims to make these benefits more accessible to consumers. There's no need for cloud service users to know all there is to know about the services they're using. To prevent business challenges, the user may combine various cloud services to create an upgraded solution using Service-Oriented Architecture (SOA). Standardised access to these resources in the SOA domain is possible via an established platform, allowing for maximum use of their resources and services. With cloud computing, parallel applications may be developed at far cheaper costs than those associated with conventional forms of computer technology. As may be seen in Figure 1, the general cloud computing paradigm.

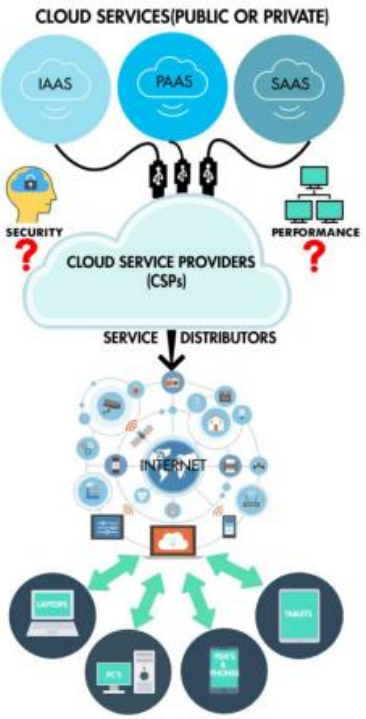

Figure 1 General Cloud Computing Metaphor

As with electricity, water, and gas, cloud computing has introduced the notion of on-demand computer resources that may be paid for as needed. In the vast majority of cases, data centres are in charge of overseeing and managing cloud services. These facilities also provide support to users in the event that they need it. With the emergence of low-cost data centres, the long-held goal of cloud computing as a utility will be realised. Making computers available as a service has been slowed by concerns about security. Cloud services rely heavily on data centres for their architecture. Cloud service providers, such as datacenters, provide a wide range of cloud services to its customers. Some of the most popular IT businesses provide pay-per-use services. Facebook, Amazon, Sales Force, Yahoo, Cisco, Microsoft, and Google are just a few of the more well-known brands on this list. The data centres for each of them are located in different parts of the globe. The licensing model known as "Software as a Service" (SaaS) makes software resources accessible on-demand. It is possible for a single server-side application to be accessed by several cloud services and many customers or clients. Using SaaS resources is beneficial to all major businesses because of the scope and cost constraints. Even though salesforce.com has been supplying cloud computing services for the last several years, it is the most well-known example

of SaaS. Google Apps and Dropbox are some well-known SaaS examples. Since the primary administration of the application is still with the cloud providers, authenticity and the security of personal data remain at stake. The safety of cloud users must also be taken into account, along with all of the other terms and conditions. Spoofing, phishing, and other forms of fraud are the most common security problems in the cloud. Cyber criminals may infiltrate a company's cloud network by using untrusted and illegal web browsers. Cloud providers are still responsible for making sure their clients are not damaged, though. A SaaS-based strategy to using cloud-based software is also considered to be the most widespread. Cloud users may strike deals with end users through the internet to save costs. The benefits and drawbacks of cloud computing should be carefully considered by everyone interested in using it, each of these elements. As a result, the cost of moving an organization's data between open or community clouds and paying for each computing asset that is utilised is likely to be higher. To estimate the expected financial impact of SaaS services, experts do a cost analysis. Customers who utilise telecom firms' end-user software solutions produce substantial income for these businesses. Many models are evaluated in order to generate the most value for the firm. Pricing is based on a variety of factors, including current market conditions. The cost-benefit analysis of amortization-based cost reductions is made easier with the help of SaaS. Data quality, integrity, and availability are out of the hands of customers that use cloud services, thus they depend on the security protocols of the cloud service providers. To guarantee that all terms and conditions are regulated in accordance with the agreed-upon settlement, a SLA must be signed. Cloud service companies have a lot of room to grow to achieve service level agreements is a major factor in their reputation (SLAs).

EXPERIMENTAL TOOLS

Different approaches and their implementation are covered in this chapter. It applied Cloudsim 3.0 and Eclipse IDE for simulation in the case of data centre analysis. Cloudsim output was discovered using Core object-oriented programming. Java is the programming language of choice for the Cloudsim simulation toolkit.

1. Cloudsim Simulator

In order to model cloud conditions, the CloudSim library was designed. Data centres, virtual machines, software, and users are just some of the things that may be defined using this framework. It also offers guidelines for scheduling and provisioning.

2. Eclipse Software

In computer programming, Eclipse is a prominent integrated development environment (IDE). Plugins may be added to the system to modify the appearance and feel of the workplace. The principal objective of Eclipse is to help in the creation of Java-based applications.

3. Java Programming Language

In order to have as few implementation dependencies as practicable, Java was meant to be a general-purpose computer programming language that is concurrent, class-based, object-oriented, and object-oriented. By enabling application developers to "write once, run everywhere" (WORA), the Java produced code may execute across all Java-supporting operating systems and devices. If you wish to launch Java programs on any Java virtual machine (JVM), you'll need to compile them into bytecode that can run on every platform.

4. Arduino Software

The Arduino IDE is a Java-based tool that operates on Windows, Mac OS X, and Linux platforms. For programming the Arduino board, it is important to employ this utility. Version 2 of the GNU General Public License regulates the IDE's source code. Using specific requirements for code structuring, the Arduino IDE is able to handle the programming languages C and C++. It came up with a strategy for scheduling jobs into fog nodes. Ford-fulkerson approach is utilized to locate the biggest data flow into our network. Using a lot of choices for scheduling jobs in the case of huge datasets is advantageous. The LM35 temperature sensor and an arduino prototype board are used in this real-time example of acquiring new data from temperature-based data.

MAXIMUM DATA FLOW

Here we utilized ford-fulkerson technique to identify maximum data flow into a network. The ford–fulkerson algorithm is a greedy method that computes the maximum flow in a flow

network. It is dubbed a "method" instead of a "algorithm" since the strategy of identifying augmenting pathways in a residual graph is not completely stated or it is specified in various implementations with varied running speeds.

In figure 2 depicts node organization, first layer is generally IoT/Sensor layer. Then, we introduced fog edge node or data center layer. Final layer is considered for calculating data that is produced from IoT/Sensor layer. This last layer comprises of master-slave model. For assessing load capacity and total output capacity in master-slave paradigm, employed ford-fulkerson method.

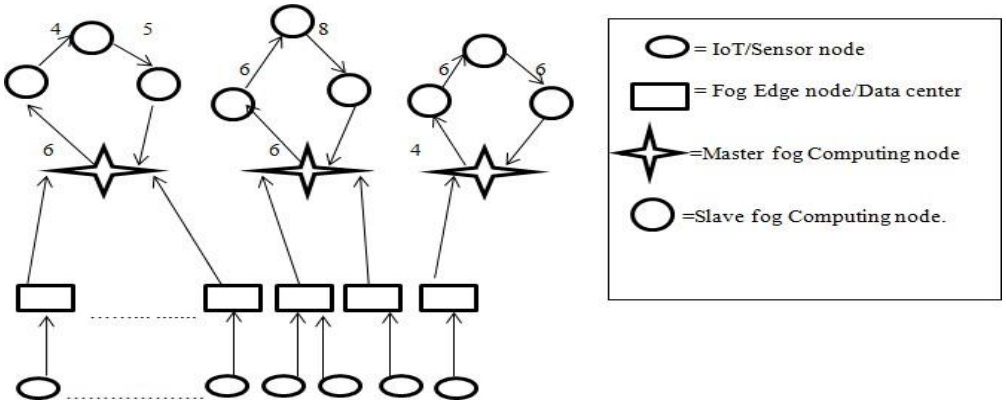

Figure 2: Node Organization into Different Layer for Fog Computing

In figure. 2 observed three master-slave model. It has supplied various load capacity in master-slave paradigm. Calculate varied load capacity based on figure 4.2 master slave model in figure. 3, 4, 5 here utilized ford-fulkerson method.

In the picture 3, 4, 5 presented several networks based on master-slave concept. It is a section of figure 4.2. In networks{\displaystyle G=(V,E)}GGG with flow capacity c, a source node su, a sink node si, u is a beginning point of edge and v is a terminating point of edge.

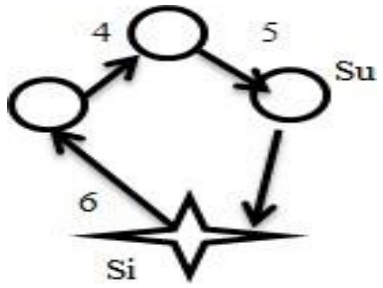

Figure. 3: Master-Slave network model with different network capacity

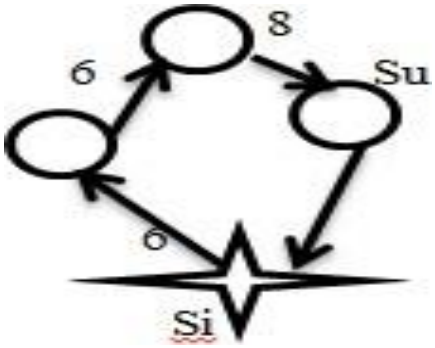

Figure 4: Master-Slave network model with different network capacity

If we use ford-fulkerson method the total network capacity and output sink node value are presented in the Table 2. The network capacity and output sink node values are helpful to figure out these nodes which are not given the duty. Moreover, the capability of the nodes that are able or not able to accomplish the allotted job.

TABLE 2: Finding Maximum Data Flow into Different Network Using Ford-Fulkerson Algorithm

Source node (Su)	Slave node (Si)	Maximum data flow in network
1	3	4
2	2	6
3	3	4

Real Time Example Fog Computing

It utilized temperature sensor LM35 that gets data from the surroundings. Here, it essentially want to check that instead of sending data to cloud if it processing data in our fog nodes in cloud load is balancing. For data transmission into fog nodes which is our local PC utilizing arduino prototype board. Using arduino ide and performing program, it gather data into our data sheet. Using distinct feature depending on different criteria, it eliminate the superfluous data. When we require data for real time services, it transmit into end user. Otherwise, we transmit into the permanent cloud storage. In this method, we have two benefits: one is superfluous data are purged into main storage, another is real time services readily given.

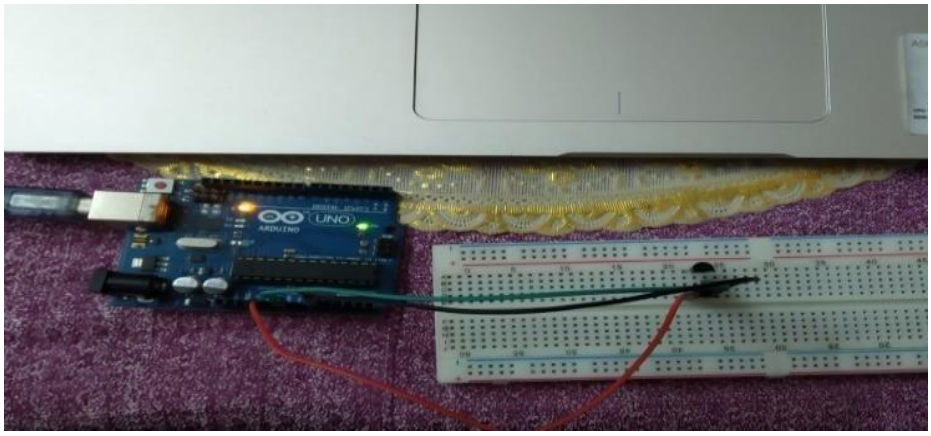

Figure. 5: Real Time Fog Computing Example.

Data Centric Analysis

In case of offering services from same data center, it is time demanding and also services is pretty bad. Moreover, upkeep of data center is more expensive. From the Table 3, it observe the same datacenter operates with two separate jobs. When one assignment is accomplished then other task is assigned. So, in this circumstance real time service offering is really extremely complex since one job depends on another activity and waited for finishing the work. In this circumstance, fog computing paradigm plays a crucial role and storage issue may be simply overcome. Because we know that fog is often operated as distributed processing.

TABLE 3: To Provide Service from Same Datacenter for Different Task

Cloudlet Id	Data Centre	Vm Id	Time	Start Time	Finish Time

0	2	0	80	0.1	80.1
1	2	1	160	0.1	160.1

TRANSMISSION TECHNOLOGY

Here we need to utilize MaaS (Monitoring as a Service) constantly since in our design, one step is dependant on another stage. Moreover, 5G internet capability is essential for data transmission over internet. Different forms of FDS (Fog Data Service) are also be required for interact conveniently in different levels of our suggested architecture. we may utilize LEACH (Low Energy adaptive Clustering Hierarchy) routing protocol for our Smart Grid node structure. TORA (Temporally Ordered Routing Algorithm) routing protocol may also be utilized for route development and management. Moreover, for our architectural communication, Bluetooth 4 may be employed in IoT devices and Fog nodes. For communication between Cloud and Fog Nodes, socket input output programming may be utilized. Many wireless communication protocols are established based on IEEE 802.15.4 Such as ZigBee for the aim of communication between fog and IoT devices. Moreover we may use another technology such as it is required to impose (caas) control as a service, (maas) monitoring as a service since one layer is controlled by another layer. Moreover data transfer from one layer to another layer 5G internet services is required. For adding security into IoT (internet of things), data block chain technology is beneficial and is a simpler solution. We may employ LORA as well as LORA WAN protocol for enabling long distance data transfer. Data gathering from IoT devices or sensors fiber optic sensors also may play a crucial role. Communication in fog and cloud layer socket input output programming is useful.

CONCLUSION

Cloud-based management information systems (MIS) have become a hot topic of debate because of the many benefits they bring. When it comes to adopting a new technology, there are several considerations that come into play. The benefits and drawbacks of cloud-based management information systems are well-known. Many benefits may be gained by using a cloud-based MIS, including the ability to manage data. Project management, strategy, and data processing all seem to be key obstacles to cloud-based MIS deployment. When businesses employ cloud-based

management information systems (MIS), customers are happier and more pleased. Additionally, data security has been a concern. In order to keep data safe, cloud service providers, businesses, and the government all have a role to play. An organization's ability to evaluate data access and operational activities is critical to understanding the system's benefits and limits. Customers benefit from interactions with small and medium-sized businesses. Customers' happiness and trust are predicted to rise dramatically as a result of cloud-based MIS. Cloud-based MIS has several advantages for corporate administration, customer interactions, data sharing, planning, and strategizing, as well as for data management and analysis. According to some scholars, only by prioritising governance and organisational duties can public service data be handled properly. Since data can be easily shared between stakeholders and business partners, inter-organizational communication has enhanced. A company's rapid expansion could be attributed to an increase in contact and proactive communication. The scale of the organisation, the needs of the industry, and the involvement of partners and stakeholders all play a role in cloud-based MIS adoption. The effect of data flow via third-party cloud service providers can only be accurately assessed with a well-structured data set and appropriate data sharing mechanisms. To implement cloud-based MIS data management, stakeholders and partners must work together.

REFERENCES

1. Ali, M, Bilal, K, Khan, S, Veeravalli, B, Li, K and Zomaya, A 2015, "DROPS: Division and replication of data in the cloud for optimal performance and security", IEEE Transactions on Cloud computing, pp. 1-14, ISSN: 2168-7161.

2. Balani, N and Ruj, S 2014, "Temporal access control with user revocation for cloud data", in 13th International Conference on Trust, Security and Privacy in Computing and Communications, pp. 336-343.

3. Bonomi, F, Milito, R, Zhu, J and Addepalli, S 2012, "Fog computing and its role in the internet of things", in Proceedings of the first edition of the MCC workshop on Mobile cloud computing, pp. 13-16.

4. Chandramohan, D, Vengattaraman, T, Rajaguru, D and Dhavachelvan, P 2016, "A new privacy preserving technique for cloud service user endorsement using multi-agents", Journal

of King Saud University Computer and Information Sciences, vol. 28, no. 1, pp. 37-54, ISSN: 1319-1578.

5. Deshmukh, AR, Mante, RV and Chatur, PN 2017, "Cloud Based Deduplication and Self Data Destruction", in IEEE International Conference on Recent Trends in Electrical, Electronics and Computing Technologies (ICRTEECT), pp. 155-158.

6. Devi, KJ and Kanimozhi, S 2014, "Efficient user revocation for dynamic groups in the cloud", International Journal of Engineering and Computer Science, vol. 3, no. 2, pp. 3938-3942, ISSN : 2319-7242.

7. a.g.ashmita et al (2018) "Data De-Duplication on Encrypted Data Lake in Cloud Environment", International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 IJERTV7IS030116 (This work is licensed under a Creative Commons Attribution 4.0 International License.) Published by : www.ijert.org Vol. 7 Issue 03, March-2018.

8. Fan, CI, Huang, SY, and Hsu, WC 2015, "Encrypted Data Deduplication in Cloud Storage", in 10th Asia Joint Conference on Information Security (AsiaJCIS), pp. 18-25.

9. Gadhve, S and Naidu, D 2016, "Self destruction model for protecting data in cloud storage based on data storage center", International Journal of Engineering and Innovative Technology, vol. 6, pp. 3, pp. 32-37, ISSN 2277-3754.

10. Gan, Q, Wang, X and Wu, D 2017, "Revocable Key-Aggregate Cryptosystem for Data Sharing in Cloud", Security and Communication Networks, vol. 2017, pp. 1-11, ISSN:1939-0114.

11. González-Manzano, L and Orfila, A 2015, "An efficient confidentiality-preserving proof of ownership for deduplication", Journal of Network and Computer Applications, vol. 50, pp. 49-59, ISSN:1084-8045.

12. Greveler, U, Justus, B and Loehr, D 2011, "A privacy preserving system for cloud computing", in 11th IEEE International Conference on Computer and Information Technology (CIT), pp. 648-653.

-o-