Operating System Assignment -1

Name: Atharva Tekade

Bits ID:2021mt70022

Group 24:

Group Members:

1.Grandhi Sai Vamsi

2. BAGATHI UDAYASRI

3. I TEJESH

4.Atharva Raman Tekade

**1.What is the problem that the thesis deals with?**

**Solution:**

**Thesis deals with the storing and retreiving files in cloud storage and the algorithm used for storing and retreiving trillions of files in cloud storage.**

2. What is the research methodology used?

The research work aims to develop an application to store and retrieve the user files in the open-source cloud environment. To build a cloud computing environment, an organization uses existing infrastructure and installs a cloud computing platform called OpenStack. OpenStack is popular open source cloud computing platform that provides different services for different task. To create public or private clouds very quickly and manage virtual machines, OpenStack is very useful. Users need not able to maintain the resources needed by them. Based on the demand, users can request the needed infrastructure resources to use. With OpenStack, there is no data loss. It has the good security features to protect all the user data. 10 10 OpenStack is an open-source cloud that manages compute, storage, and networking resources in a data centre. The service used in this research is infrastructure as a service and storage as a service. Through a dashboard or command line, the user can access the services provided by OpenStack. The important OpenStack component used in this thesis is shown in figure 1.3. Figure 1.3 shows the working principle of the proposed work in OpenStack. Client sends the request to the horizon dashboard, using keystone component, the username of the user and password is checked. Nova allocates the instances in the compute node. Volume is allocated to the user using cinder component, where user can insert and search their files. The main component used in this thesis is cinder, where cinder is a block storage service used for storing and retrieving a large number of files or applications. Cinder allocates, persistent, readable, and writable block storage that could be utilized by the user for creating compute instances

3. What is the proposed solution??

**Solution:**

**In the cloud storage system, more user files are stored on different compute nodes. To organize and maintain files in the cloud storage system, digit binate algorithm and digit compact prefix algorithms are used in the backend of the compute node and controller node. The algorithm time complexity is analysed and discussed. The performance evaluation done for both digit binate and digit compact prefix algorithms. The implementation is done in OpenStack private cloud. The digit binate algorithm shows the minimum time in storing and retrieving compared to digit binate algorithm with different compute node instances and files. The implementation is done in OpenStack private cloud.**

**The Internet is an important factor in the cloud to provide services to users. Router is needed on the Internet to determine the packet’s next-hop route. This is achieved by checking the destination address of the incoming IP packet in a routing table. Because of a higher number of nodes in the cloud scenario, there is a frequent change in the network topology. The new node information is to be updated in all routing tables available in the network. The frequent updating in the routing table affects the process of finding the most matching entry due to which a bottleneck created in the iv iv routers on the Internet. The thesis describes the fast and efficient hybrid lookup algorithm that prevents this bottleneck by splitting the octets into four. The implementation is done in SDN Ryu controller and Mininet virtual topology creator. The proposed algorithm shows minimum lookup time and search time compared with the binary search tree**

**The digit compact prefix is a data structure that uses the digits to organize and search the files in the cloud storage volume. When the volume is large with more number of users, searching their filename is very difficult and time-consuming. In an open-source cloud system, all the details about the 100 100 filename are stored in the controller node. When the number of users storing their files increases, it takes more time to search and retrieve the file details. In the digit compact prefix method, storing and indexing provides a better solution. To store the files in digit compact prefix method, the filename given by the user is converted to a random hash value. User name, the virtual machine and compute node that the user is accessing are stored as metadata in the controller node. The hash value of the filename is organized and stored in the digit compact prefix data structure. Search is performed in the controller node by giving the username and the filename. The controller finds the virtual machine and the compute node where the users have stored their files already. Insertion operation The open source cloud storage provides on-demand services to the user to save their data at anytime and anywhere. To insert files into the cloud storage, four things are very important, namely cloud user, controller, compute node and storage. The interaction is shown in the figure 5.2. Filenames are considered as the keys to organize the proposed data structure. User should be first authenticated in the cloud system by the keystone service, once authenticated the user should choose the virtual machine instances. The chosen virtual machine can be allocated to any of the compute nodes by the controller. The compute node number and the virtual machine number are stored in the database of the compute node. The details are specific for the particular user. Now, the user can use the services provided by the cloud. When using the storage service, the files will be stored in the instance volume in cinder. In the backend, the file names are organized using the digit compact prefix structure. 10**

**In implementing the digit compact prefix method, linked list is used to create nodes For the Insert and search function, filename with key . ree cases shown in the value is passed as an argument. In insert function, the (algorithm. case (if the node value not match with any value(case (ii , if this is the first node, create a new node down to the root node till the hash value of the corresponding filename(case(iii , if file name key value is present, traverse down from the root In the search function, two cases shown in the algorithm . case (i)f the filename key value is present in the root node and child nodes case(ii) if the filename key value is not present OpenStack is installed on a personal computer running the Windows 7 operating system with the Intel i5 processor, 8GB RAM and 500 GB hard disk. VirtualBox is used to create Virtual Machine as guest OS. The installation is based on the OpenStack Icehouse. The OpenStack platform is implemented with many projects and services as shown in figure 5.4. Each project service enables different functions within the cloud scenario. Some of the projects are computing, block storage, image, identity, and web user interface. The OpenStack is deployed in two-node setup. One node is the controller and another node is a compute node.**

4. What is the future scope of this?

Solution:

1. 1. **Better Cloud Services**

To begin with, the future of cloud computing in education and industries around this domain will be able to witness the power of sound cloud computing services. Better cloud services can be categorized into three types. You have infrastructure as a service, platform as a service, and software as a service. These are three important types of services used by user-friendly and old organizations. When more and more services are able to support these facilities, it becomes easier for organizations to shift. In the long run, more schools and colleges will start to leverage the perks of cloud computing.

1. 2. **Security**

One of the most important areas of discussion around cloud computing technology would be security. It is important for service providers to ensure that the data is stored both safely and securely. This calls for more skill and knowledge around cloud computing. Indeed, this is one of the major reasons why the

The future of cloud computing jobs is bright. Companies will need skilled specialists who can promise security in all stages of cloud services. It is also important for service providers to ensure that cyber attacks are kept at bay. Even small companies that don’t prioritize security need to change their business model. Studies and technologies that stress the improvement of the future scope of cloud computing security will find bigger and better platforms to prove their theories in the upcoming days.

1. 3. **Modular Software Development**

Modular software development is considered a cornerstone that can bend or break the future of cloud computing technology. Why? For companies to be able to leverage the cloud technology code and applications have to break into smaller chunks, which are not coupled with one another. This makes it easier for developers to upload their works into the cloud. Also, it makes it simpler for people to access the content from Remote locations. On the other hand, security and the availability of features improve when code is modular. In the long run, companies that invest in modular code can save more money too.

1. 4. **Market Growth**

By the end of 2018, the market value of cloud computing crossed 128 billion USD. This is not going to drop in the upcoming days. Mainly because the future use of cloud computing is only going to increase and not decline. Even big brands and Fortune 500 companies are migrating their services and products to the cloud. Meanwhile, the spending patterns of companies around on-premise infrastructure have reduced. These are pointers for a better and stronger future in the cloud. Undeniably, the market and the future of cloud computing a long term forecast, is inevitable.

1. 5. **Virtualization**

Finally, we need to focus on the future of virtualization and cloud computing. This is an interesting bond that has plenty of scope for research and development. To begin with, when cloud computing and virtualization are brought together, a unique architecture is required. This could be a development that maps to the qualities of the Computing Cell. And, the computing cells is known for its consistent need for finer and sophisticated software infrastructure, which is paired with intricate features like encryption, third party authentication, efficient and reliable network segmentation, and data management.

All these improvements need to be provided across all channels in the cloud. And, it is important for the service providers to ensure that their cloud technology meets these standards. Virtualization and cloud can bring to light a hybrid IT system, which is a challenge and a big problem today. Another important reason why virtualization and cloud computing is crucial would be budgeting. When hardware expenses are limited and cut down to the services offered by the cloud, there will be more cost-cutting and savings. In fact, the way data needs to be stored can be controlled using algorithms. It will no longer need the support of humans.

1. B. **Scope of Cloud Computing**

The above predictions prove that the scope of growth for cloud computing is immense. More and more organizations need to prioritize the use of this technology. In fact, they need to restructure and invest in coding standards that can support seamless migration into the cloud. Also, cloud computing is strongly associated with concepts like the internet of things. When data gets stored in the cloud, it becomes easier for IoT to ensure performance, security, and functionality. The only limitation would be the speed of the network, which controls the pace at which data is gathered and processed. If the network is fast, everything else about the use of cloud computing will fall in place.

5) Propose suggestions or alternate solutions to the problem that the thesis tries to solve.

It combines the private and public cloud storage services for storage of data. The selection of cloud system is based on the sensitivity and importance of data to be stored. If a data is related to legal compliance or organization’s sensitive information, it is stored on a private cloud storage. However, if the data is harmless to the user’s privacy and security, it can be uploaded to a public cloud. This helps in saving the cost of purchasing unnecessary cloud storage space.

Cloud databases provide facilities for large scale data storage and retrieval of distributed data. However, the current access control techniques provided in database systems for maintaining security are not sufficient to secure the private data stored in public cloud databases. In this paper, a new secured data storage algorithm for effective maintenance of confidential data is proposed. To perform storage and retrieval operations of data in the cloud data storage effectively, map reduce algorithms are developed in this work which performs data reduction and fast processing. In order to consider the temporal nature of documents to be retrieved, we propose a new algorithm called Temporal Secured Cloud Map Reduced Algorithm which integrates temporal constraints with map reduce algorithms and also the chaining Hill Cipher encryption algorithms which is proposed newly in this work. The main advantage of the proposed algorithm is that they reduce the processing time and maintains security effectively. The experimental results obtained from this work depict that the proposed model is optimizing cost and it ensures data security.