An Approach to Live Application Server Migration in AWS Public Cloud

Nisha Rathod¹ Prof. Sandip Chauhan²
¹M.E. Scholar ²Assistant Professor
¹²Department of Computer Engineering
¹²KITRC, Kalol (N. G.), Gujarat, India

Abstract— virtualization technology is at the heart of cloud computing technology. One particular physical host may contain more than one virtual machine. In cloud computing, Virtual machine may require to migrate on other physical host to perform maintenance of physical host, to run server consolidation process, load balancing and fault tolerance. Online migration can help the client to get uninterrupted service while offline migration increases the downtime. From all available cloud service provider in the market, in this paper Amazon has been chosen to implement the live VM migration. In this paper we described different procedure to perform live migration, an approach to implement live migration in AWS public cloud.

Keywords: Virtual Machine Migration, Data center, Cloud Computing, AWS (Amazon Web Service)

I. INTRODUCTION

Cloud computing is giving its services via internet on pay per use basis. Cloud computing has various service models like Infrastructure as a Service (IAAS) providing computing, storage and network services. Platform as a Service (PAAS) providing required development and runtime environment and Software as a Service (SAAS) providing ready made utility services to end users. In IAAS model, cloud service provider launches more than one virtual machines in a single datacenter and running their application in virtual machine host. This offers infrastructure services like ready to use network, storage and computing facility resources. Service provider manages all the underlying complexity of installation and maintains hardware or software resources. To perform scheduled maintenance activity on a host, to balance the load, to survive in faulty situation or to consolidate server, virtual machine running on that particular host needs to be migrated to another host. In live migration downtime must be smallest so that client getting service from Virtual host does not feel any interruption in their services. Cloud infrastructure consists of varied, distributed and virtual resources.

In Cloud computing there are various deployment models – Private Cloud, Public Cloud, Community Cloud and Hybrid Cloud.

A. Public cloud [12]

These types of clouds are made accessible to the overall population by a service provider who is responsible for hosting the cloud. By and large, open cloud suppliers like Amazon AWS, Microsoft and Google own and manages the infrastructure offer access over the Internet. With this model, clients have zero ability to see or control over where their resources are located. It is essential to note that all clients on public cloud share the pooled resources.

B. Private cloud [12]

Private cloud will be cloud framework committed to a specific association. Private cloud permit organizations to host applications in the cloud, while tending to concerns with respect to information security and control, which is frequently lacking in a public cloud environment. It is not imparted to different associations, whether oversaw inside or by an outsider, and it can be facilitated inside or remotely.

C. Community cloud:[12]

A community cloud is a is a multi-tenant cloud service model that is shared among several organizations and that is governed, managed and secured commonly by all the participating organizations or a third party managed service provider. Clouds are a hybrid form of private clouds built and operated specifically for a targeted group. These communities have similar cloud requirements and their ultimate goal is to work together to achieve their business objectives. The goal of community clouds is to have participating organizations realize the benefits of a public cloud with the added level of privacy, security, and policy compliance usually associated with a private cloud. Community clouds can be either on-premise or off-premise.

D. Hybrid Cloud [12]

Hybrid Clouds are a structure of two or more clouds (private, group or open) that stay unique substances however are bound together offering the upsides of various arrangement models. In a hybrid cloud, you can influence outsider cloud suppliers in either a full or fractional way; expanding the adaptability of computing. Enlarging a customary private cloud with the assets of an public cloud can be utilized to deal with any surprising surges in workload.

From this all deployment model we have chosen public cloud model to implement the live migration facility because this deployment generally model covers more users.

In this paper, section II describes various techniques to perform live VM migration, section III describes introduction to Amazon Web Services and EC2 instance launch procedure and section IV describes implementation strategy to live VM migration. Section V describes conclusion and future work.

E. Live VM migration advantages: [8]

1) Reduce IT costs and improve flexibility with server consolidation.
2) Decrease downtime and improve reliability with business continuity and disaster recovery.
3) Increase energy efficiency by running fewer servers and dynamically powering down unused servers with our green IT solutions
4) Accessing more processing power (in the sense of load balancing).
5) Exploitation of resource locality (for performance), resource sharing (meaning sharing of expensive or rare resources - such as telescopes or medical equipment – or large amounts or free memory by processes over a network), fault resilience.

6) Simplified system administration and mobile computing (for instance as used by commuters from office to home).

II. LIVE VM MIGRATION TECHNIQUES

Live VM migration supports minimum downtime and thus very less or no interruption to the client. It can be able to manage consistency of source host VM at destination host. Live migration transfers the whole OS state including application running on that VM, network connection, source host VM main memory and CPU state. To perform live migration various techniques has been introduced. Some of the techniques described below.

A. Pre-copy memory approach: [1]

In these approach after reserving the necessary resources at destination host, initially all the memory pages are transferred to the target host while VM is still running at source host. Meanwhile the pages which are being updated at source being marked as dirty pages and dirty pages again required to be sent to the target. If the application is more write intensive than more than one round (iteratively) of transferring dirty page is required. Once some minimum dirty page threshold count is reached at source than small stop and copy phase will be started. During this phase source host is suspended and remaining dirty pages at source are being copied to the target to synchronize the source and destination. Once VM is ready at target, it can be resumed.

B. Adaptive - Memory Compression [2]

In this approach, first pages to be transferred are being compressed at source and at destination again decompress to get the original page information. Extra overhead is there due to compression and de-compression but less amount of data to be transferred so advantageous in low bandwidth network.

C. Improved-Pre copy Time-Series Based Approach [3]

In this approach, extra bit map is being added to reduce the number of pages to be transferred in each phase. Other variations to these algorithm is specified either to reduce downtime, to reduce amount of data to be transferred or to reduce number of pages to be transferred in each iteration. But for simplicity and minimum downtime as a preference we chose pre-copy approach to perform live VM migration.

D. Post-Copy Memory Approach: [4]

Post-copy migration defers the transfer of a VM’s memory contents until after its processor state has been sent to the target host. This deferral is in contrast to the traditional pre-copy approach, which first copies the memory state over multiple iterations followed by a final transfer of the processor state. The post-copy strategy can provide a “win-win” by reducing total migration time while maintaining the liveliness of the VM during migration.

Comparison is based on the experiment conducted on Web server, in which in guest VM six different clients are working on the VM.[13]

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Pre-copy</th>
<th>Post-copy</th>
<th>Adaptive Memory Compression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pages Transferred</td>
<td>180000 to 200000 bytes</td>
<td>120000 to 140000 bytes</td>
<td>600 mb</td>
</tr>
<tr>
<td>Total Down Time</td>
<td>600 ms</td>
<td>800 ms</td>
<td>600 ms</td>
</tr>
<tr>
<td>Total Migration Time</td>
<td>10 sec</td>
<td>8 sec</td>
<td>100 sec</td>
</tr>
</tbody>
</table>

After observing different parameter values mentioned in the above table, we will consider the algorithm which gives the minimum downtime so that migration can be performed in a transparent way to the client.

Downtime is similar in pre-copy and compression algorithm, but at the same time total migration time is also less in pre-copy approach. So we will be giving preference to pre-copy approach to implement live VM migration in AWS public cloud.

III. AMAZON WEB SERVICES [14]

Amazon Web Services offers a broad set of global compute, storage, database, analytics, application, and deployment services that help organizations move faster, lower IT costs, and scale applications. These services are trusted by the largest enterprises and the hottest start-ups to power a wide variety of workloads including: web and mobile applications, data processing and warehousing, storage, archive, and many others.

Amazon Web Services (AWS) provides trusted, cloud-based solutions to help you meet your business needs. Running your solutions in the AWS Cloud can help you get your applications up and running faster while providing the same level of security that organizations like Pfizer, Intuit, and the US Navy rely on. AWS also provides resources around the world, so you can deploy your solutions where your customers are.

![List of available services provided by AWS](image-url)

Above mentioned list describes various services like Elastic Cloud Computing (EC2), Elastic Block Storage (EBS), and Elastic Load Balancing (ELB) and similarly it provides various API for different languages to develop application.

To perform live VM migration we have to first launch a VM instance. Following figures shows the way to create an EC2 VM instance graphically.

Steps to launch a new VM instance in AWS:
- Choose EC2 service from AWS list shown in figure 3.1
- Choose an appropriate Amazon machine image.
- Choose an instance type like micro, small, large, etc.
- Configure instance details like network, subnet, shutdown behavior, monitoring, tenancy etc.
- Add required storage and set storage configuration like size, type (SSD or magnetic storage), edit the settings of the root level
- Tag the instance with key-value pair.

- Configure a security group via a set of firewall rules which controls the traffic to the created instance. We can set protocol type, port range or source IP address.

After configuring all above mentioned details it will ask for review instance launch.

While clicking a launch button it asks for choosing the existing key pair or creating the new key pair. It is asymmetric way to securely connect to the instance. In which AWS stores the public key and a private key file stored by user. To log into the windows image private key pair file is required to obtain the password. For Linux Machine Image, private key file is used to securely SSH into the instance.

After successful completion of the above mentioned steps in our AWS account we can see the running
instance with its entire predefined configuration.

Fig. 3.2: Successful launch of an instance in AWS account
To launch an instance programmatically instead of graphically, AWS provides a toolkit containing API required for developing the code to launch an instance. We have chosen Java as a development language and Eclipse Luna as an IDE. Development machine should be configured with the latest version of Java SDK and Eclipse Luna. To get support of AWS Java SDK toolkit, configure AWS toolkit with Eclipse Luna. Following screen shows the coding screen with Eclipse supporting AWS toolkit indicated by orange color logo in toolbar.

IV. APPROACH TO PERFORM LIVE APPLICATION SERVER MIGRATION IN AWS PUBLIC CLOUD
We have launched a VM with Ubuntu OS image in our account. VM instance has configured with Apache web server and Tomcat application server. To connect to the Ubuntu image instance from a Windows desktop, Putty is required. Putty asks for a private key-pair file in .pk format so PuttyGen is required to convert .perm file to .pk file.

Fig. 3.3: Connected Ubuntu image
Fig. 3.3 shows the connected Ubuntu machine image. Apache web server has been started. To perform live application server migration right now following scenario is implemented.

<table>
<thead>
<tr>
<th>Application Server Instance 1</th>
<th>Application Server Instance 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web Server forwards request to the application server</td>
<td></td>
</tr>
</tbody>
</table>

User gives request to the Web Server
Over here it is assumed that both the application server instances have similar configuration. Both instances are deployed with two simple Java server pages to check the responding instance. Whenever a user gives a request to the web server, it will forward the request to the server instance 1 and it will respond. Following figure shows the user request and response generated by instance 1.

Fig. 3.4: Response served by instance 1
As shown in fig. 3.4 right now user request is served by instance 1. It also maintains session detail containing session ID and user name.

If planned shut down for maintenance purpose or auto-shutdown of instance 1 happens, following figure shows the planned shutdown of instance 1.

Fig. 3.5: Planned Shutdown of instance 1
As per live migration of application server user should start to get response from application server instance 2. Negligible downtime may occur while migrating session data from instance 1 to instance 2. It is mentioned as negligible because the user does not feel any interruption while movement happens. Following figure shows after shutting down instance 1 user started to get response from instance 2.
It is also possible to programmatically fetch and list id of all running VM in a particular account and then generate request to migrate a particular VM by entering VM id. Following figure shows the listing of all available ids in particular configured account and asking for VM id to migrate.

![Listing ids of running VM in particular account](image)

**Fig. 3.6:** Listing ids of running VM in particular account

V. CONCLUSION

In this paper one approach which is similar to precopy approach has mentioned to perform application server migration which gives negligible downtime in AWS public cloud. Thus we can say it a live application server migration. This work can be expanded by securing a migration data because cloud environment is susceptible for various security risks.

REFERENCES


[14]aws.amazon.com