"DARWINBOX"

-

"VPN CONNECTION BETWEEN AZURE AND AWS USING GATEWAYS."

PREPARED BY

Kunal Sharma (2018BTechCSE109)

FACULTY GUIDE

Dr. Suman Saha

Department of Compute Science Engineering
Institute of Engineering and Technology (IET)
JK Lakshmipat University Jaipur

May 2022

"DARWINBOX"

-

"VPN CONNECTION BETWEEN AZURE AND AWS USING GATEWAYS."

Practice School-2

Submitted in partial fulfilment of the requirements

for the degree of

Bachelor of Technology in Computer Science Engineering

By:
Kunal Sharma (2018BTechCSE109)

Faculty Guide:
Dr. Suman Saha



Department of Computer Science Engineering
Institute of Engineering and Technology (IET)
JK Lakshmipat University Jaipur

May 2022

TABLE OF CONTENTS

Topic	Page No.
CERTIFICATE	iii
ACKNOWLEDGEMENTS	iv
ABSTRACT	v
LIST OF ACRONYMS	vi
CHAPTER 1: INTRODUCTION	
1.1 About Organization	1
1.2 Purpose	2
1.3 Objectives	2
1.4 Solutions	3
CHAPTER 2: TECHNOLOGY STACK	
3.1 Microsoft Azure	5
3.2 Amazon Web Services (AWS)	5
3.3 Amazon Redshift Cluster	5
3.4 Virtual Network Gateway (Azure)	5
3.5 Customer Gateway (AWS)	5
3.6 Virtual Private Gateway (AWS)	5
3.7 Site To Site VPN (Azure/AWS)	5
3.8 Local Network Gateway (Azure)	6
3.9 Internet Gateway (AWS)	6
3.10 Route table (AWS)	6
CHAPTER 3: WORKFLOW / RESULTS	7
CHAPTER 4: CONCLUSION	29
REFERENCES	30

CERTIFICATE

This is to certify that the Practice School-2 project work entitled "**DarwinBox-VPN Connection between Azure and AWS using gateways**" submitted by **Kunal Sharma (2018BTechCSE109)**, towards the partial fulfillment of the requirements for the degree of **Bachelor of Technology in Computer Science Engineering** of JK Lakshmipat University Jaipur is the record of work carried out by them under my supervision and guidance. In my opinion, the submitted work has reached a level required for being accepted for Practice School-II examination.

Dr. Suman Saha
Associate Professor,
Department of Computer Science Engineering
Institute of Engineering & Technology (IET)
JK Lakshmipat University, Jaipur

Dr. Gireesh Kaushik
PS-2 Co-ordinator and Asst. Professor,
Department of Computer Science Engineering
Institute of Engineering & Technology (IET)
JK Lakshmipat University, Jaipur

Date of Submission: 27-05-2022

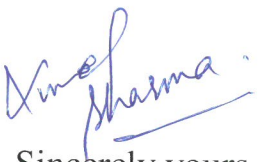
Prof. Dr. – Sanjay Goel
Director
Institute of Engineering & Technology
JK Lakshmipat University Jaipur

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to Dr. Suman Saha, Associate Professor, Department of Computer Science & Engineering for his valuable time and guidance that made the project work a success.

My heartfelt thanks to Mr. Harsh Goyal for his support during the project.. Thanking all my friends and all those who have helped me carrying out this work directly or indirectly without whom working on this project work was not possible.

I would also like to thank sincerely to Dr. Sonal Jain, Head of CSE Department and Dr. Sanjay Goel, Director, Institute of Engineering and Technology, JK Lakshmipat University for giving us a platform to carry out the project.



Sincerely yours,

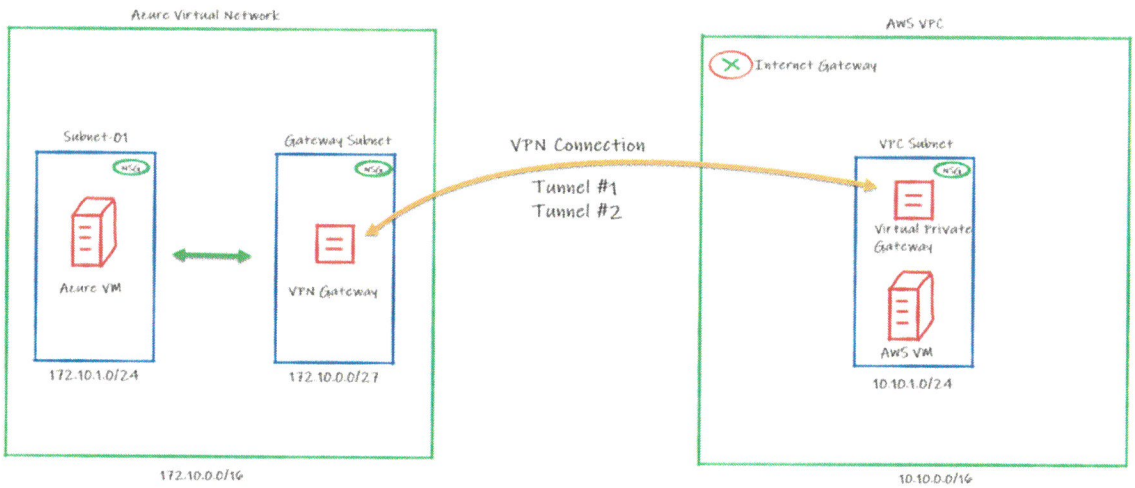
Kunal Sharma (2018BTechCSE109)

ABSTRACT

Multi-cloud architectures are an ever-increasing common design in the IT world and there will likely become a point when you want to setup a fast-direct resilient VPN connection between your public cloud providers such as Azure and AWS. Before the 6th February 2019, AWS did not support IKEv2 so to setup a site to site VPN with Azure you normally had to set the VPN connection with a 3rd party network appliance or a VM running software that supported it such as RRAS. Since the change you control the IKE version by updating the initiator end (Azure) and the AWS side endpoint will negotiate the session on the same protocol.

We will build resources listed below for both Azure and AWS.

1. Azure Virtual Network Gateways = AWS Virtual Private Gateways
2. Azure Local Network Gateways = AWS Customer Gateways
3. Azure Connections = AWS Site-to-Site VPN Connections



Regarding the high availability, please note that on AWS, by default a VPN connection always will have 2 Public IPs, one per tunnel. On Azure it doesn't happens by default and in this case you will be using Active/Passive from Azure side.

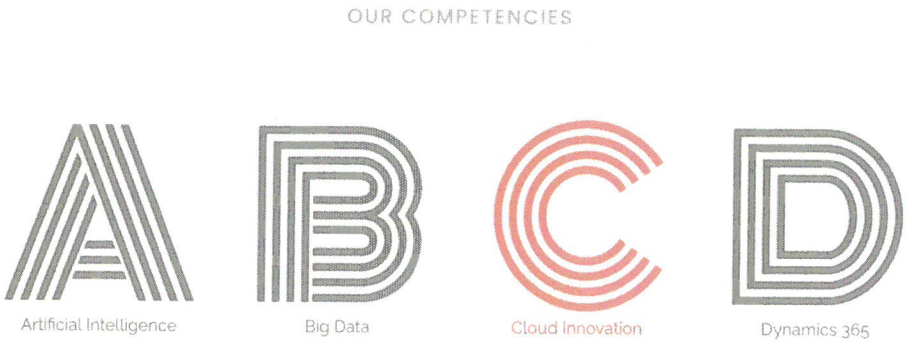
This means that we will be setting only one "node" from Azure VPN Gateway to establish two VPN connections with AWS. In case of a failure, the second node from Azure VPN Gateway will connect to AWS in a Active/Passive mode.

List of Acronyms

Abbreviation	Full form
RG	Resource group
db	Darwinbox
az	Azure
vnet	Virtual network
vnetgw	Virtual network gateway
pip	Public Ip
aws	Amazon web services
customgw	Custom gateway
Vpgw	Virtual private gateway
lng	Local network gateway
S2s	Site to site
Internetgw	Internet gateway
vm	Virtual machine

CHAPTER 1 : INTRODUCTION

1.1 About Organization



A- ARTIFICIAL INTELLIGENCE
B- BIG DATA
C- CLOUD INNOVATION
D- DYNAMIC 365

Celebal Technologies is a premier software services company in the field of Data Science, Big Data, and Enterprise Cloud. We help you achieve a competitive advantage with intelligent data solutions, built using cutting-edge technology. Our core offerings are around Cloud Innovation, Supply Chain Analytics, Chat Bots, Power Platforms, and Data Analytics. Our solutions can help you accelerate decision-making and take giants leaps in your digital transformation journeys.



1.2 Purpose



Darwinbox is the leading provider of cloud-based Human Resources Management Software (HRMS). Their SaaS platform engages and empowers employees across the entire lifecycle (hire-to-retire) with a smarter, simpler & mobile-first HR Tech experience powered by Artificial Intelligence (AI) and Machine Learning (ML).

The objective is to build a modern data platform that can help Darwinbox to get better insights from all the data being generated to enable them to make informed decisions in charting out their growth story.

- The physical design will employ Microsoft and Celebal best practices for performance, networking, and security.
- We will build production environments and assist in promoting changes to each environment.
- Scripts will be created for all objects to ensure a stable environment for Production.

Setting up the Azure environment to connect with and move data from Redshift for multiple tenants.

1.3 Objectives

OBJECTIVE 1: Setting up the Azure environment to connect with and move data from Redshift for multiple tenants.

OBJECTIVE 2: Setting up the Power BI embedded server on the Azure VNet.

OBJECTIVE 3: Establishing site to site VPN connection between Microsoft azure and Aamazon Redshift cluster (AWS)

1.4 Solution

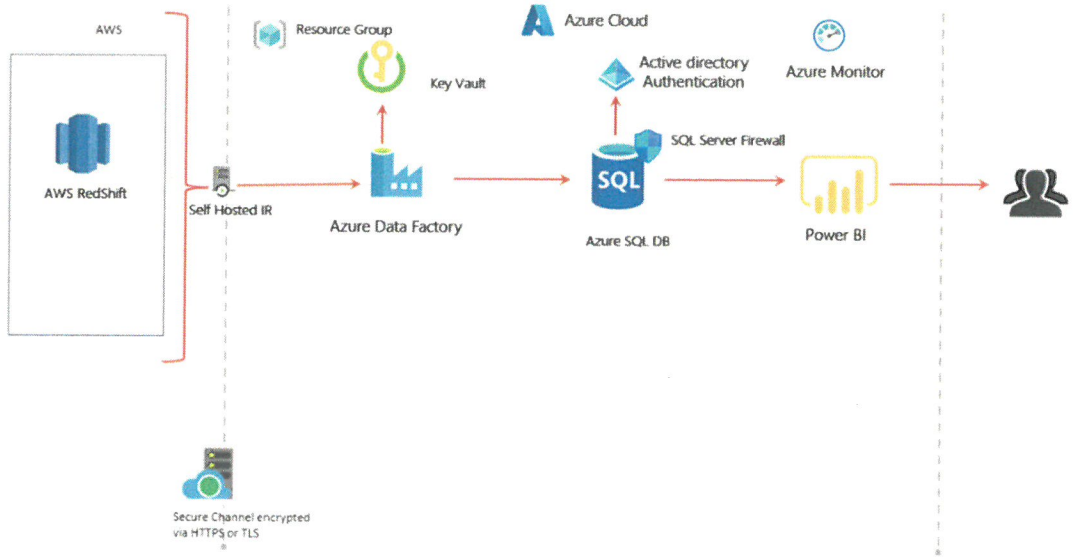


Figure 1.3: Environment Architecture

Naming convention are "AZ" denotes azure, "ED" denotes Subscription ID first two letter , CI denotes Region i.e. central India then Next denotes Service Name continuing with Number

- **Resource group (Az-ED-CI-RG-001)** - We have to create a single resource group with given naming convention. That holds all the related resources to the solution, in this case it will be holding ADF and Azure SQL database.

- **Azure Data Factory (AZ-ED-CI-ADF-001)** - Azure Data Factory is Azure's cloud ETL service for scale-out serverless data integration and data transformation. Both the linked services be it SHIR or Azure SQL DB services will be created in ADF.

- **Self-Hosted Integration Runtime (SHIR) Linked Service (Az-ED-CI-SHIR-LS-001)** - SHIR serves as a secure gateway between internal and external networks, it passes encrypted credentials to access internal resources and encrypts outbound data. In this case on successful registration the connection between on-prem and ADF will be established.

- **Redshift Linked Service (AZ-ED-CI-RDS-LS-001)** – This allow us to establish a connection between the Amazon redshift to Azure data factory, make sure to whitelist VM on redshift for the successful connection.

- **Azure SQL DB Linked Service (AZ-ED-CI-SDB-LS-001)** – This serves as a gateway to connect to our SQL DB that we created on the Azure portal.

- **Key vault (Az-ED-CI-KV-001)** - Azure Key Vault is a cloud service for securely storing and accessing secrets. A secret is anything that you want to tightly control access to, such as API keys, passwords, certificates, or cryptographic keys. Key Vault service supports two types of containers: vaults and managed hardware security module(HSM) pools.

- **Azure SQL DB (AZ-ED-CI-SQL-DB-001)** – We have to create a SQL database ,Azure SQL Database is a fully managed database service, which means that Microsoft operates SQL Server for you, and ensures its availability and performance. SQL Database also includes innovative features to enhance your business continuity, such as built-in high availability.

CHAPTER 2 : TECHNOLOGY STACK

2.1 Microsoft Azure

At its core, Azure is a public cloud computing platform—with solutions including Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) that can be used for services such as **analytics, virtual computing, storage, networking, and much more**.[2]

2.2 Amazon Web Services (AWS)

Amazon Web Services, Inc. is a subsidiary of Amazon that provides on-demand cloud computing platforms and APIs to individuals, companies, and governments, on a metered pay-as-you-go basis. These cloud computing web services provide distributed computing processing capacity and software tools via AWS server farms.[3]

2.3 Amazon Redshift Cluster

An Amazon Redshift data warehouse is a collection of computing resources called nodes, which are organized into a group called a cluster. Each cluster runs an Amazon Redshift engine and contains one or more databases.[4]

2.4 Virtual Network Gateway (Azure)

A virtual network gateway is composed of two or more VMs that are automatically configured and deployed to a specific subnet you create called the *gateway subnet*. The gateway VMs contain routing tables and run specific gateway services. You can't directly configure the VMs that are part of the virtual network gateway, although the settings that you select when configuring your gateway impact the gateway VMs that are created.[5]

2.5 Customer Gateway (AWS)

A customer gateway device is a physical or software appliance that you own or manage in your on-premises network (on your side of a Site-to-Site VPN connection). You or your network administrator must configure the device to work with the Site-to-Site VPN connection.[6]

2.6 Virtual Private Gateway (AWS)

A virtual private gateway is a logical, fully redundant distributed edge routing function that sits at the edge of your VPC. As it is capable of terminating VPN connections from your on-prem or customer environments, the VPG is the VPN concentrator on the Amazon side of the Site-to-Site VPN connection.[7]

2.7 Site To Site VPN (Azure/AWS)

A site-to-site setup is where two (or more) different networks are connected together using one OpenVPN tunnel. In this connection model, devices in one network can reach devices in the other network, and vice versa. The implementation of this is, for as far as Access Server is involved in this, relatively simple.[8]

2.8 Local Network Gateway (Azure)

The local network gateway is a specific object that represents your on-premises location (the site) for routing purposes. You give the site a name by which Azure can refer to it, then specify the IP address of the on-premises VPN device to which you'll create a connection.[9]

2.9 Internet Gateway(AWS)

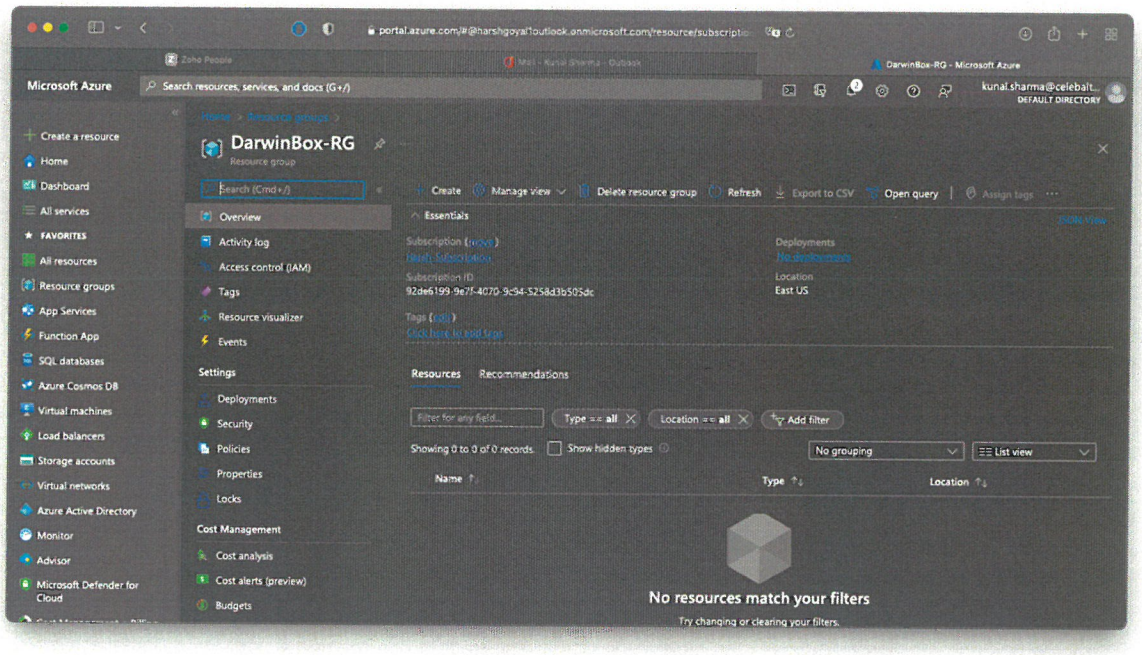
An internet gateway is a horizontally scaled, redundant, and highly available VPC component that allows communication between your VPC and the internet. An internet gateway enables resources (like EC2 instances) in your public subnets to connect to the internet if the resource has a public IPv4 address or an IPv6 address. Similarly, resources on the internet can initiate a connection to resources in your subnet using the public IPv4 address or IPv6 address. For example, an internet gateway enables you to connect to an EC2 instance in AWS using your local computer.[10]

2.10 Route Tables (AWS)

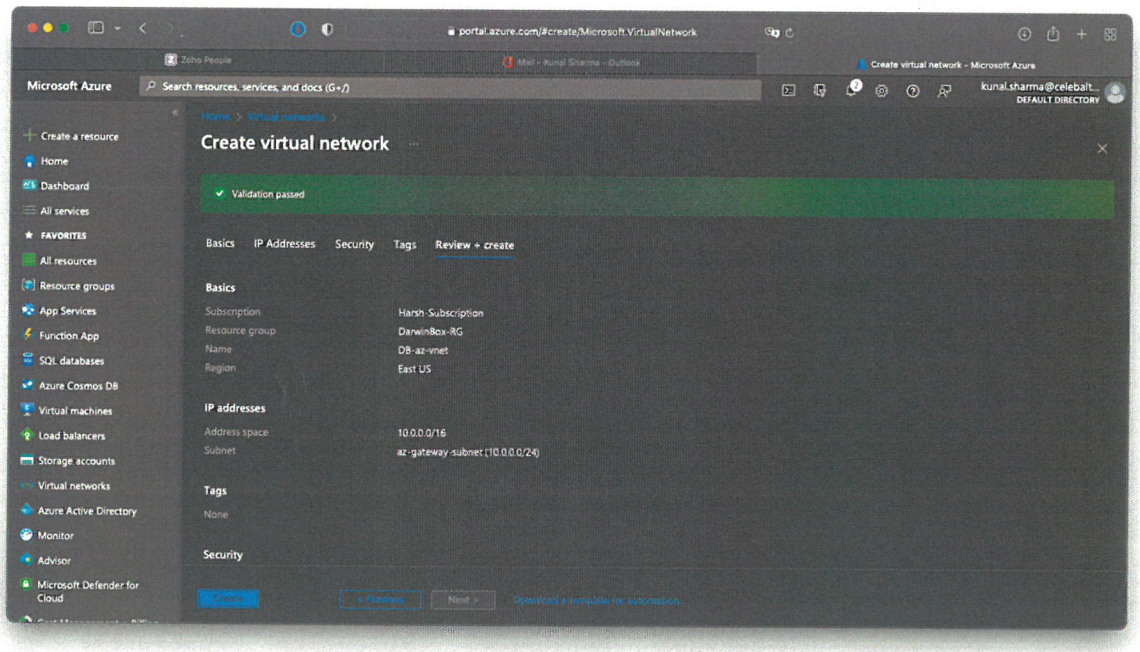
A route table contains a set of rules, called routes, that determine where network traffic from your subnet or gateway is directed.[11]

CHAPTER 3 : FLOW OF WORK

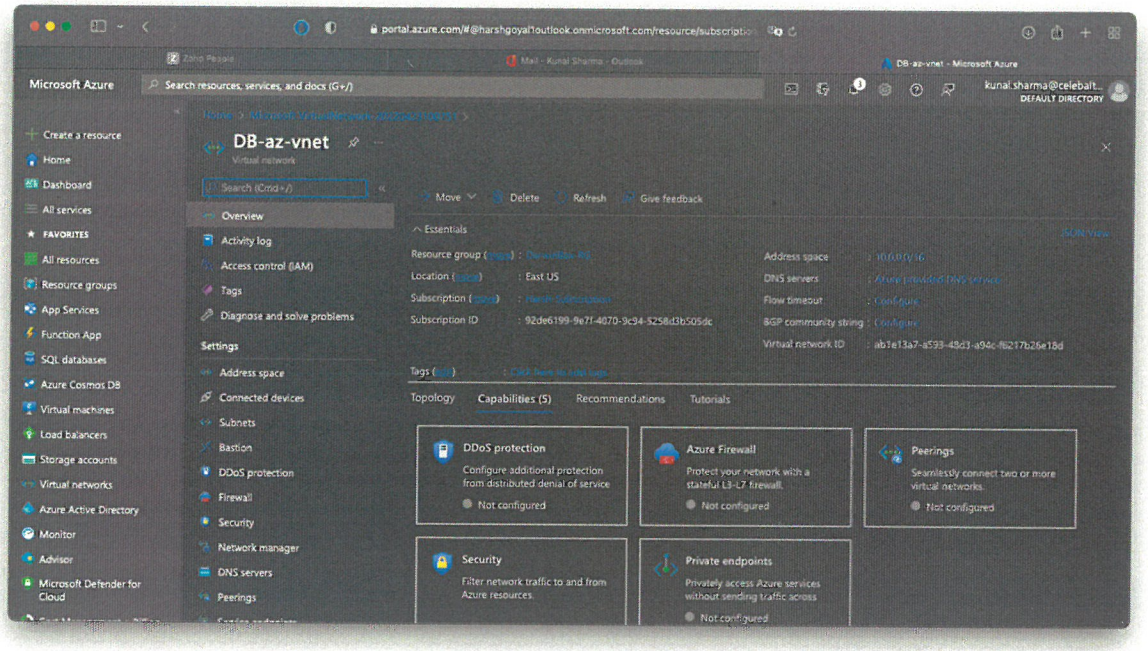
1. Create a resource group on Azure to deploy the resources on that.



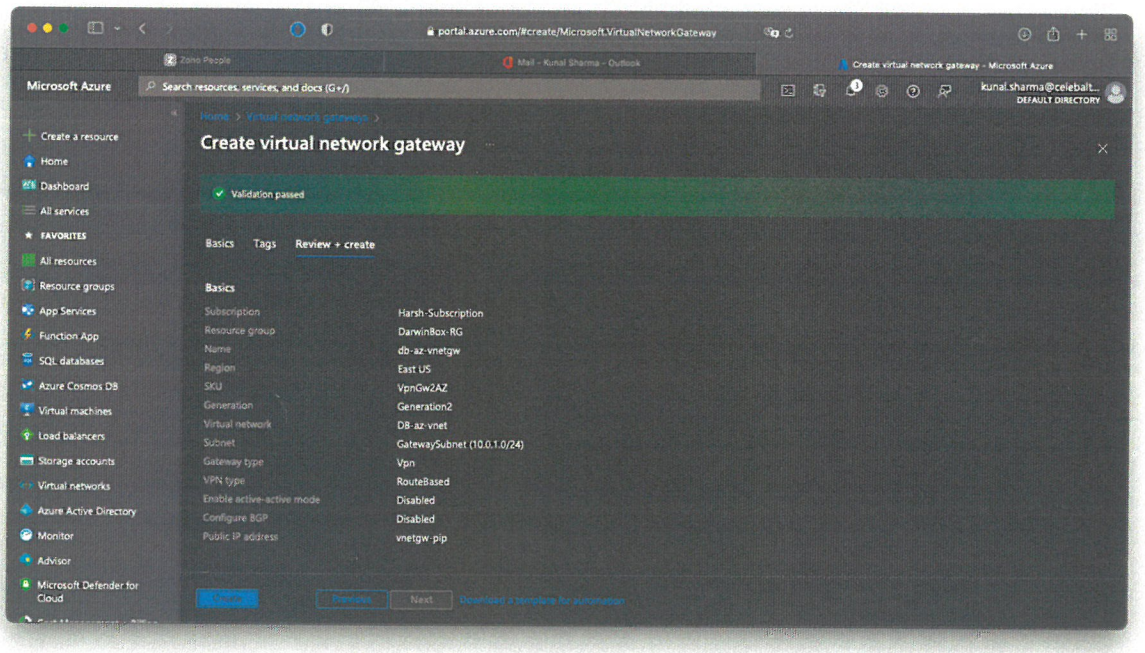
2. Create a virtual network with name DB-az-vnet address space 10.0.0.0/16 and change name of default subnet to az-gateway-subnet.



3. Virtual network created.

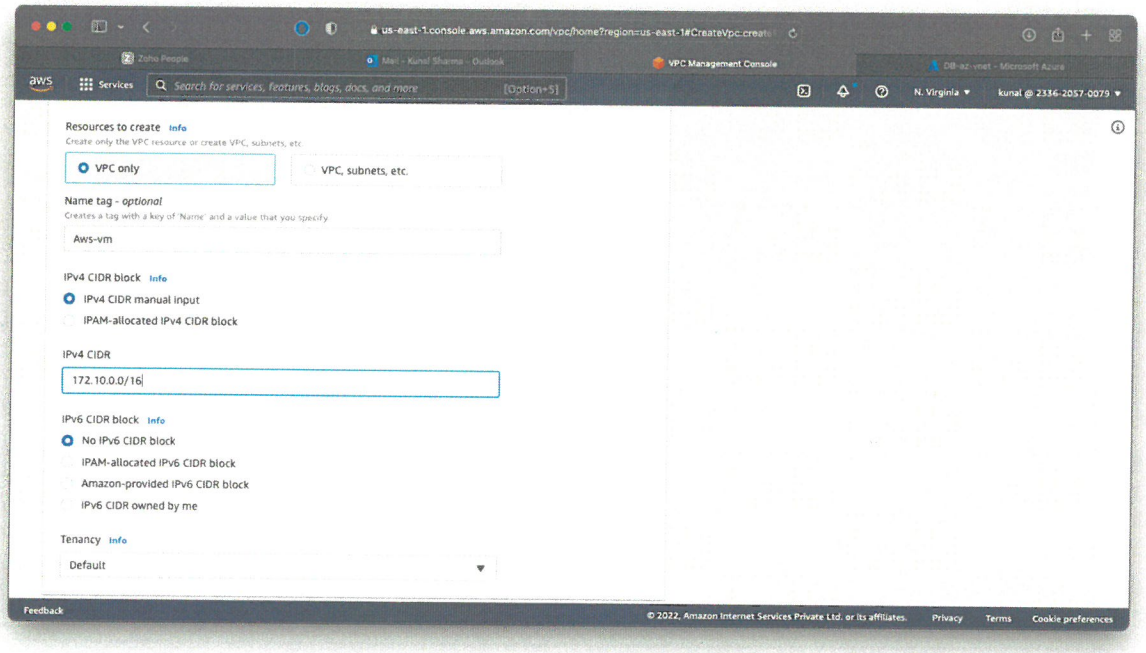


4. Create a virtual network gateway (db-az-vnetgw) using same vnet and subnet created , followed by a public IP address (vnetgw-pip).

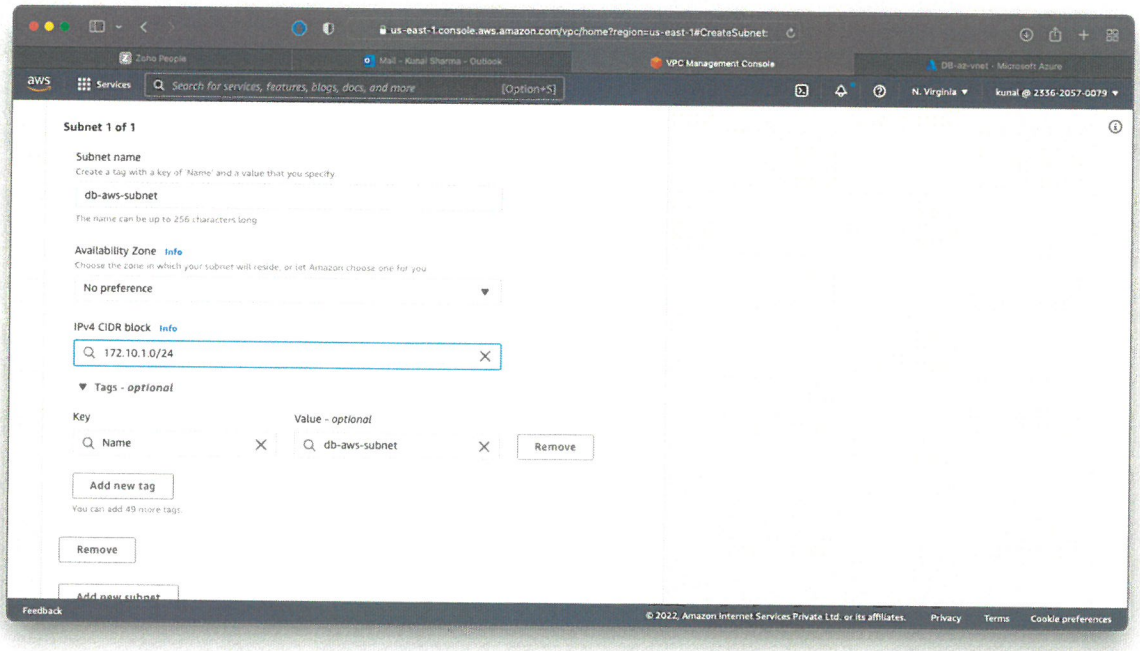


Configuring AWS

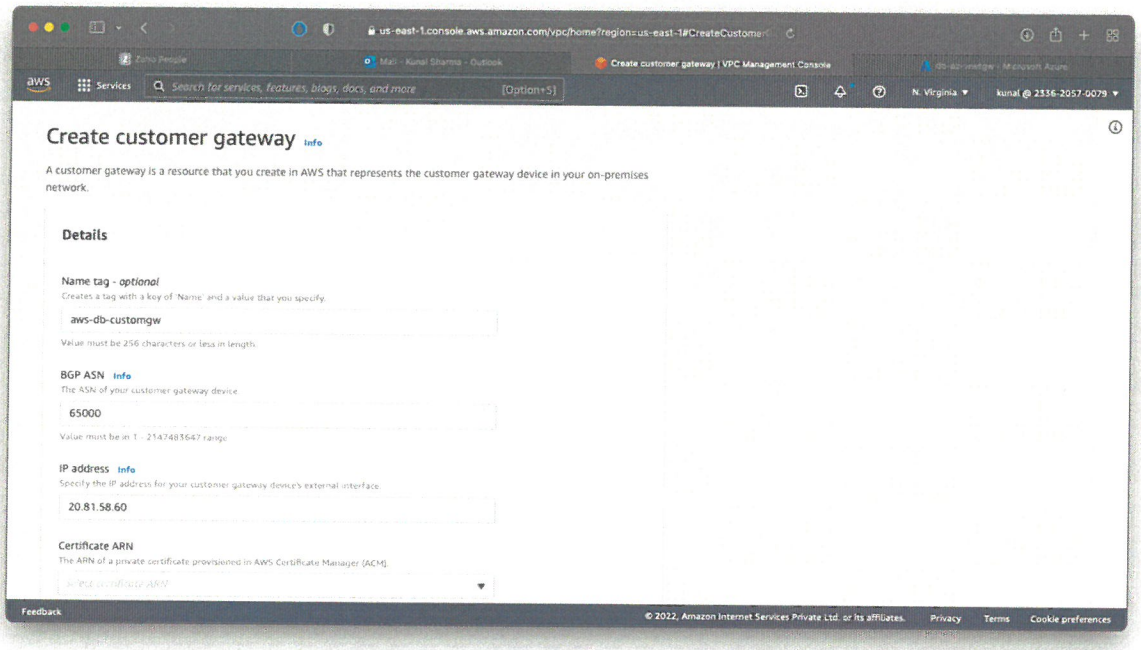
5. Create a Virtual private cloud (VPC) names Aws-vm and use IPv4 CIDR as 172.10.0.0/16 and click on create keeping everything default.



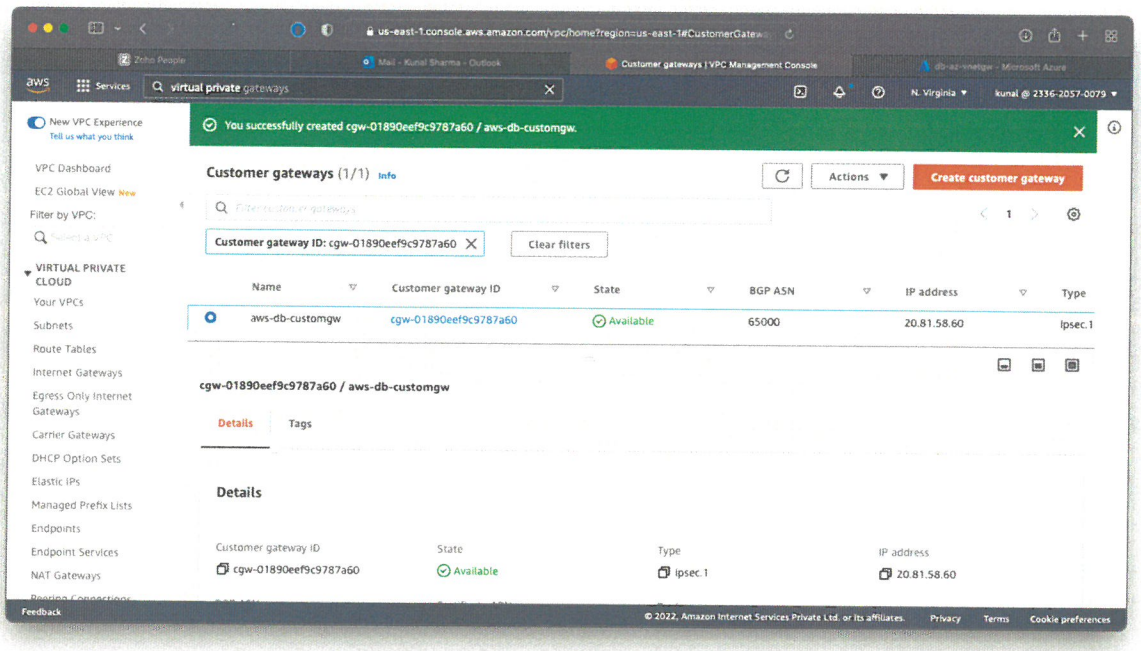
6. Create a subnet (db-aws-subnet) with IPv4 CIDR block as 172.10.1.0/24 and click on create.



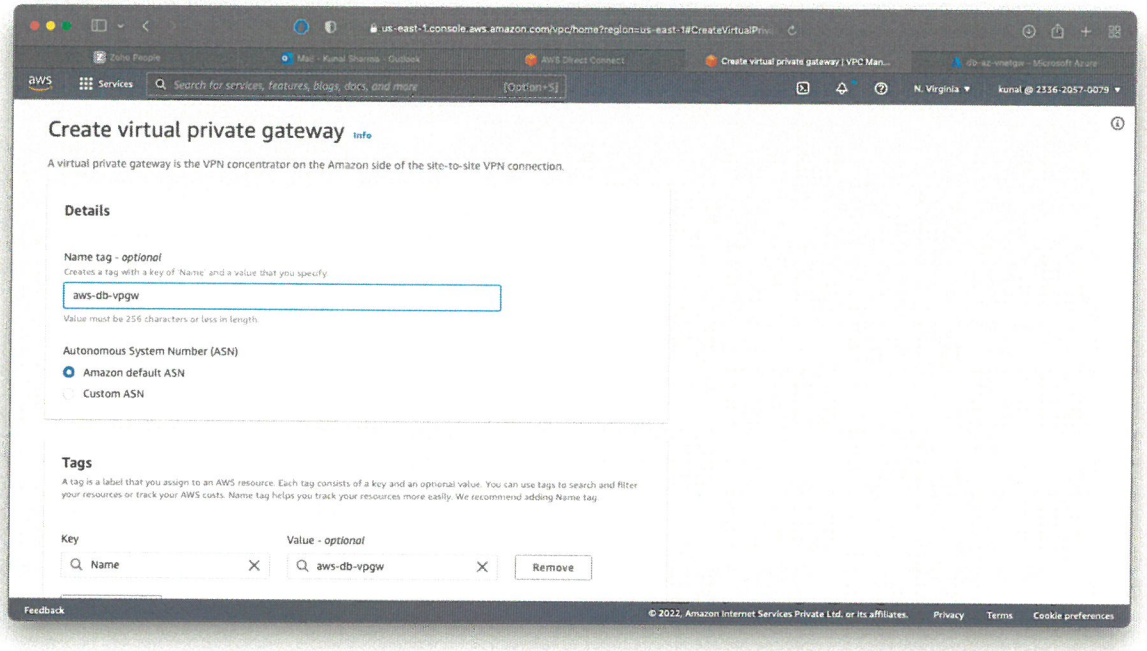
7. Create a customer gateway (aws-db-customgw) pointing to the public IP address of Azure VPN Gateway (20.81.58.60)
   The Customer Gateway is an AWS resource with information to AWS about the customer gateway device, which in this case is the Azure VPN Gateway.



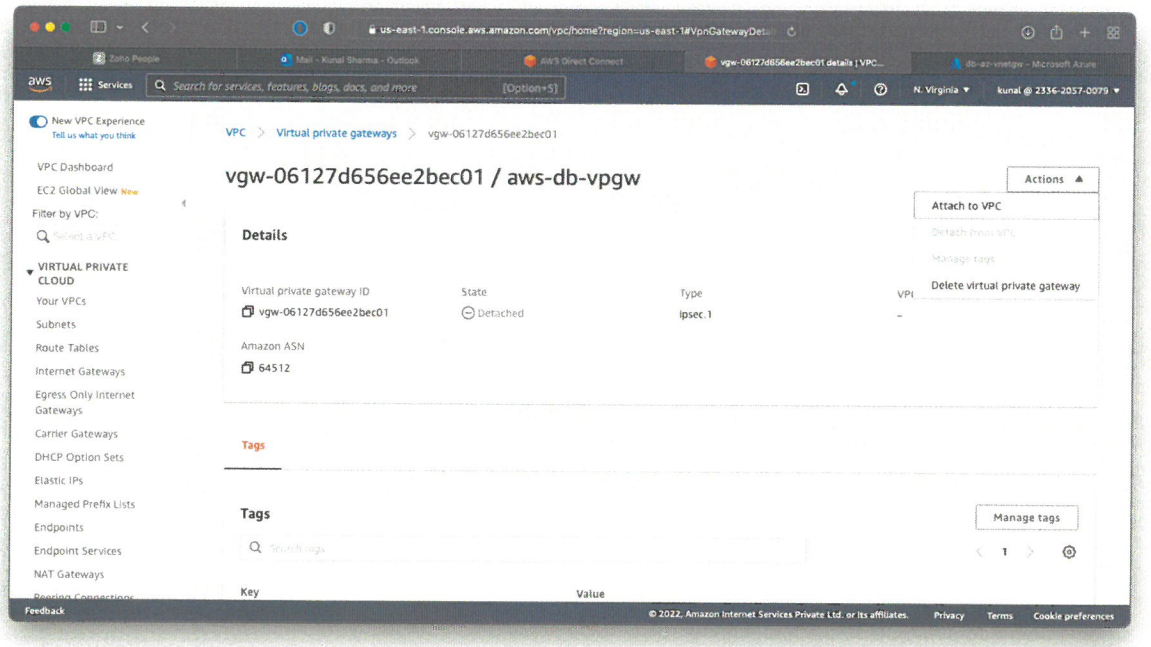
8. Customer gateway created and is in available state.



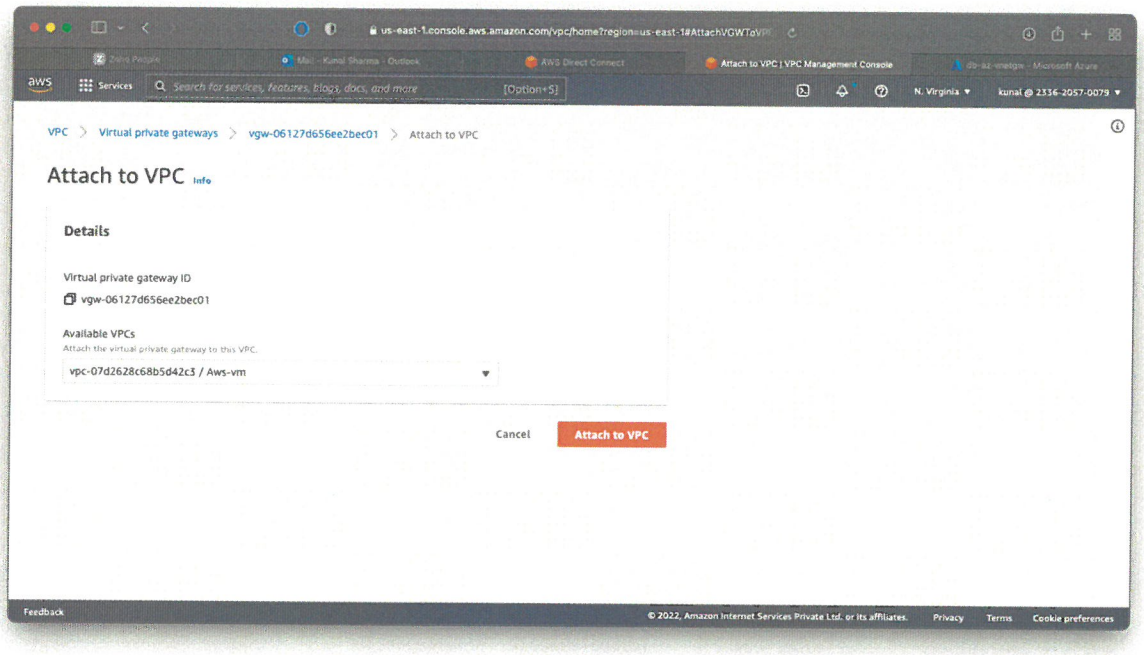
9. Create the Virtual Private Gateway (aws-db-vpgw) then attach to the VPC.



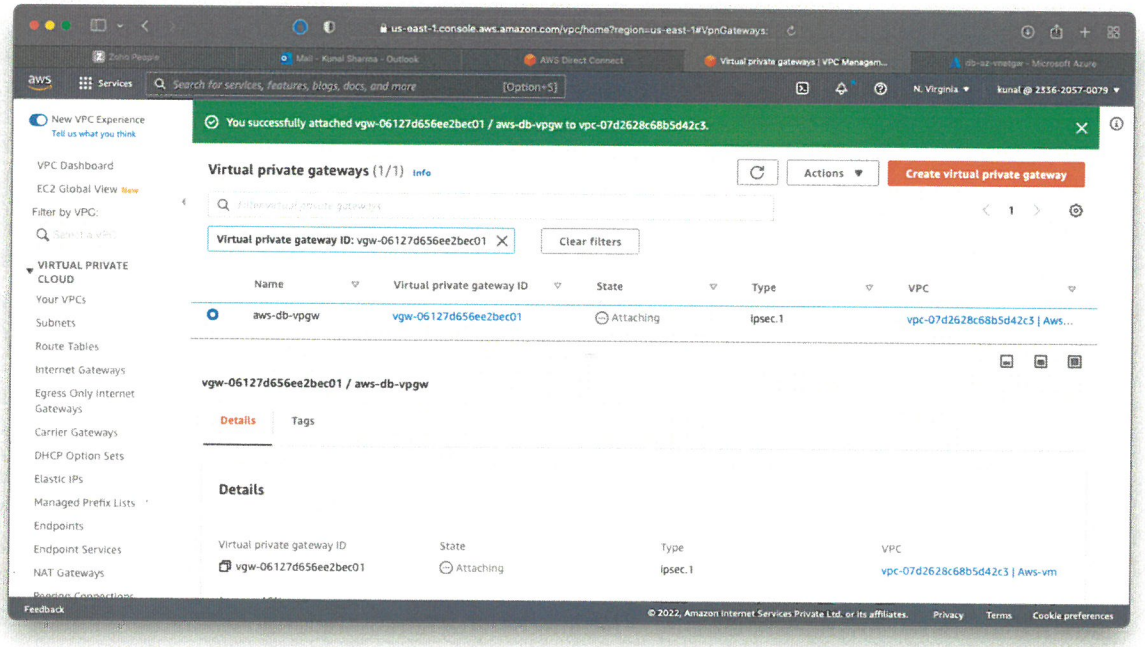
10. Go to the resource and click on action to choose attach to VPC option.



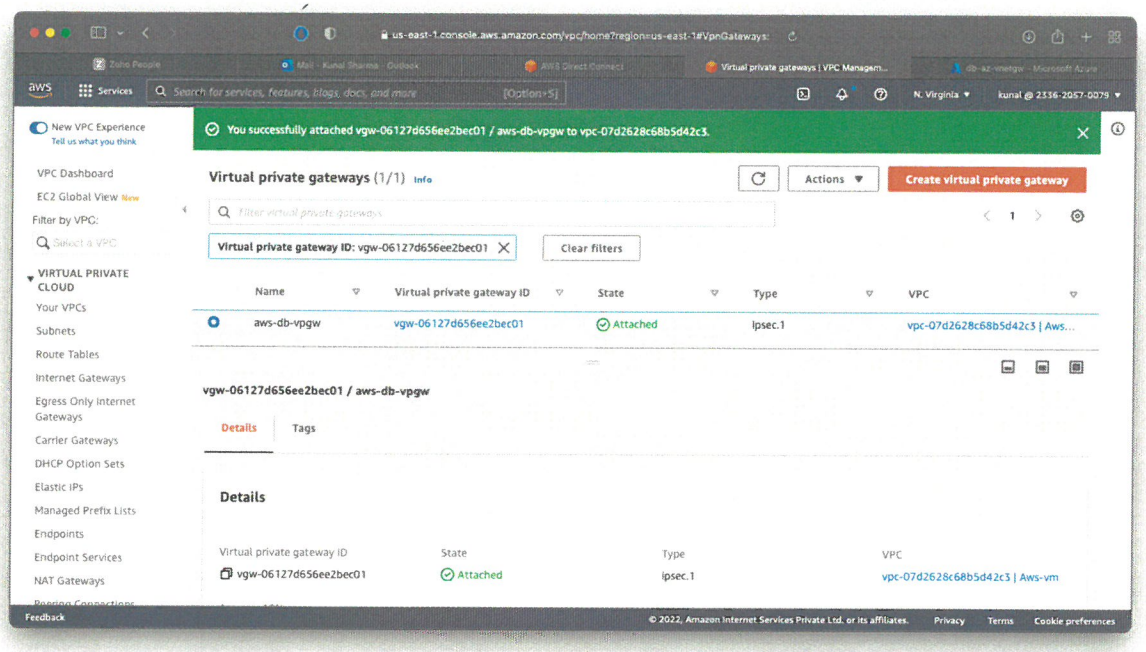
11. Select the VPC and click on attach to VPC.



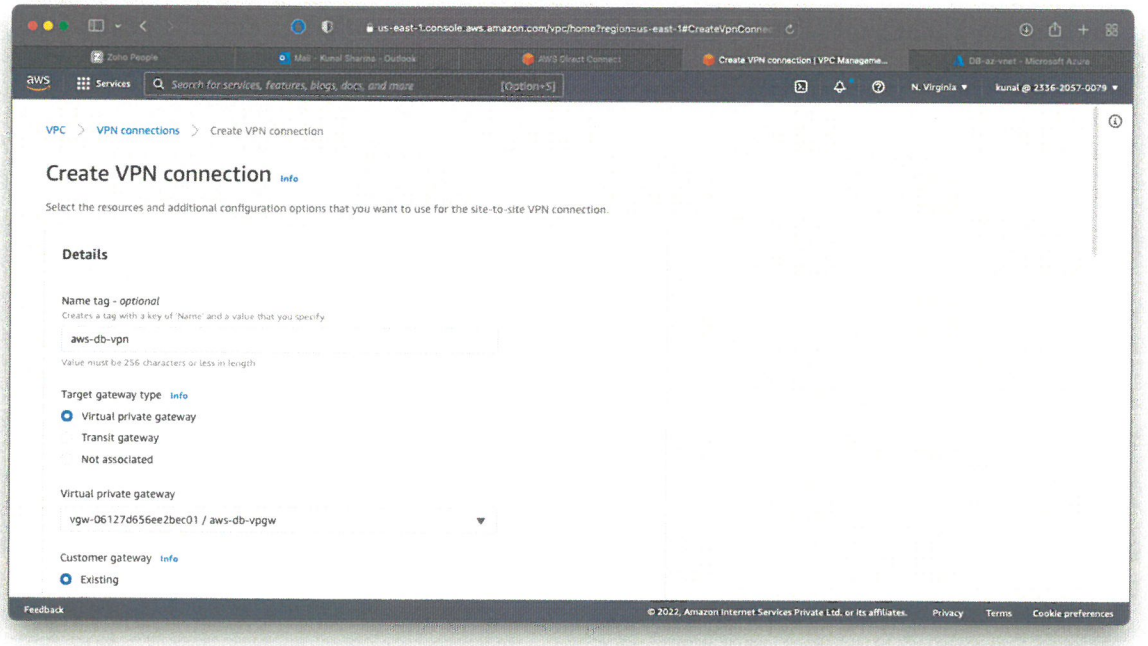
12. State of virtual private gateway is attaching.



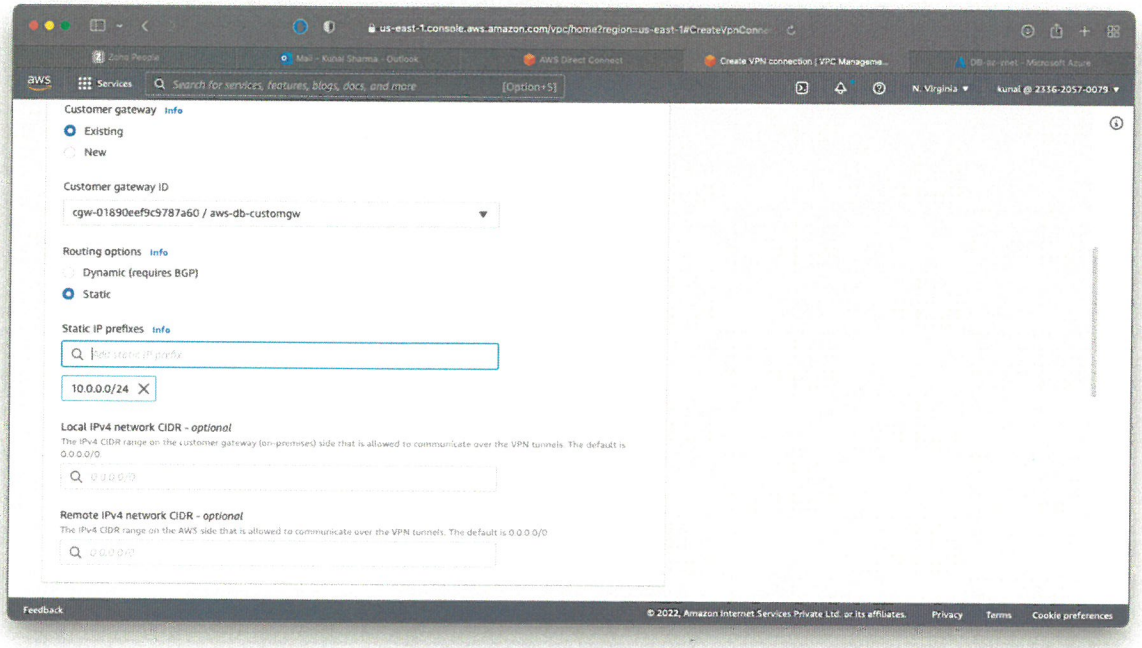
13. State changed from attaching to attached.



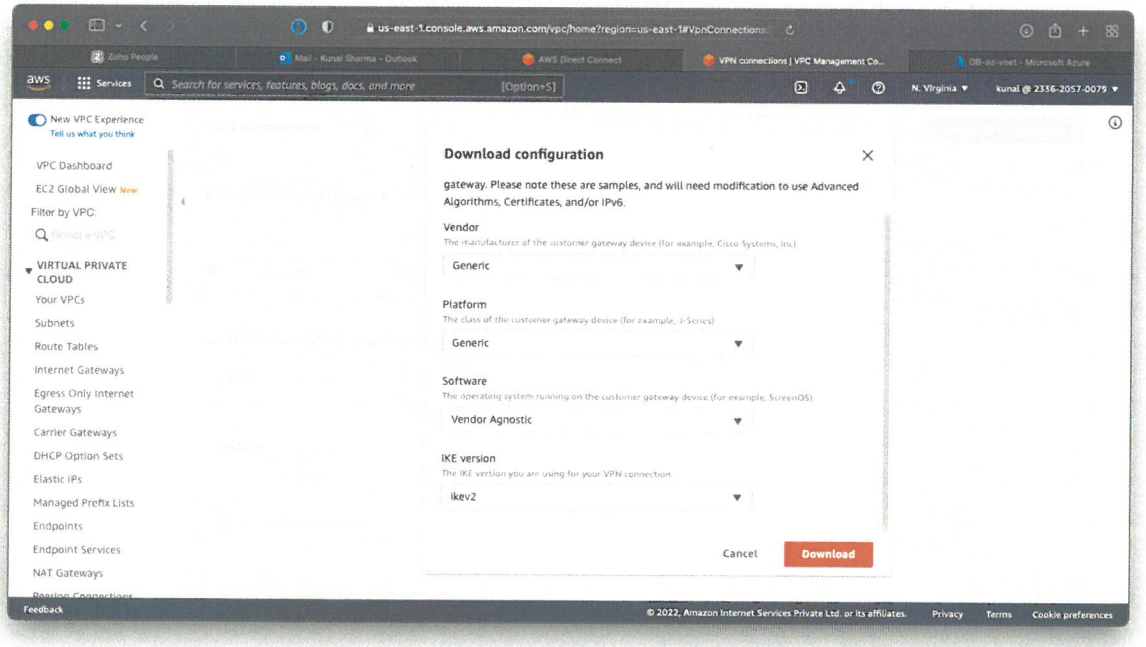
14. Create a site-to-site VPN Connection (aws-db-vpn) and use the vpg created by us previously.



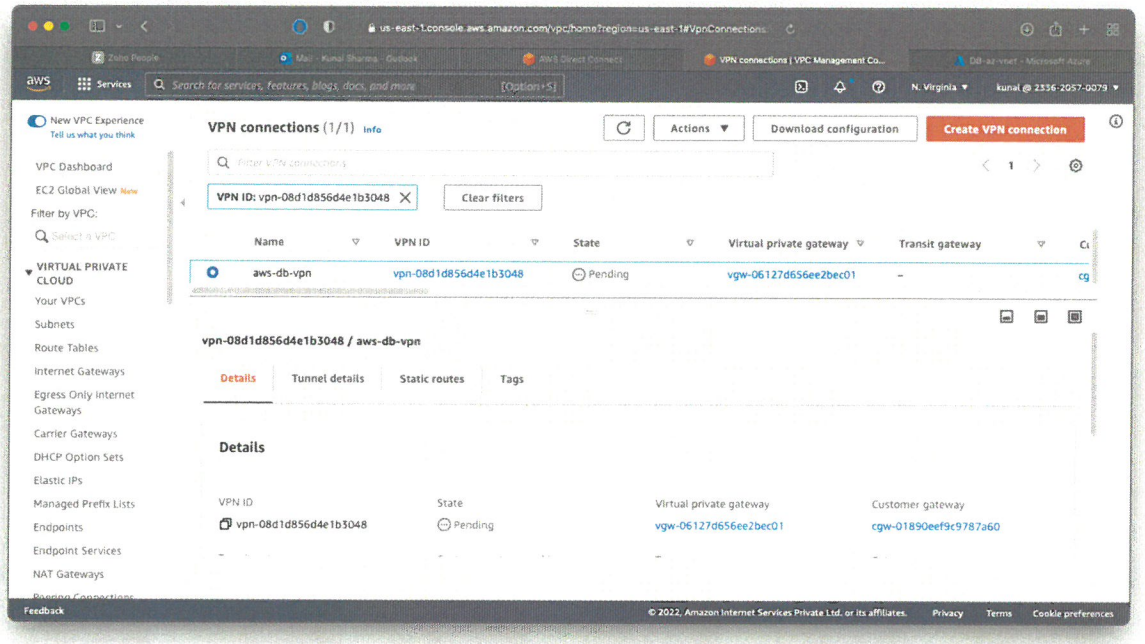
15. Switch routing options to static from dynamic and use IP of subnet (10.0.0.0/24) as static IP prefixes.



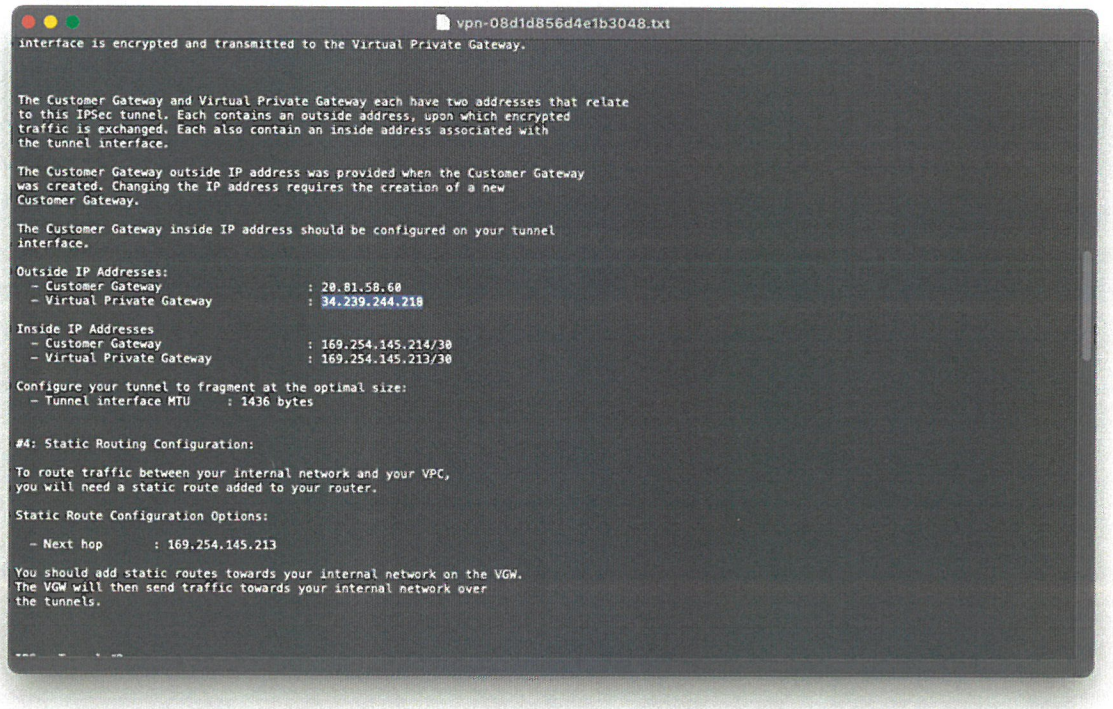
16. Click on download configuration and select vendor as generic and keeping everything static, switch the IKE version from ikev1 to ikev2 and download the file.



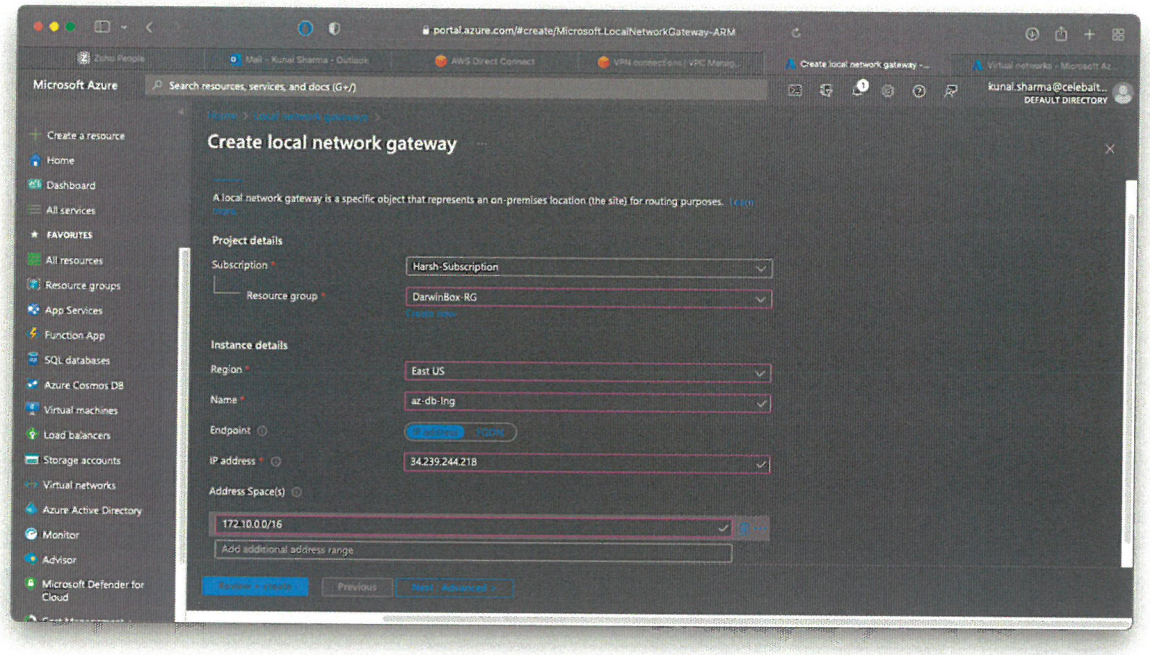
17. VPN connection is established and the state is pending.



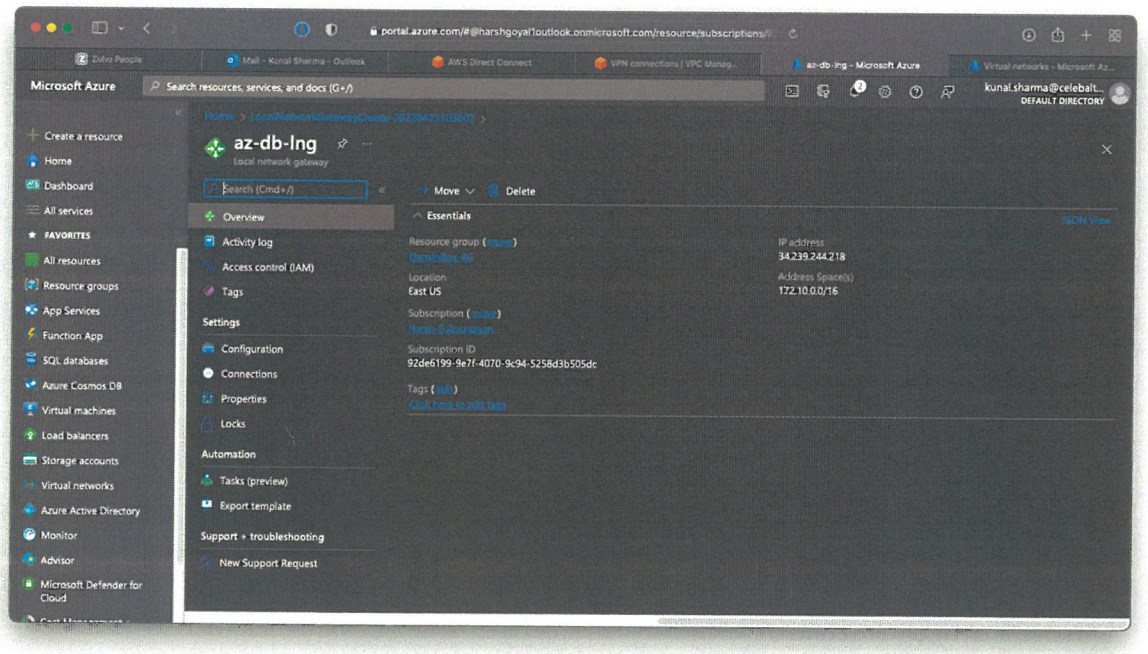
18. Open the file we downloaded and copy the virtual private gateway (34.239.244.218) of outside IP Addresses, we will be using this as the address space while creating local network gateway in next step.



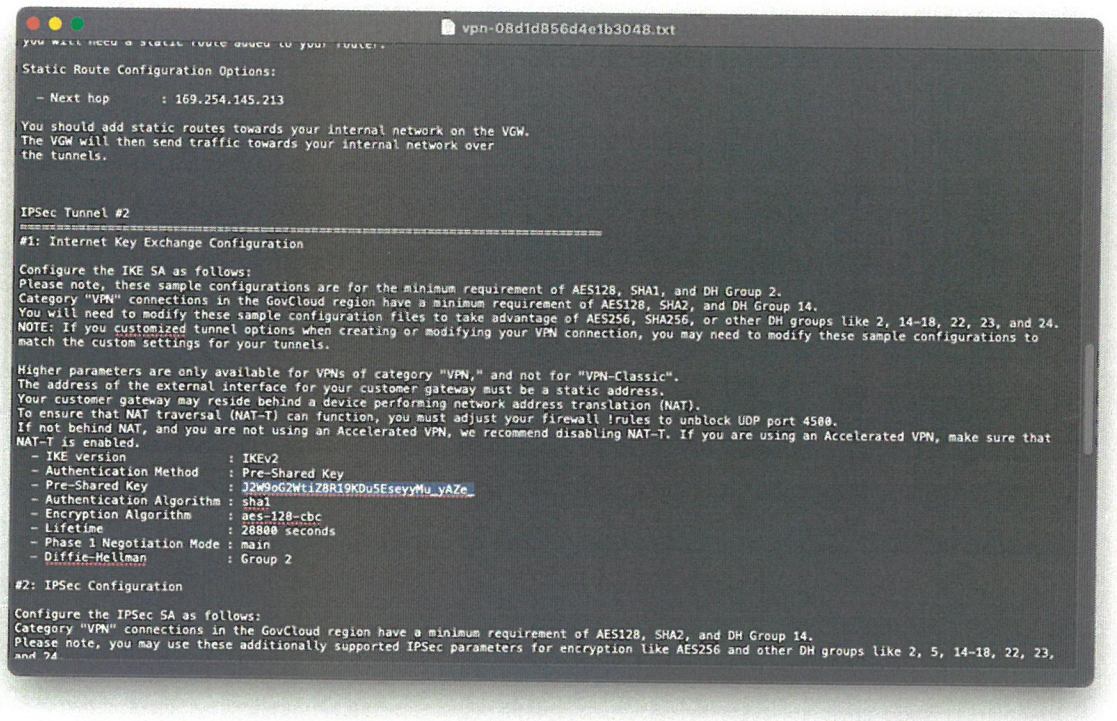
19. Switch back to azure and create a local network gateway (az-db-lng) and use the address we copied in last step in the addresses space of the lng.



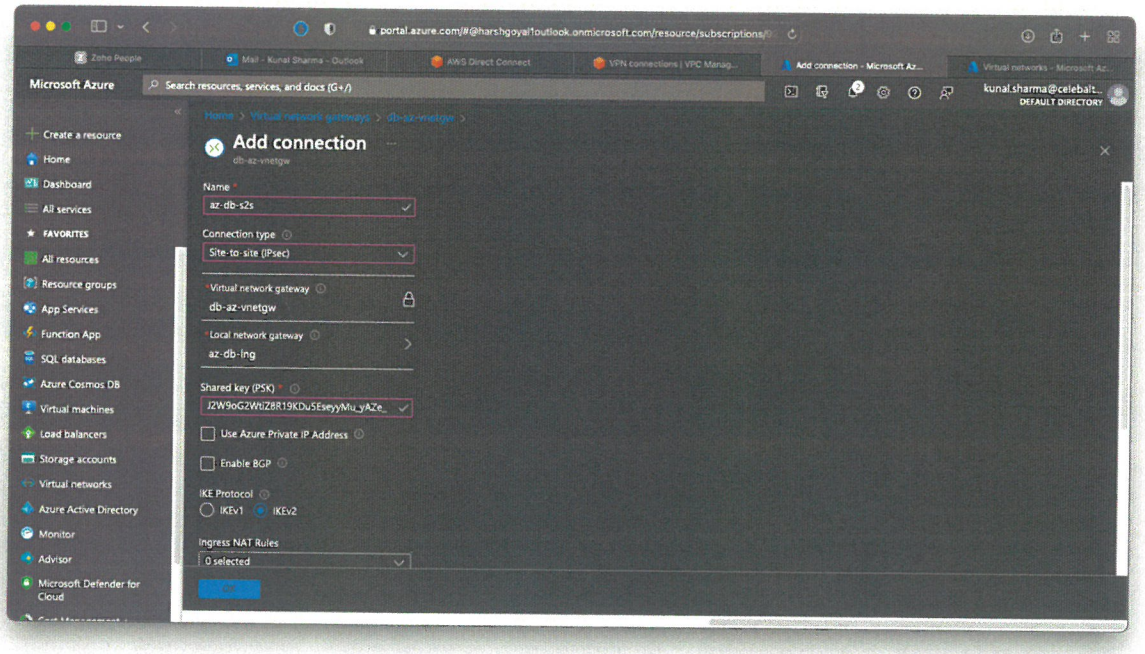
20. Local network gateway is created, now click on connections from left panel to create a s2s connection.



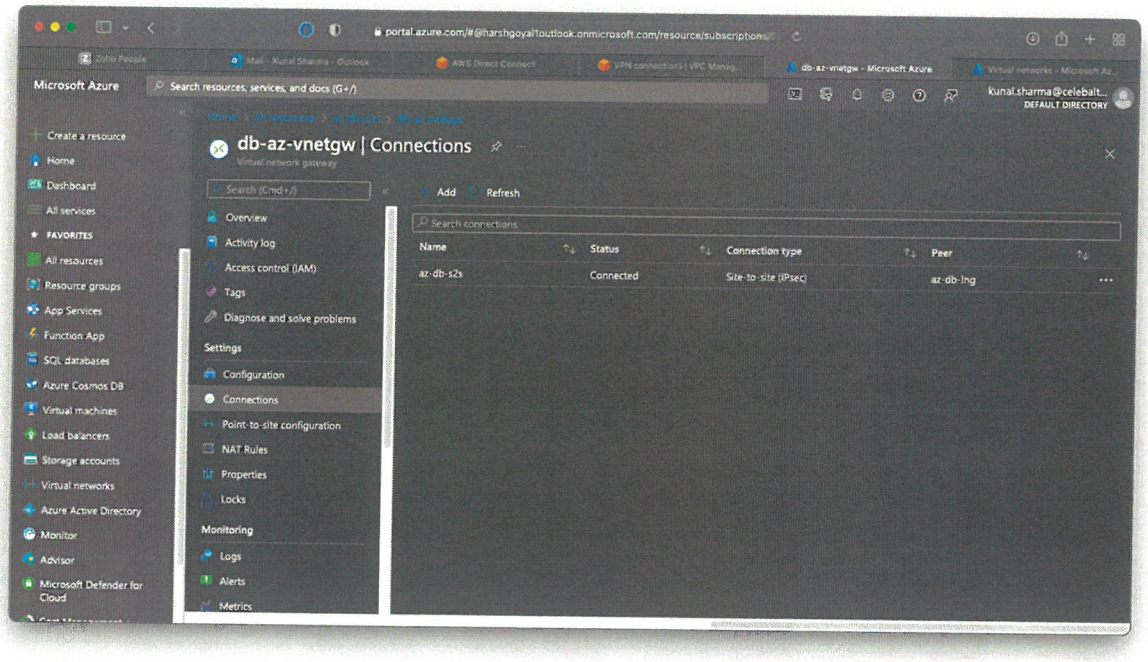
21. Now copy the pre shared key from the tunnel-1 (in image it is showing from tunnel-2 which is wrong, scroll up and you will find tunnel-1 as well), we will be using this while creating the s2s connection.



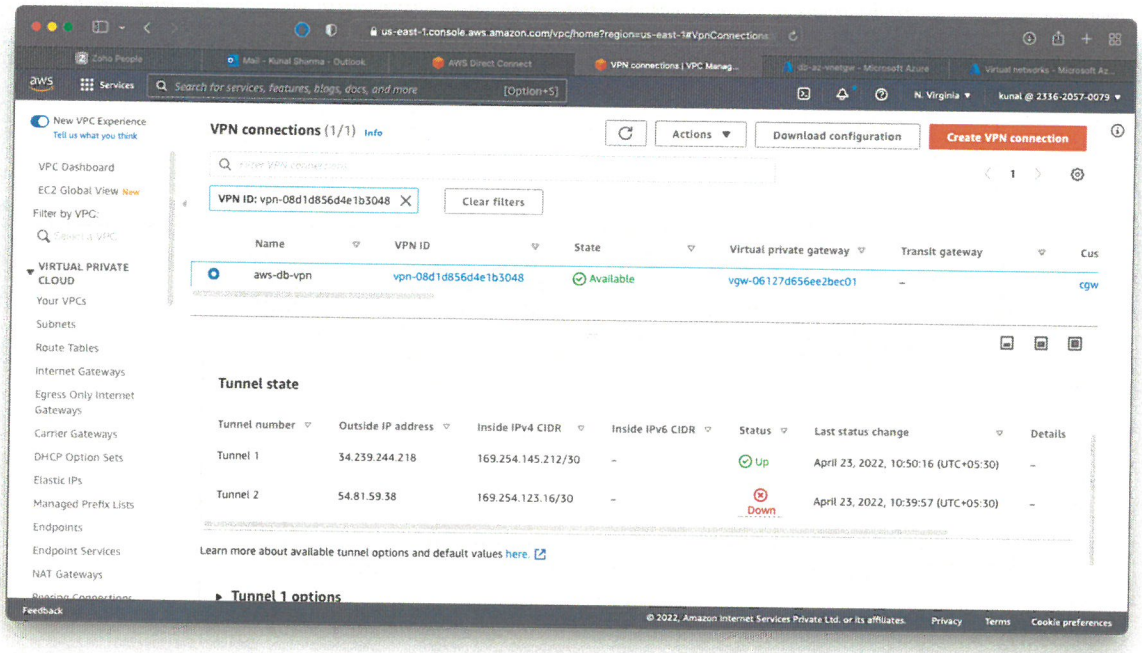
22. Apply the name (az-db-s2s) and switch connection type to site-to-site, use the local network gateway created and apply the key we copied in last step and click on ok.



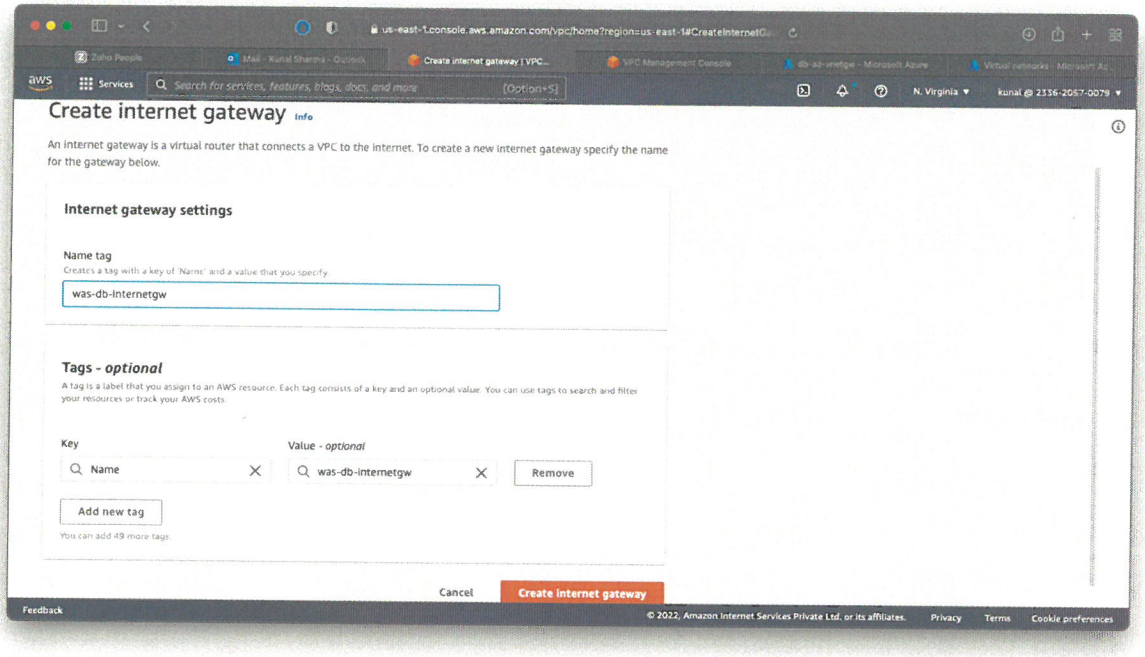
23. Site to site connection has been established and the status is connected.



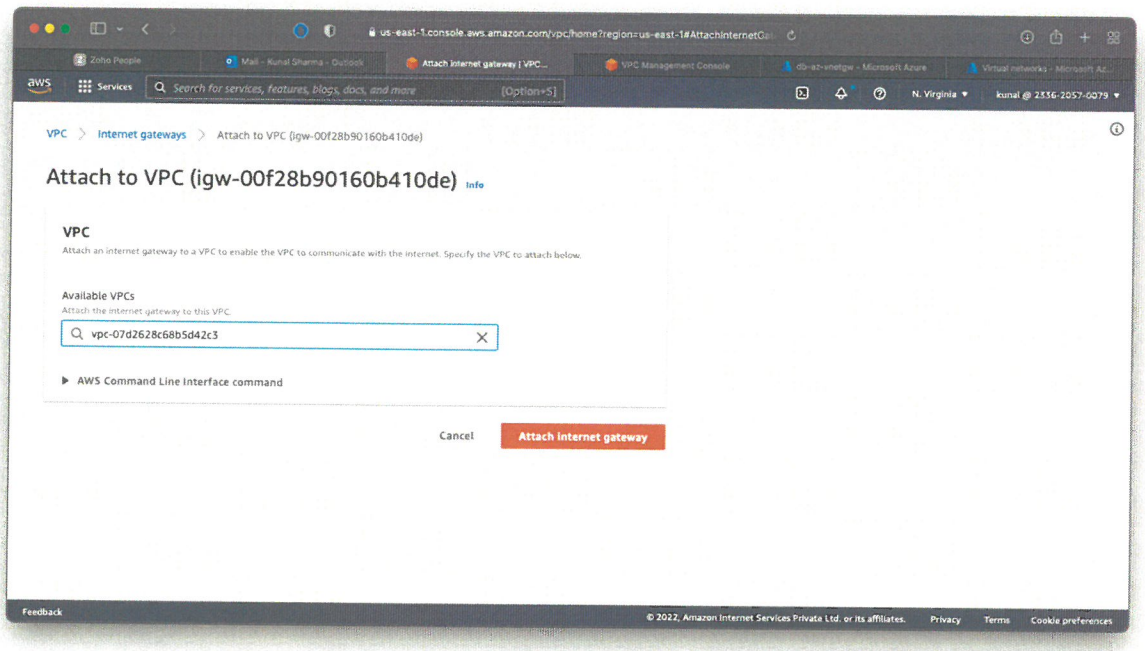
24. We can cross check the same from AWS portal where we can see that the status of tunnel-1 is up.



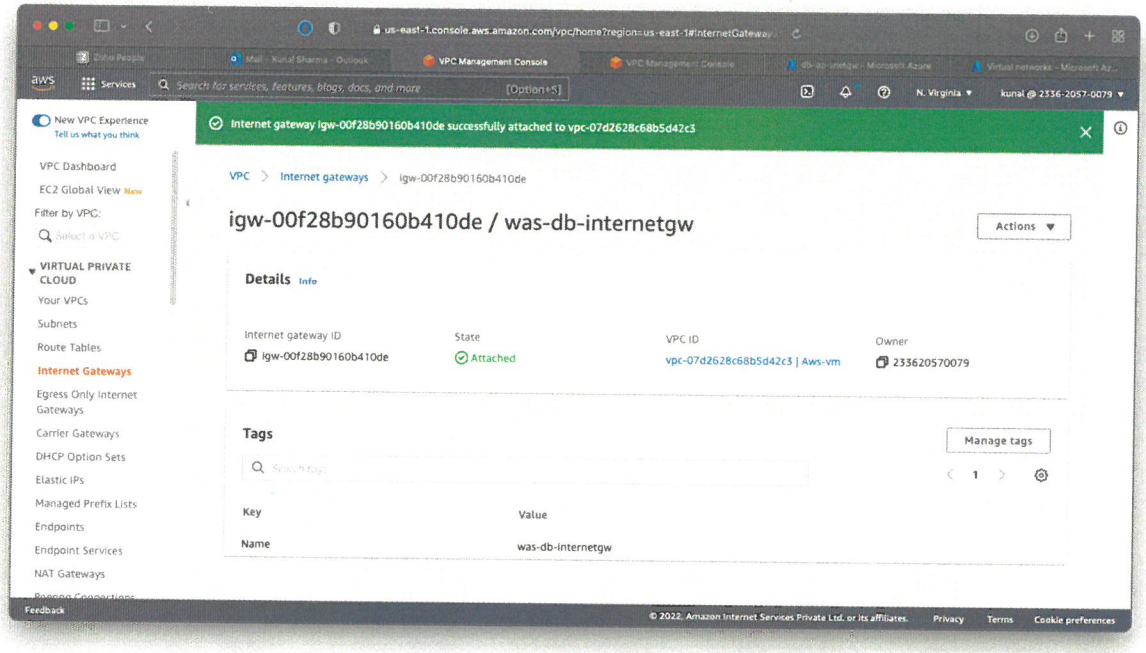
25. Create an internet gateway (aws-db-internetgw) and click on create gateway.



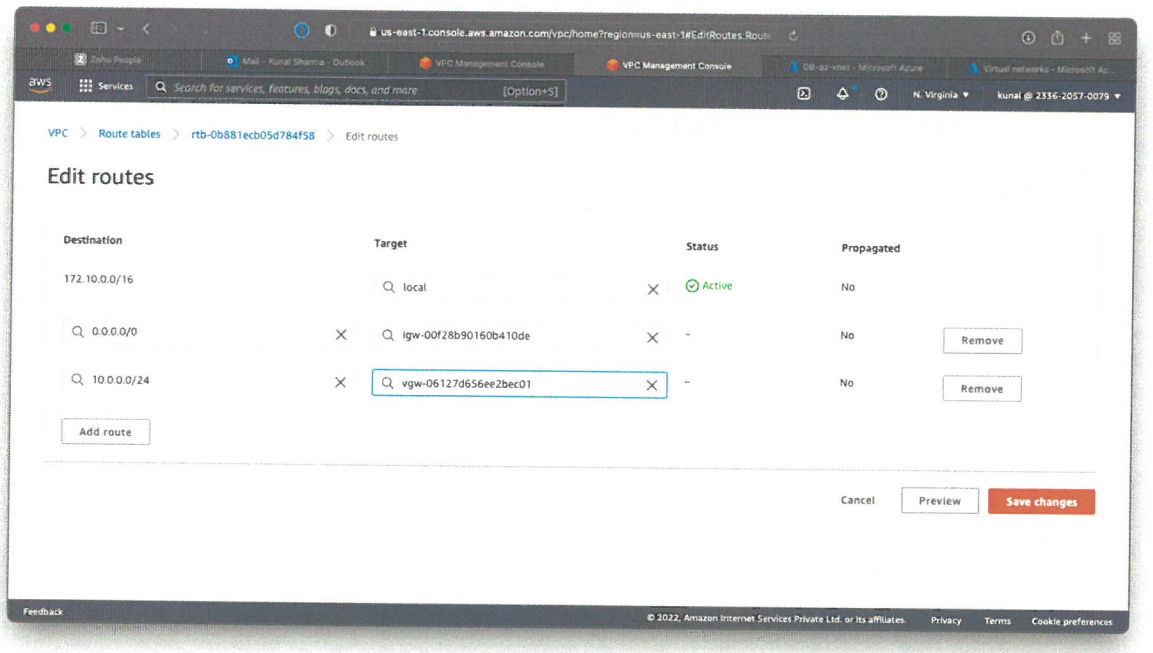
26. Attach the internet gateway to the vpc created by us.



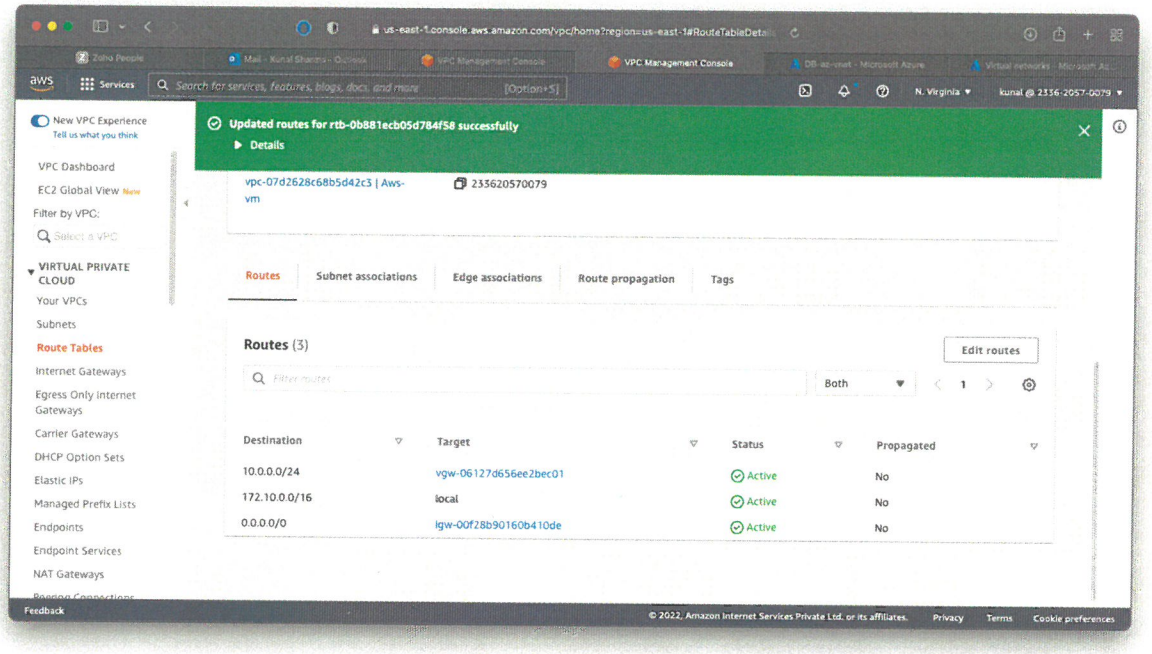
27. Internet gateway is created.



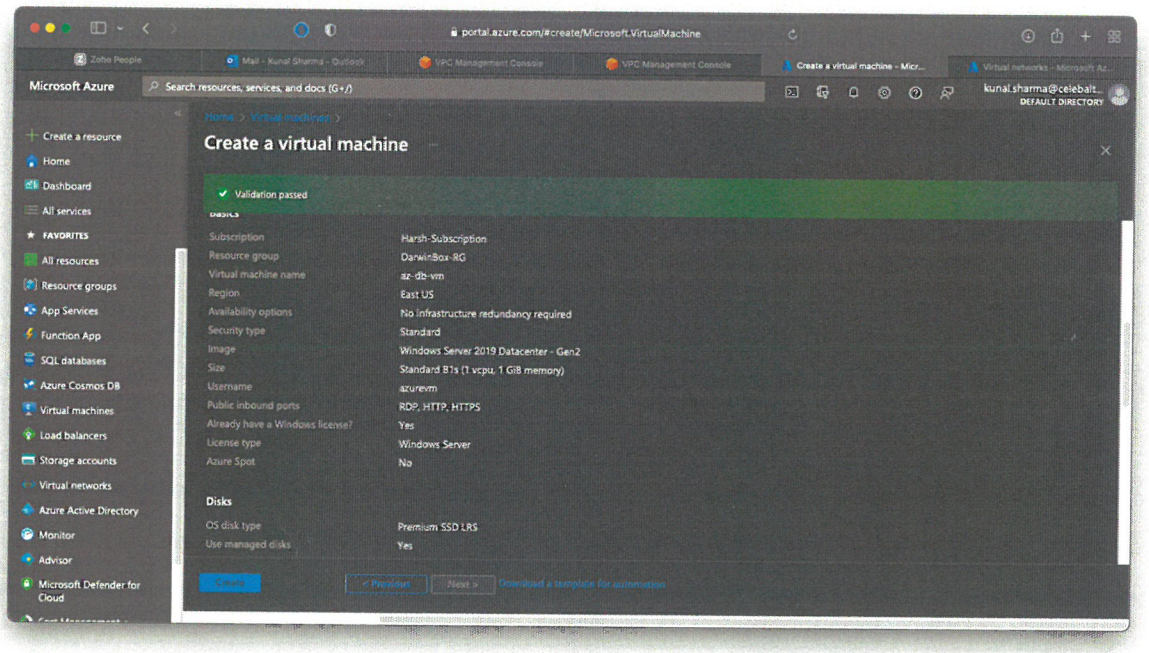
28. Click on edit route and use 0.0.0/0 as destination and select target as the internet gateway created by us, similarly use the IP of subnet created on azure (gateway-subnet) and use the target as Virtual gateway.



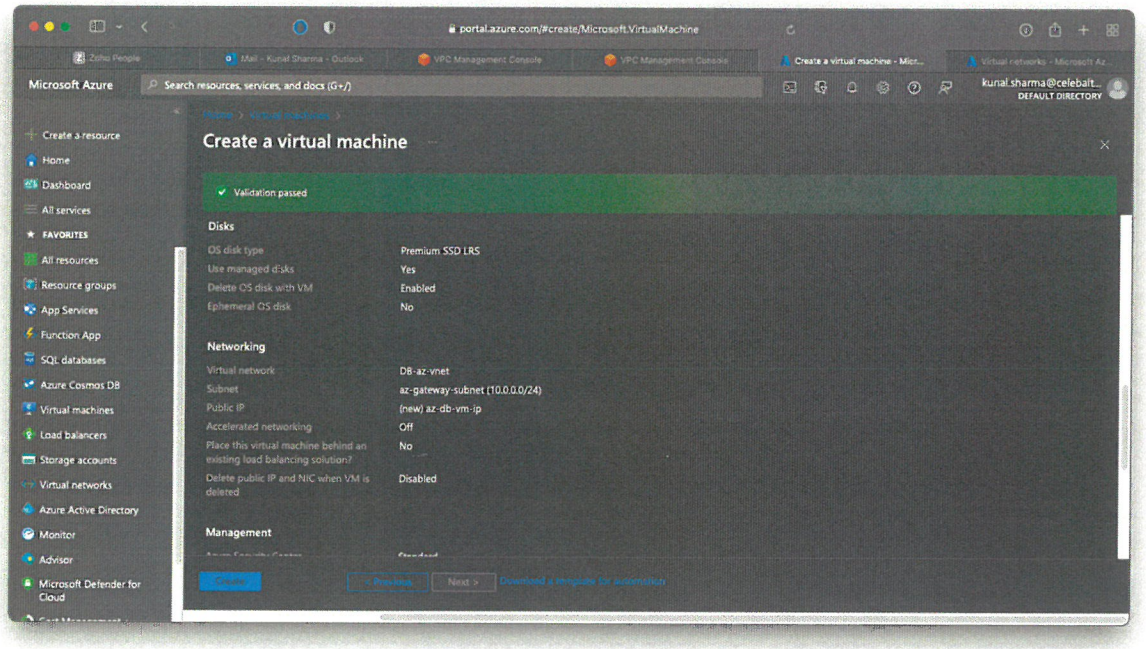
29. Route tables updated. And status of all the routes are active.



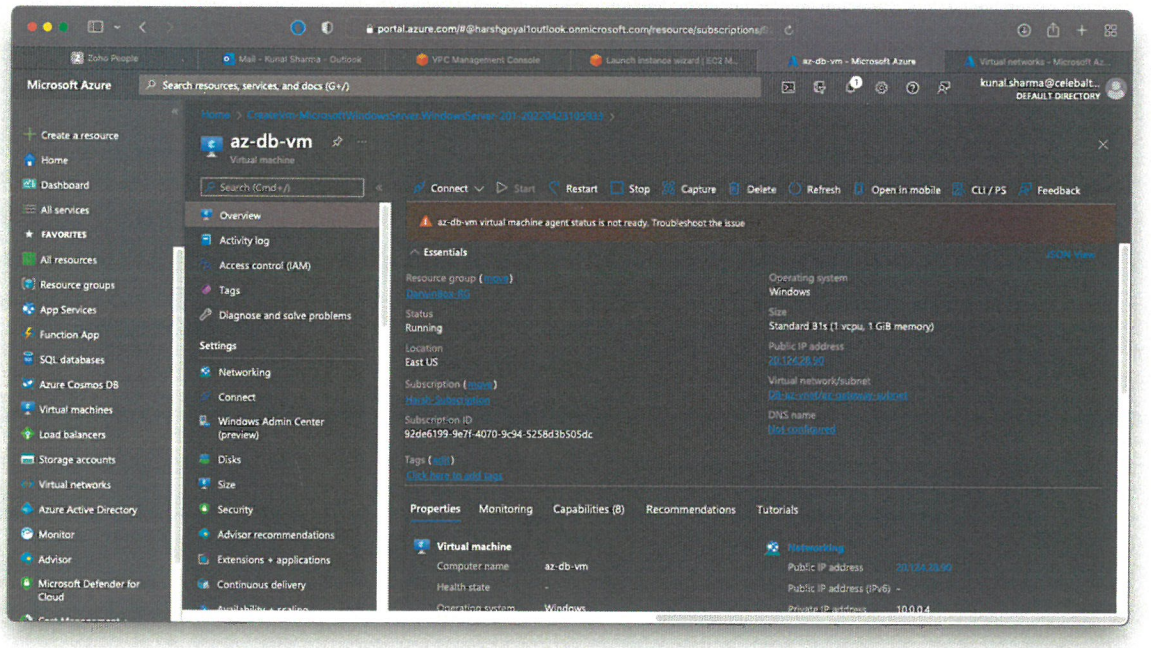
30. Create a virtual machine on azure (az-db-vm) using the image and size given in the figure and using public inbounds (RDP,HTTP,HTTPS).



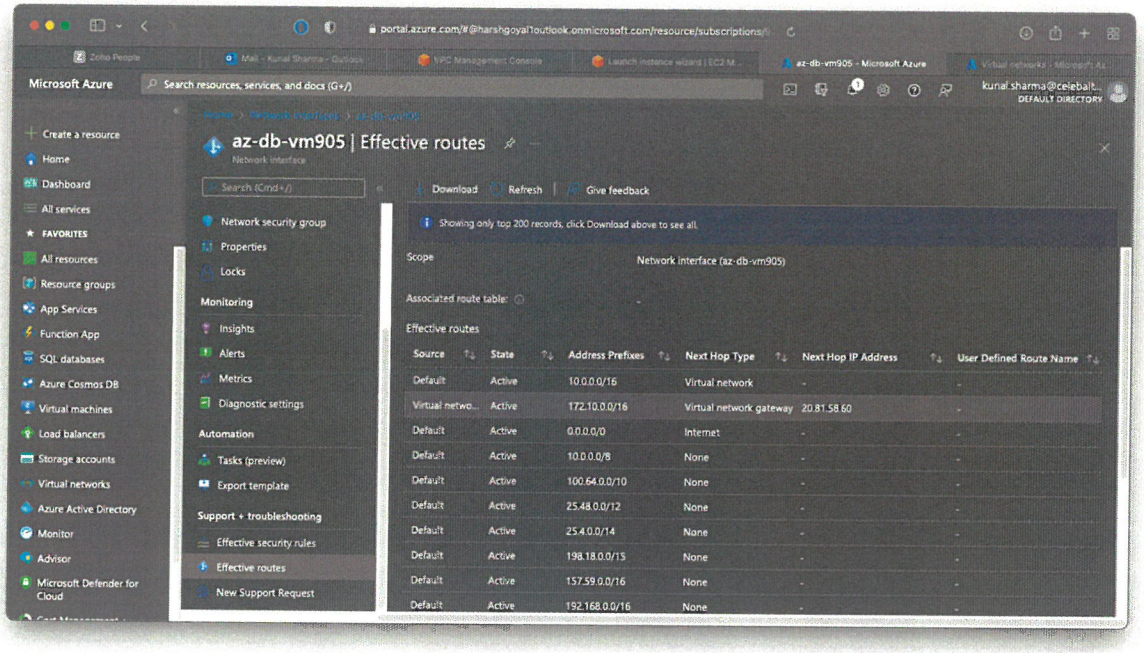
31. Use the same vnet and subnet we have already created and create a new IP (az-db-vm-IP) and make it static. Click on create.



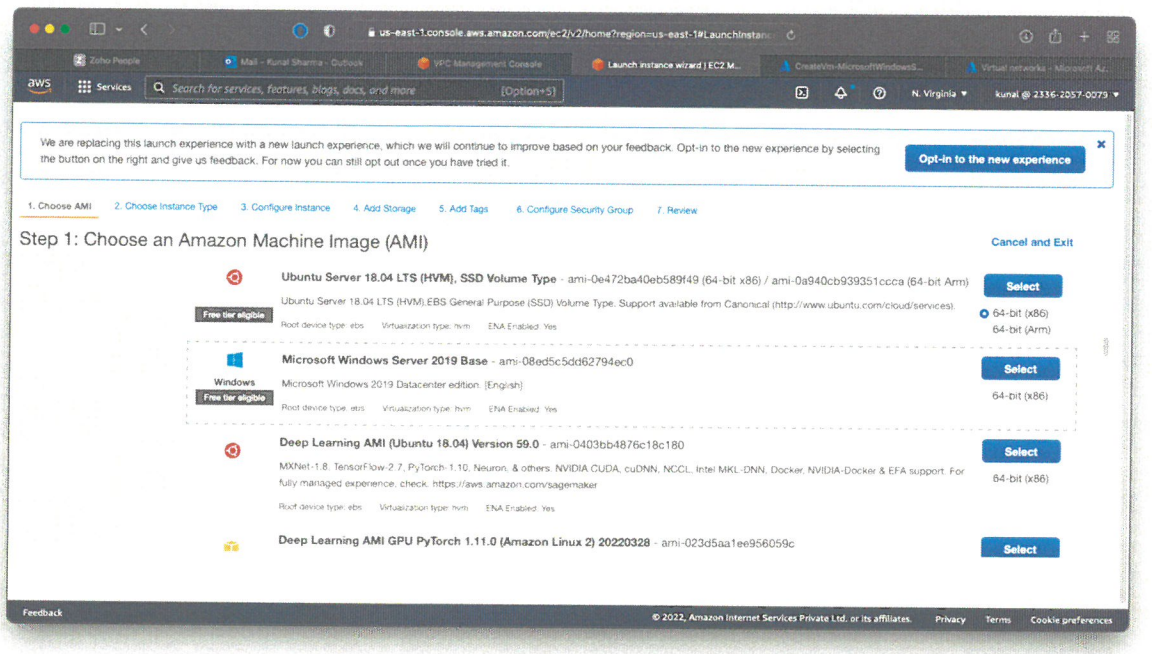
32. Azure vm created.



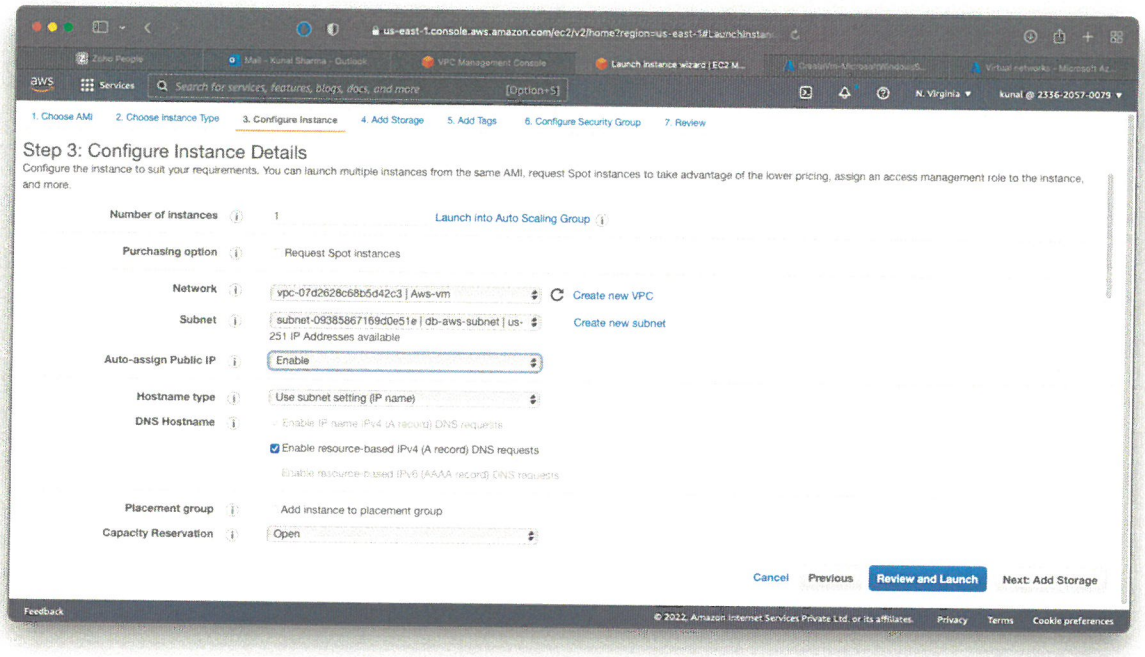
33. On Azure the route was automatically created. We can check selecting the Azure VM > Networking > Network Interface > Effective routes.



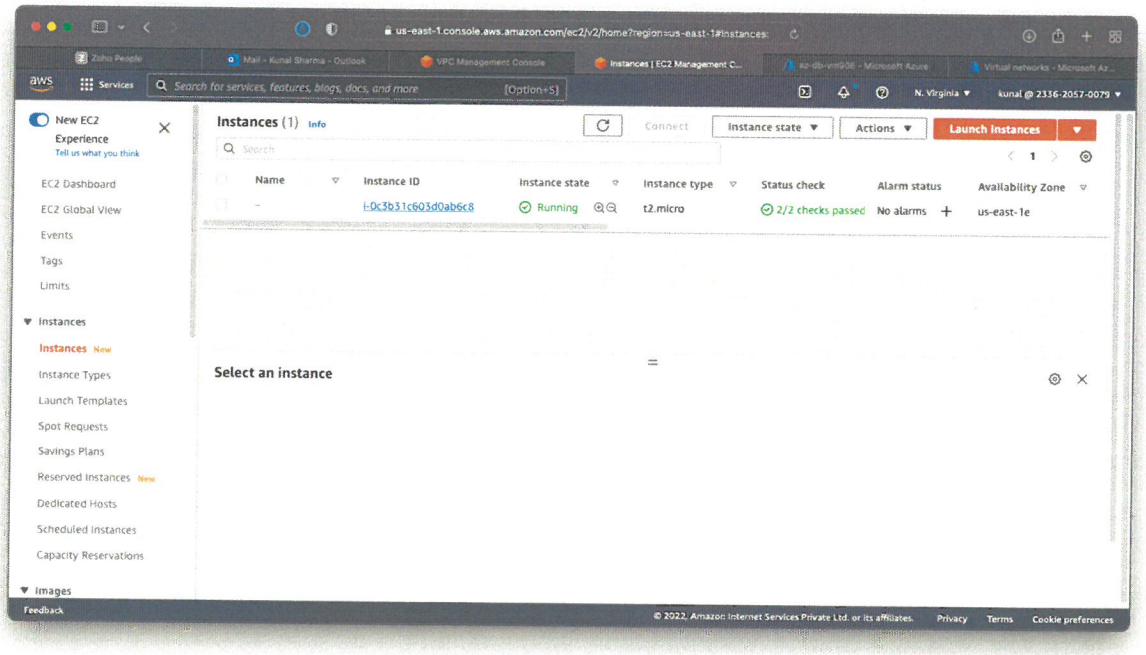
34. We will now be creating an EC2 instance, select windows image.



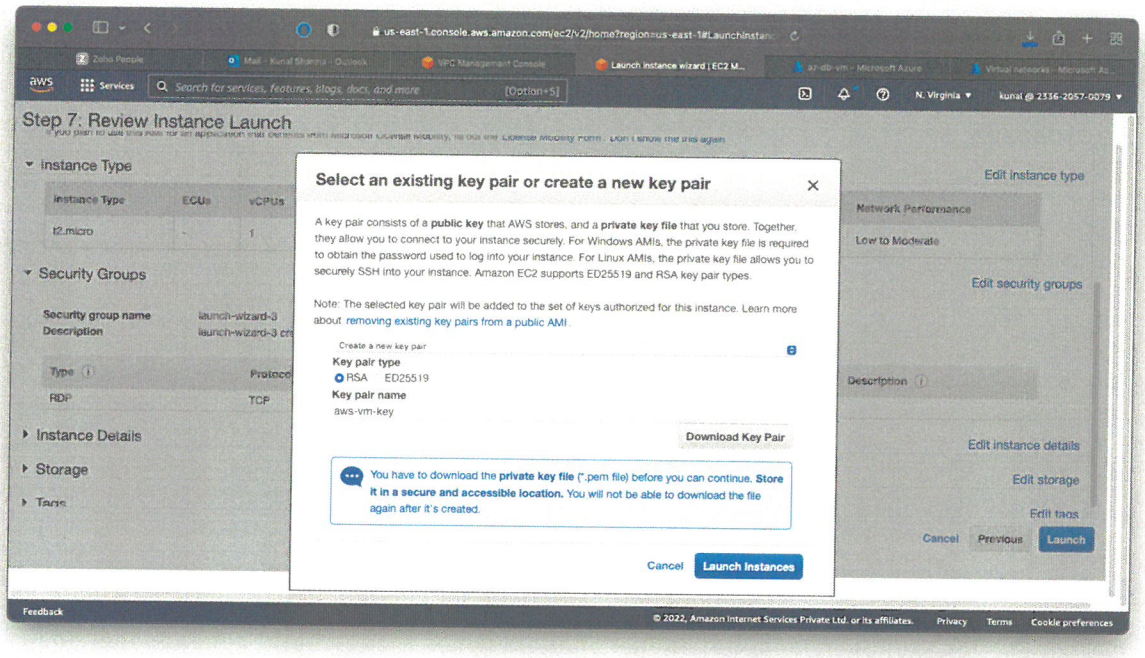
35. Use same vpc, subnet we have created and enable auto assign public IP and click on review and launch.



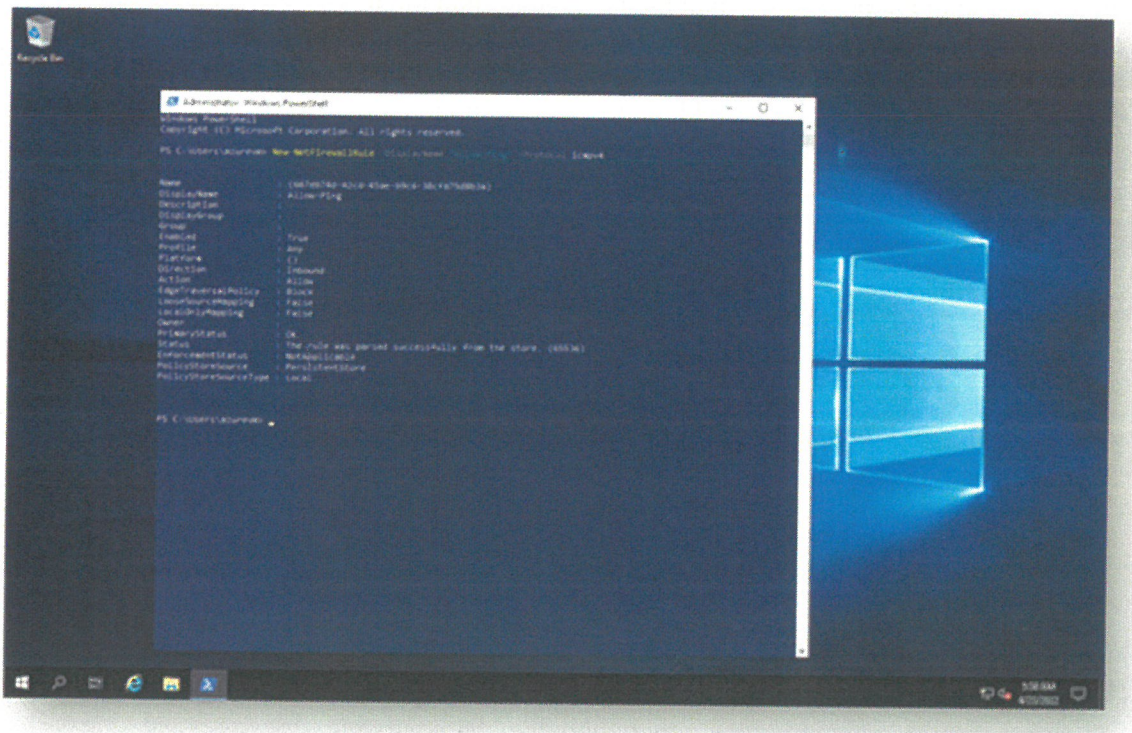
36. Instance created and status is 2/2 pass.



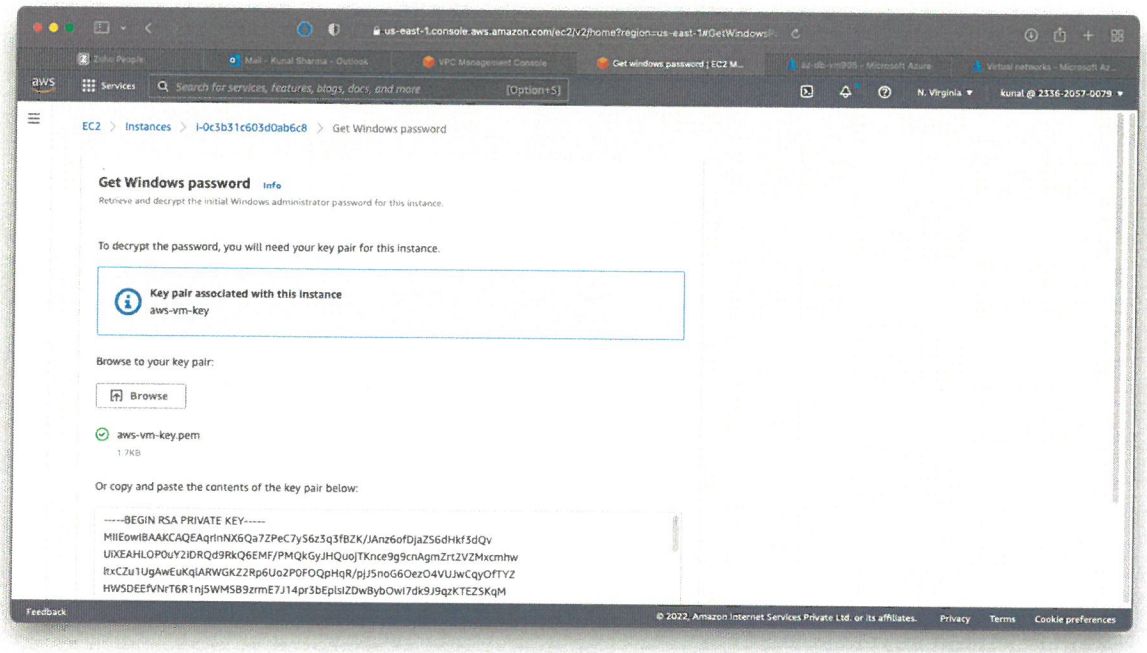
37. Create a new key pair for the instance created and download it.



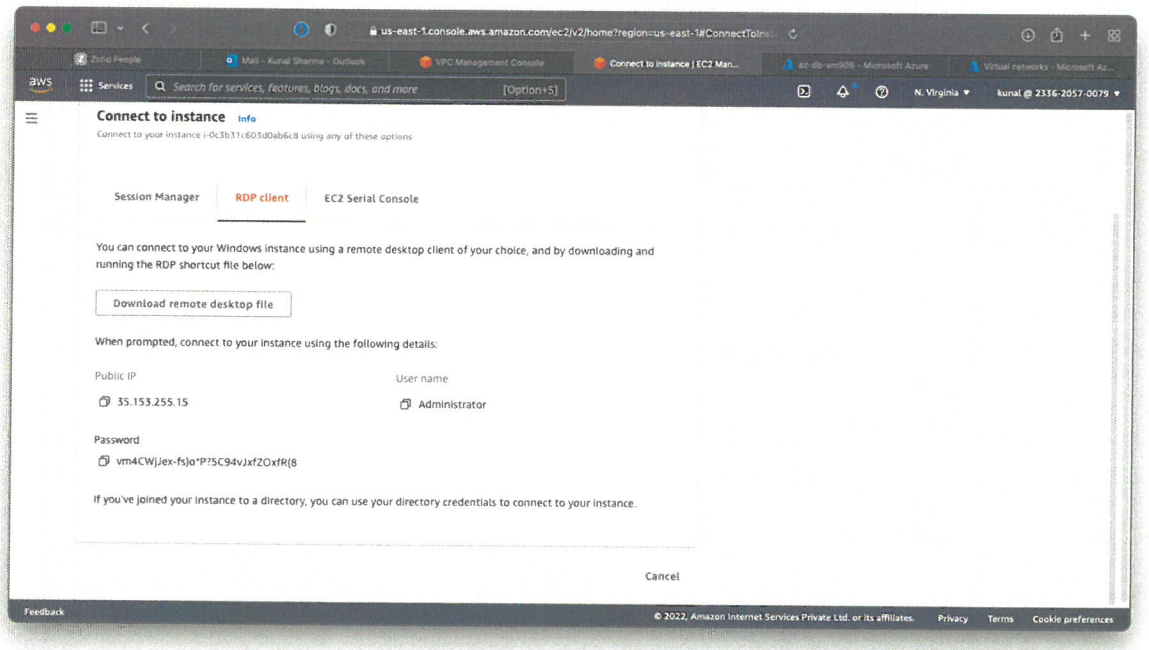
38. Open the RDP of azure vm and apply the following command on windows PowerShell to enable the ping functionality of the server.
New-NetFirewallRule –DisplayName “Allow-Ping” -Protocol icmpv4



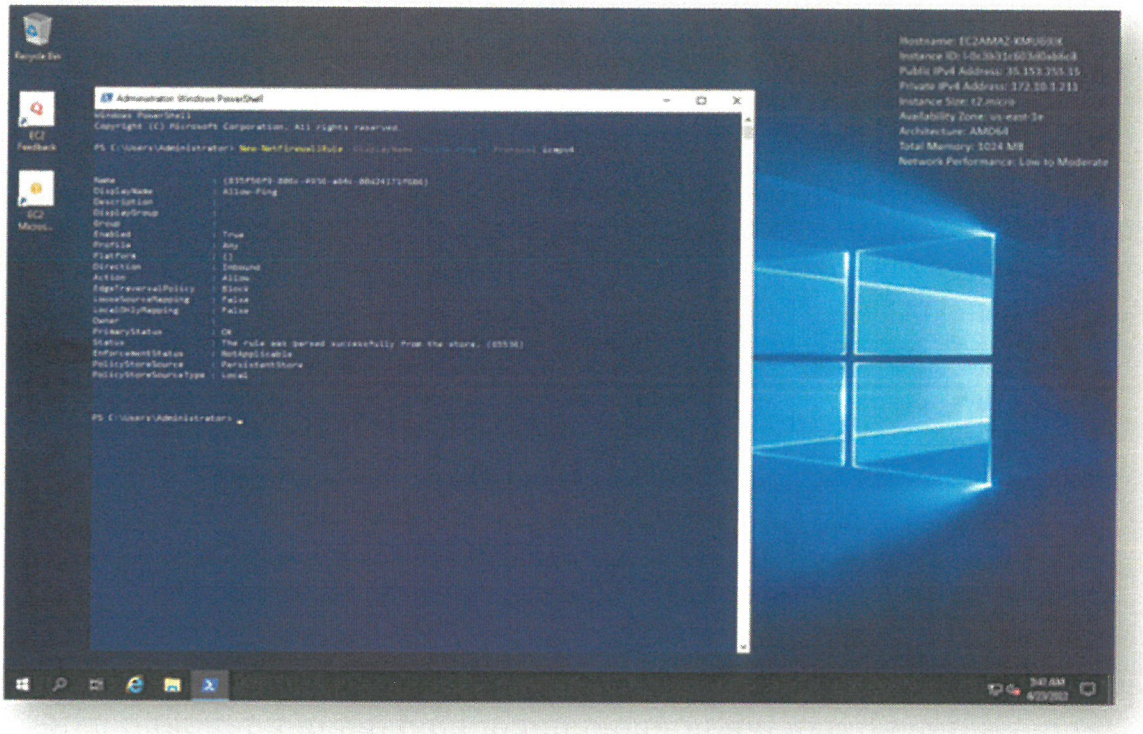
39. Click on connect button on the ec2 instance created.
Go to the rdp client and upload the key pair file downloaded that will decrypt the password for vm.



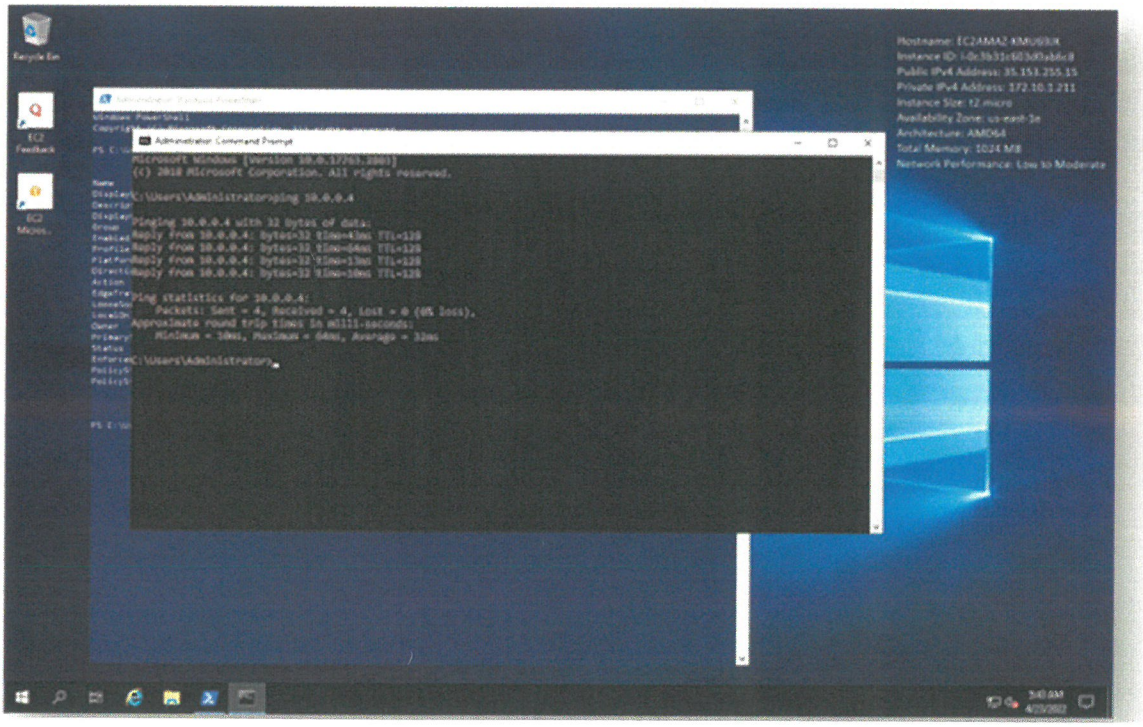
40. Username and password for the instance in generated, click on download remote desktop file .



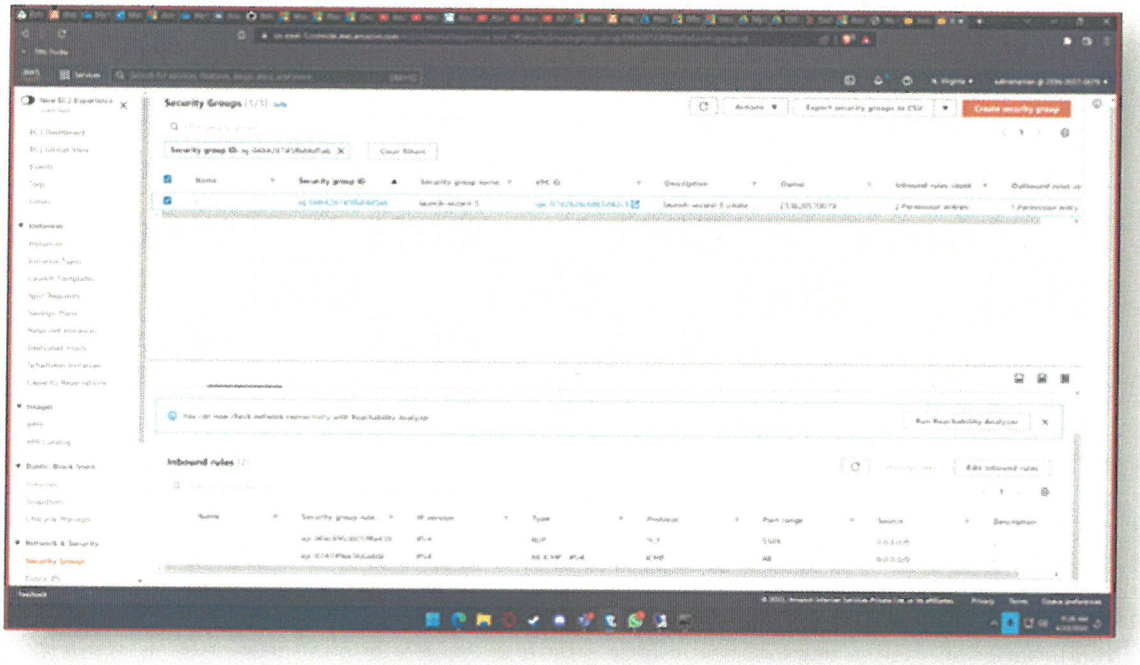
41. Open the RDP of aws vm and apply the following command on windows PowerShell to enable the ping functionality of the server.
New-NetFirewallRule –DisplayName “Allow-Ping” -Protocol icmpv4



42. Open the terminal of aws-vm and try to ping the private IP (10.0.0.4) of azure vm we have created , we should get a Reply from az vm to aws vm as shown.



43. If we try to ping from azure vm to aws vm directly we will get an timeout error because in AWS we need to create everything manually , so in NSG we only had RDP allow , add ICMP inbound it got solved.



RESULTS:

Since we have created a site to site VPN tunnel between AWS and Azure now we can transfer data between both using ADF pipelines and by setting up SHIR, this in short reduces our efforts. We can also scaleup the process of data migration using this.

This will also help us to shift our prem data from one cloud service to another, we are no more restricted to use a single cloud service now.

CHAPTER 4 : CONCLUSION

4.1 Benefits of site to site

Site-to-site VPNs are in use by many organizations. The reason for this is that they provide a number of benefits to enterprises and their employees, such as:

- **Secure Connectivity:** All traffic flowing over a site-to-site VPN is encrypted. This means that any business data crossing over the public Internet is encrypted, protecting it against eavesdropping and modification.

- **Simplified Network Architecture:** Organizations commonly use internal IP address ranges for devices within their LANs. These addresses need to be converted to external IP addresses to be accessible from the public Internet. With site-to-site VPNs, traffic from one LAN to another remains "internal", meaning that all sites can use internal addresses for each others' resources.

- **Access Control:** Some network resources are intended to only be accessible internally, meaning that employees at other sites should have access but not external users. Since site-to-site VPN users are "internal" users, access control rules are simpler to define because any traffic not originating from inside the network or entering via VPN tunnels can be blocked from accessing these resources.

4.2 Limitations of Site to Site

- **Limited Scalability:** A VPN provides point-to-point connectivity, meaning that a unique connection is required for each pair of connected sites. As a result, the number of VPNs required for a fully-connected network grows exponentially with the number of sites.

- **Inefficient Routing:** The limited scalability and lack of built-in security of VPNs drives some organizations to implement a "hub and spoke" network architecture, where all connections pass through the headquarters site for security inspection. While this reduces the number of VPN tunnels required within an organization, it can create significant network latency and additional load on the headquarters network.

- **Fragmented Visibility:** Each site-to-site VPN connection is independent from all of the others. This means that it can be difficult for an organization to maintain full, integrated visibility into its network traffic. As a result, attacks distributed across the corporate WAN may be more difficult to detect and respond to effectively.

4.3 SHIR

The integration runtime (IR) is the compute infrastructure that Azure Data Factory and Synapse pipelines use to provide data-integration capabilities across different network environments

A self-hosted integration runtime can run copy activities between a cloud data store and a data store in a private network. It also can dispatch transform activities against compute resources in an on-premises network or an Azure virtual network. The installation of a self-hosted integration runtime needs an on-premises machine or a virtual machine inside a private network.

- You can use a single self-hosted integration runtime for multiple on-premises data sources. You can also share it with another data factory within the same Azure Active Directory (Azure AD) tenant.

- The self-hosted integration runtime doesn't need to be on the same machine as the data source. However, having the self-hosted integration runtime close to the data source reduces the time for the self-hosted integration runtime to connect to the data source. We recommend that you install the self-hosted integration runtime on a machine that differs from the one that hosts the on-premises data source. When the self-hosted integration runtime and data source are on different machines, the self-hosted integration runtime doesn't compete with the data source for resources.

- Use a self-hosted integration runtime to support data integration within an Azure virtual network.

REFERENCES

1. https://celebaltech.com
2. https://en.wikipedia.org/wiki/Microsoft_Azure
3. https://en.wikipedia.org/wiki/Amazon_Web_Services
4. https://docs.aws.amazon.com/redshift/latest/mgmt/working-with-clusters.html#working-with-clusters-overview
5. https://docs.microsoft.com/en-us/azure/vpn-gateway/vpn-gateway-about-vpngateways
6. https://docs.aws.amazon.com/vpn/latest/s2svpn/your-cgw.html
7. https://aviatrix.com/learn-center/glossary/vgw/
8. https://openvpn.net/vpn-server-resources/site-to-site-routing-explained-in-detail/
9. https://docs.microsoft.com/en-us/azure/vpn-gateway/tutorial-site-to-site-portal
10. https://docs.aws.amazon.com/vpc/latest/userguide/VPC_Internet_Gateway.html
11. https://docs.aws.amazon.com/vpc/latest/userguide/VPC_Route_Tables.html

CELEBAL TECHNOLOGIES

hr@celebaltech.com
enterprisesales@celebaltech.com
www.celebaltech.com
0141-4700659
India: A-2 7th floor, UDB Corporate Tower, J.L.N. Marg, Jaipur

STRICTLY CONFIDENTIAL

Dated: 26th May, 2022

Ongoing Internship Certificate

To Whomsoever It May Concern

This is to certify that **Kunal Sharma** had undertaken Internship with Celebal Technologies, Jaipur, starting from 17th January 2022 in **Cloud Infra & Security** Department.

During his 4 months of Internship, he was working under the mentorship of **Mr. Harsh Goyal.**

Best Regards

Tejal

Tejal Mangal
Sr. Executive
Human Resources
Celebal Technologies