

AWS Assignment 4

1. Explain the Snowball concept.

Ans- Snowball is a petabyte-scale data transport solution that uses secure appliances to transfer large amounts of data into and out of the AWS cloud. Using Snowball addresses common challenges with large-scale data transfers including high network costs, long transfer times, and security concerns.

2. Make a distinction between NAT Gateways and NAT Instances.

Ans - The following is a high-level summary of the differences between NAT gateways and NAT instances.

Attribute	NAT gateway	NAT instance
Availability	Highly available. NAT gateways in each Availability Zone are implemented with redundancy. Create a NAT gateway in each Availability Zone to ensure zone-independent architecture.	Use a script to manage failover between instances.
Bandwidth	Scale up to 45 Gbps.	Depends on the bandwidth of the instance type.
Maintenance	Managed by AWS. You do not need to perform any maintenance.	Managed by you, for example, by installing software updates or operating system patches on the instance.
Performance	Software is optimized for handling NAT traffic.	A generic AMI that's configured to perform NAT.
Cost	Charged depending on the number of NAT gateways you use, duration of usage, and amount of data that you send through the NAT gateways.	Charged depending on the number of NAT instances that you use, duration of usage, and instance type and size.
Type and size	Uniform offering; you don't need to decide on the type or size.	Choose a suitable instance type and size, according to your predicted workload.
Public IP addresses	Choose the Elastic IP address to associate with a public NAT gateway at creation.	Use an Elastic IP address or a public IP address with a NAT instance. You can change the public IP address at any time by

Attribute	NAT gateway	NAT instance
		associating a new Elastic IP address with the instance.
Private IP addresses	Automatically selected from the subnet's IP address range when you create the gateway.	Assign a specific private IP address from the subnet's IP address range when you launch the instance.
Security groups	You cannot associate security groups with NAT gateways. You can associate them with the resources behind the NAT gateway to control inbound and outbound traffic.	Associate with your NAT instance and the resources behind your NAT instance to control inbound and outbound traffic.
Network ACLs	Use a network ACL to control the traffic to and from the subnet in which your NAT gateway resides.	Use a network ACL to control the traffic to and from the subnet in which your NAT instance resides.
Flow logs	Use flow logs to capture the traffic.	Use flow logs to capture the traffic.
Port forwarding	Not supported.	Manually customize the configuration to support port forwarding.
Bastion servers	Not supported.	Use as a bastion server.
Traffic metrics	View CloudWatch metrics for the NAT gateway .	View CloudWatch metrics for the instance.
Timeout behavior	When a connection times out, a NAT gateway returns an RST packet to any resources behind the NAT gateway that attempt to continue the connection (it does not send a FIN packet).	When a connection times out, a NAT instance sends a FIN packet to resources behind the NAT instance to close the connection.
IP fragmentation	Supports forwarding of IP fragmented packets for the UDP protocol. Does not support fragmentation for the TCP and ICMP protocols. Fragmented packets for these protocols will get dropped.	Supports reassembly of IP fragmented packets for the UDP, TCP, and ICMP protocols.

3. Describe the essential components of Amazon Web Services (AWS).

Ans- AWS Components are:

1. Data Management and Data Transfer

To run HPC applications in the AWS cloud, you need to move the required data into the cloud. There are several data transport solutions designed to securely transfer huge amounts of data. This overcomes issues like a long time for transfer, high network costs, and security concerns. Also, you can automate the movement of data between the AWS cloud and on-premises storage. There are options for establishing a private connection to the AWS from your premises. This increases bandwidth to provide more throughput, reduces the cost of the network, and provides a consistent network experience.

2. Compute & Networking

There are several compute instances types that can be customized according to your needs. It also handles monitoring your application and adjusting its capacity for maintaining a steady and predictable performance at an affordable cost. Also, setting up application scaling across multiple services for multiple resources takes a few minutes. Enhanced networking options from AWS allow achieving lower inter-instance latency and higher bandwidth.

3. Storage

When looking for an HPC solution, you need to consider the storage options and cost. There are several flexible blocks, object, and file storage options in AWS services that allow permanent and transient data storage. It allows allocating storage volumes according to the size you need. You can store and access and data type over the cloud without doing a data migration project. Also, with AWS services, you can transfer your workload to the cloud from on-premises.

4. Automation and Orchestration

For using the infrastructure efficiently, you need to automate scheduling submitted jobs and the job submission process. AWS services allow you to run thousands of batch computing jobs through the dynamic provision of the computer resources on the basis of the requirements.

5. Operations and Management

As a system administrator, you are responsible for avoiding cost overruns and monitoring the infrastructure. There are several management and monitoring services that allow you to optimize utilization of resources, monitor the application, get a complete view of operational health, and respond to the performance changes.

6. Visualization

With the AWS services, you can easily visualize the engineering simulations' results without moving huge amounts of data. Now, you can access the interactive applications remotely over a standard network and deliver application sessions to any workstation.

7. Security and Compliance

For running applications on the cloud, you need to have an understanding of regulatory compliance and security management. There are several quick-launch templates and security related services offered by

AWS that helps in protecting data and customer privacy by putting strong safeguards in the AWS infrastructure.

4. When should you utilize a spin-up server? Use examples to demonstrate your point?

Ans - It is a general term relating to starting a new instance of particular service on Amazon Web Services. Let's say that you need a Linux based server to host your WordPress website. One way you might do this is to "spin up" a new EC2 instance (a virtual machine running Linux). Through the Amazon console, **you would be reserving a part of a physical server** where you could install your own software and start that server as you wish.

Since Amazon has literally hundreds of different types of services (EC2, RDS for databases, Elasticache for caching, etc) to "spin up" an AWS instance refers to starting any of those services offered on AWS.

5. Explain the concept of outlier car scaling.

Ans – Many machine learning algorithms perform better when numerical input variables are scaled to a standard range.

This includes algorithms that use a weighted sum of the input, like linear regression, and algorithms that use distance measures, like k-nearest neighbors.

Standardizing is a popular scaling technique that subtracts the mean from values and divides by the standard deviation, transforming the probability distribution for an input variable to a standard Gaussian (zero mean and unit variance). Standardization can become skewed or biased if the input variable contains outlier values.

To overcome this, the median and interquartile range can be used when standardizing numerical input variables, generally referred to as robust scaling