



**edureka!**  
LEARNING CENTER



# Cloud and DevOps Program

Developing Cloud Solutions using AWS

Interview Preparation



# AWS Interview Preparation

Amazon Web Service (AWS) is an online platform providing cost-effective, scalable cloud computing solutions. In this AWS Interview Preparation document, the most frequently asked AWS interview questions have been compiled.

Let us first begin with some Basic AWS Interview Questions.

**Q1. Try this AWS scenario-based interview question. I have some private servers on my premises, also I have distributed some of my workloads on the public cloud. What is this architecture called?**

- A. Virtual Private Network
- B. Private Cloud
- C. Virtual Private Cloud
- D. Hybrid Cloud

**Ans: D.**

**Explanation:** This type of architecture would be a hybrid cloud. Why? Because we are using both, the public cloud, and your on premises servers i.e. the private cloud. To make this hybrid architecture easy to use, wouldn't it be better if your private and public cloud were all on the same network(virtually). This is established by including your public cloud servers in a virtual private cloud, and connecting this virtual cloud with your on premise servers using a VPN(Virtual Private Network).

**Q2. What does the following command do with respect to the Amazon EC2 security groups?**

**ec2-create-group CreateSecurityGroup**

- A. Groups the user created security groups into a new group for easy access.
- B. Creates a new security group for use with your account.
- C. Creates a new group inside the security group.
- D. Creates a new rule inside the security group.

**Ans: B**

**Explanation:** A Security group is just like a firewall, it controls the traffic in and out of your instance. In AWS terms, the inbound and outbound traffic. The command mentioned is pretty straight forward, it says create security group, and does the same. Moving along, once your security group is created, you can add different rules in it. For example, you have an RDS instance, to access it, you have to add the public IP address of the machine from which you want access the instance in its security group.

**Q3. Here is AWS scenario-based interview question. You have a video trans-coding application. The videos are processed according to a queue. If the processing of a video is interrupted in one instance, it is resumed in another instance. Currently there is a huge back-log of videos which needs to be processed, for this you need to add more instances, but you need these instances only until your backlog is reduced. Which of these would be an efficient way to do it?**

**Ans:** You should be using an **On Demand** instance for the same. Why? First of all, the workload has to be processed now, meaning it is urgent, secondly you don't need them once your backlog is cleared, therefore Reserved Instance is out of the picture, and since the work is urgent, you cannot stop the work on your instance just because the spot price spiked, therefore Spot Instances shall also not be used. Hence On-Demand instances shall be the right choice in this case.

**Q4. You have a distributed application that periodically processes large volumes of data across multiple Amazon EC2 Instances. The application is designed to recover gracefully from Amazon EC2 instance failures. You are required to accomplish this task in the most cost-effective way.**

**Which of the following will meet your requirements?**

- A. Spot Instances
- B. Reserved instances
- C. Dedicated instances
- D. On-Demand instances

**Ans: A**

**Explanation:** Since the work we are addressing here is not continuous, a reserved instance shall be idle at times, same goes with On Demand instances. Also it does not make sense to launch an On Demand instance whenever work comes up, since it is expensive. Hence Spot Instances will be the right fit because of their low rates and no long term commitments.

**Q5. How is stopping and terminating an instance different from each other?**

**Ans:** Starting, stopping and terminating are the three states in an EC2 instance, let's discuss them in detail:

- **Stopping and Starting** an instance: When an instance is stopped, the instance performs a normal shutdown and then transitions to a stopped state. All of its Amazon EBS volumes remain attached, and you can start the instance again at a later time. You are not charged for additional instance hours while the instance is in a stopped state.
- **Terminating** an instance: When an instance is terminated, the instance performs a normal shutdown, then the attached Amazon EBS volumes are deleted unless the volume's *deleteOnTermination* attribute is set to false. The instance itself is also deleted, and you can't start the instance again at a later time.

**Q6. If I want my instance to run on single-tenant hardware, which value do I have to set the instance's tenancy attribute to?**

- A. Dedicated
- B. Isolated
- C. One
- D. Reserved

**Ans: A.**

**Explanation:** The Instance tenancy attribute should be set to Dedicated Instance. The rest of the values are invalid.

**Q7. When will you incur costs with an Elastic IP address (EIP)?**

- A. When an EIP is allocated.
- B. When it is allocated and associated with a running instance.
- C. When it is allocated and associated with a stopped instance.
- D. Costs are incurred regardless of whether the EIP is associated with a running instance.

**Ans: C.**

**Explanation:** You are not charged if only one Elastic IP address is attached to your running instance. But you do get charged in the following conditions:

- When you use more than one Elastic IPs with your instance.
- When your Elastic IP is attached to a stopped instance.
- When your Elastic IP is not attached to any instance.

**Q8. How is a Spot instance different from an On-Demand instance or Reserved Instance?**

**Ans:** First of all, let's understand that Spot Instance, On-Demand instance and Reserved Instances are all models for pricing. Moving along, spot instances provide the ability for customers to purchase compute capacity with no upfront commitment, at hourly rates usually lower than the On-Demand rate in each region. Spot instances are just like bidding, the bidding price is called Spot Price. The Spot Price fluctuates based on supply and demand for instances, but customers will never pay more than the maximum price they have specified. If the Spot Price moves higher than a customer's maximum price, the customer's EC2 instance will be shut down automatically. But the reverse is not true, if the Spot prices come down again, your EC2 instance will not be launched automatically, one has to do that manually. In Spot and On demand instance, there is no commitment for the duration from the user side, however in reserved instances one has to stick to the time period that he has chosen.

**Q9. Are the Reserved Instances available for Multi-AZ Deployments?**

- A. Multi-AZ Deployments are only available for Cluster Compute instances types
- B. Available for all instance types
- C. Only available for M3 instance types
- D. D. Not Available for Reserved Instances

**Ans: B**

**Explanation:** Reserved Instances is a pricing model, which is available for all instance types in EC2.

**Q10. How to use the processor state control feature available on the c4.8xlarge instance?**

**Ans:** The processor state control consists of 2 states:

- The C state – Sleep state varying from c0 to c6. C6 being the deepest sleep state for a processor
- The P state – Performance state p0 being the highest and p15 being the lowest possible frequency.

Now, why the C state and P state. Processors have cores, these cores need thermal headroom to boost their performance. Now since all the cores are on the processor the temperature should be kept at an optimal state so that all the cores can perform at the highest performance.

Now how will these states help in that? If a core is put into sleep state it will reduce the overall temperature of the processor and hence other cores can perform better. Now the same can be synchronized with other cores, so that the processor can boost as many cores it can by timely putting other cores to sleep, and thus get an overall performance boost.

Concluding, the C and P state can be customized in some EC2 instances like the c4.8xlarge instance and thus you can customize the processor according to your workload.

**Q11. What kind of network performance parameters can you expect when you launch instances in cluster placement group?**

**Ans:** The network performance depends on the instance type and network performance specification, if launched in a placement group you can expect up to

- 10 Gbps in a single-flow,
- 20 Gbps in multiframe i.e full duplex
- Network traffic outside the placement group will be limited to 5 Gbps(full duplex).

**Q12. To deploy a 4-node cluster of Hadoop in AWS, which instance type can be used?**

**Ans:** First let's understand what actually happens in a Hadoop cluster, the Hadoop cluster follows a master slave concept. The master machine processes all the data, slave machines store the data and act as data nodes. Since all the storage happens at the slave, a higher capacity hard disk would be recommended and since master does all the processing, a higher RAM and a much better CPU is required. Therefore, you can select the configuration of your machine depending on your workload. For e.g. – In this case c4.xlarge will be preferred for master machine whereas for slave machine we can select i2.large instance. If you don't want to deal with configuring your instance and installing hadoop cluster manually, you can straight away launch an Amazon EMR (Elastic Map Reduce) instance which automatically configures the servers for you. You dump your data to be processed in S3, EMR picks it from there, processes it, and dumps it back into S3.

**Q13. Where do you think an AMI fits, when you are designing an architecture for a solution?**

**Ans:** AMIs(Amazon Machine Images) are like templates of virtual machines and an instance is derived from an AMI. AWS offers pre-baked AMIs which you can choose while you are launching an instance, some AMIs are not free, therefore can be bought from the AWS Marketplace. You can also choose to create your own custom AMI which would help you save space on AWS. For example if you don't need a set of software on your installation, you can customize your AMI to do that. This makes it cost efficient, since you are removing the unwanted things.

**Q14. How do you choose an Availability Zone?**

**Ans:** Let's understand this through an example, consider there's a company which has user base in India as well as in the US.

Let us see how we will choose the region for this use case :



<b>Regions</b>	• Mumbai/N Virginia
<b>Instance Type</b> (Reserved Instance)	• e.g. amazon ec2- m4.4xlarge 16(vCPU), 64 GB RAM
<b>Pricing(1 Year)</b>	• Mumbai - \$691/monthly - \$0.9 hourly • N Virginia - \$480/monthly - <b>\$0.6 hourly</b>
<b>Latency</b>	• From USA to India - <b>Low</b> • From India to USA - High

So, with reference to the above figure the regions to choose between are, Mumbai and North Virginia. Now let us first compare the pricing, you have hourly prices, which can be converted to your per month figure. Here North Virginia emerges as a winner. But, pricing cannot be the only parameter to consider. Performance should also be kept in mind hence, let's look at latency as well. Latency basically is the time that a server takes to respond to your requests i.e the response time. North Virginia wins again!

So concluding, North Virginia should be chosen for this use case.

#### Q15. Is one Elastic IP address enough for every instance that I have running?

**Ans:** Depends! Every instance comes with its own private and public address. The private address is associated exclusively with the instance and is returned to Amazon EC2 only when it is stopped or terminated. Similarly, the public address is associated exclusively with the instance until it is stopped or terminated. However, this can be replaced by the Elastic IP address, which stays with the instance as long as the user doesn't manually detach it. But what if you are hosting multiple websites on your EC2 server, in that case you may require more than one Elastic IP address.

#### Q16. What are the best practices for Security in Amazon EC2?

**Ans:** There are several best practices to secure Amazon EC2. A few of them are given below:

- Use AWS Identity and Access Management (IAM) to control access to your AWS resources.

- Restrict access by only allowing trusted hosts or networks to access ports on your instance.
- Review the rules in your security groups regularly, and ensure that you apply the principle of least
- Privilege – only open up permissions that you require.
- Disable password-based logins for instances launched from your AMI. Passwords can be found or cracked, and are a security risk.

**Q17. Another scenario-based interview question. You need to configure an Amazon S3 bucket to serve static assets for your public-facing web application. Which method will ensure that all objects uploaded to the bucket are set to public read?**

- A. Set permissions on the object to public read during upload.
- B. Configure the bucket policy to set all objects to public read.
- C. Use AWS Identity and Access Management roles to set the bucket to public read.
- D. Amazon S3 objects default to public read, so no action is needed.

**Ans: B**

**Explanation:** Rather than making changes to every object, its better to set the policy for the whole bucket. IAM is used to give more granular permissions, since this is a website, all objects would be public by default.

**Q18. A customer wants to leverage Amazon Simple Storage Service (S3) and Amazon Glacier as part of their backup and archive infrastructure. The customer plans to use third-party software to support this integration. Which approach will limit the access of the third party software to only the Amazon S3 bucket named “company-backup”?**

- A. A custom bucket policy limited to the Amazon S3 API in three Amazon Glacier archive “company-backup”
- B. A custom bucket policy limited to the Amazon S3 API in “company-backup”
- C. A custom IAM user policy limited to the Amazon S3 API for the Amazon Glacier archive “company-backup”.
- D. A custom IAM user policy limited to the Amazon S3 API in “company-backup”.

**Ans: D**

**Explanation:** Taking queue from the previous questions, this use case involves more granular permissions, hence IAM would be used here.

**Q19. Can S3 be used with EC2 instances, if yes, how?**

**Ans:** Yes, it can be used for instances with root devices backed by local instance storage. By using Amazon S3, developers have access to the same highly scalable, reliable, fast, inexpensive data storage infrastructure that Amazon uses to run its own global network of web sites. In order to execute systems in the Amazon EC2 environment, developers use the tools provided to load their Amazon Machine Images (AMIs) into Amazon S3 and to move them between Amazon S3 and Amazon EC2.

Another use case could be for websites hosted on EC2 to load their static content from S3.

**Q20. A customer implemented AWS Storage Gateway with a gateway-cached volume at their main office. An event takes the link between the main and branch office offline. Which methods will enable the branch office to access their data?**

- A. Restore by implementing a lifecycle policy on the Amazon S3 bucket.
- B. Make an Amazon Glacier Restore API call to load the files into another Amazon S3 bucket within four to six hours.
- C. Launch a new AWS Storage Gateway instance AMI in Amazon EC2, and restore from a gateway snapshot.
- D. Create an Amazon EBS volume from a gateway snapshot, and mount it to an Amazon EC2 instance.

**Ans: C**

**Explanation:** The fastest way to do it would be launching a new storage gateway instance. Why? Since time is the key factor which drives every business, troubleshooting this problem will take more time. Rather than we can just restore the previous working state of the storage gateway on a new instance.

**Q21. When you need to move data over long distances using the internet, for instance across countries or continents to your Amazon S3 bucket, which method or service will you use?**

- A. Amazon Glacier
- B. Amazon CloudFront
- C. Amazon Transfer Acceleration
- D. Amazon Snowball

**Ans: C**

**Explanation:** You would not use Snowball, because for now, the snowball service does not support cross region data transfer, and since, we are transferring across countries, Snowball cannot be used. Transfer Acceleration

shall be the right choice here as it throttles your data transfer with the use of optimized network paths and Amazon's content delivery network upto 300% compared to normal data transfer speed.

**Q22. How can you speed up data transfer in Snowball?**

**Ans:** The data transfer can be increased in the following way:

- By performing multiple copy operations at one time i.e. if the workstation is powerful enough, you can initiate multiple cp commands each from different terminals, on the same Snowball device.
- Copying from multiple workstations to the same snowball.
- Transferring large files or by creating a batch of small file, this will reduce the encryption overhead.
- Eliminating unnecessary hops i.e. make a setup where the source machine(s) and the snowball are the only machines active on the switch being used, this can hugely improve performance.

**Q23. If you want to launch Amazon Elastic Compute Cloud (EC2) instances and assign each instance a predetermined private IP address you should:**

- A. Launch the instance from a private Amazon Machine Image (AMI).
- B. Assign a group of sequential Elastic IP address to the instances.
- C. Launch the instances in the Amazon Virtual Private Cloud (VPC).
- D. Launch the instances in a Placement Group.

**Ans: C**

**Explanation:** The best way of connecting to your cloud resources (for ex- ec2 instances) from your own data center (for eg- private cloud) is a VPC. Once you connect your datacenter to the VPC in which your instances are present, each instance is assigned a private IP address which can be accessed from your

datacenter. Hence, you can access your public cloud resources, as if they were on your own network.

**Q24. Can I connect my corporate datacenter to the Amazon Cloud?**

**Ans:** Yes, you can do this by establishing a VPN(Virtual Private Network) connection between your company's network and your VPC (Virtual Private Cloud), this will allow you to interact with your EC2 instances as if they were within your existing network.

**Q25. Is it possible to change the private IP addresses of an EC2 while it is running/stopped in a VPC?**

**Ans:** Primary private IP address is attached with the instance throughout its lifetime and cannot be changed, however secondary private addresses can be unassigned, assigned or moved between interfaces or instances at any point.

**Q26. Why do you make subnets?**

- A. Because there is a shortage of networks
- B. To efficiently utilize networks that have a large no. of hosts.
- C. Because there is a shortage of hosts.
- D. To efficiently utilize networks that have a small no. of hosts.

**Ans: B**

**Explanation:** If there is a network which has a large no. of hosts, managing all these hosts can be a tedious job. Therefore we divide this network into subnets (sub-networks) so that managing these hosts becomes simpler.

**Q27. Which of the following is true?**

- A. You can attach multiple route tables to a subnet
- B. You can attach multiple subnets to a route table
- C. Both A and B
- D. None of these.

**Ans: B**

**Explanation:** Route Tables are used to route network packets, therefore in a subnet having multiple route tables will lead to confusion as to where the packet has to go. Therefore, there is only one route table in a subnet, and since a route table can have any no. of records or information, hence attaching multiple subnets to a route table is possible.

**Q28. In CloudFront what happens when content is NOT present at an Edge location and a request is made to it?**

- A. An Error “404 not found” is returned
- B. CloudFront delivers the content directly from the origin server and stores it in the cache of the edge location
- C. The request is kept on hold till content is delivered to the edge location
- D. The request is routed to the next closest edge location

**Ans: B**

**Explanation:** CloudFront is a content delivery system, which caches data to the nearest edge location from the user, to reduce latency. If data is not present at an edge location, the first time the data may get transferred from the original server, but from the next time, it will be served from the cached edge.

**Q29. If I'm using Amazon CloudFront, can I use Direct Connect to transfer objects from my own data center?**

**Ans:** Yes. Amazon CloudFront supports custom origins including origins from outside of AWS. With AWS Direct Connect, you will be charged with the respective data transfer rates.

**Q30. If my AWS Direct Connect fails, will I lose my connectivity?**

**Ans:** If a backup AWS Direct connect has been configured, in the event of a failure it will switch over to the second one. It is recommended to enable Bidirectional Forwarding Detection (BFD) when configuring your connections to ensure faster detection and failover. On the other hand, if you have configured a backup IPsec VPN connection instead, all VPC traffic will failover to the backup VPN connection automatically. Traffic to/from public resources such as Amazon S3 will be routed over the Internet. If you do not have a backup AWS Direct Connect link or a IPsec VPN link, then Amazon VPC traffic will be dropped in the event of a failure.

**Q31. If I launch a standby RDS instance, will it be in the same Availability Zone as my primary?**

- A. Only for Oracle RDS types
- B. Yes
- C. Only if it is configured at launch
- D. No

**Ans: D**

**Explanation:** No, since the purpose of having a standby instance is to avoid an infrastructure failure (if it happens), therefore the standby instance is stored in a different availability zone, which is a physically different independent infrastructure.



**Q32. When would I prefer Provisioned IOPS over Standard RDS storage?**

- A. If you have batch-oriented workloads
- B. If you use production online transaction processing (OLTP) workloads.
- C. If you have workloads that are not sensitive to consistent performance
- D. All of the above

**Ans: A.**

**Explanation:** Provisioned IOPS deliver high IO rates but on the other hand it is expensive as well. Batch processing workloads do not require manual intervention they enable full utilization of systems, therefore a provisioned IOPS will be preferred for batch oriented workload.

**Q33. How are Amazon RDS, DynamoDB, and Redshift different?**

**Ans:** Let us see how Amazon RDS, DynamoDB, and Redshift are different from each other.

- Amazon RDS is a database management service for relational databases, it manages patching, upgrading, backing up of data etc. of databases for you without your intervention. RDS is a Db management service for structured data only.
- DynamoDB, on the other hand, is a NoSQL database service, NoSQL deals with unstructured data.
- Redshift, is an entirely different service, it is a data warehouse product and is used in data analysis.

**Q34. If I am running my DB Instance as a Multi-AZ deployment, can I use the standby DB Instance for read or write operations along with primary DB instance?**

- A. Yes
- B. Only with MySQL based RDS
- C. Only for Oracle RDS instances
- D. No

**Ans: D**

**Explanation:** No, Standby DB instance cannot be used with primary DB instance in parallel, as the former is solely used for standby purposes, it cannot be used unless the primary instance goes down.

**Q35. Your company's branch offices are all over the world, they use a software with a multi-regional deployment on AWS, they use MySQL 5.6 for data persistence.**

**The task is to run an hourly batch process and read data from every region to compute cross-regional reports which will be distributed to all the branches. This should be done in the shortest time possible. How will you build the DB architecture in order to meet the requirements?**

- A. For each regional deployment, use RDS MySQL with a master in the region and a read replica in the HQ region
- B. For each regional deployment, use MySQL on EC2 with a master in the region and send hourly EBS snapshots to the HQ region
- C. For each regional deployment, use RDS MySQL with a master in the region and send hourly RDS snapshots to the HQ region
- D. For each regional deployment, use MySQL on EC2 with a master in the region and use S3 to copy data files hourly to the HQ region

**Ans: A**

**Explanation:** For this we will take an RDS instance as a master, because it will manage our database for us and since we have to read from every region, we'll put a read replica of this instance in every region where the data has to be read from. Option C is not correct since putting a read replica would be more efficient than putting a snapshot, a read replica can be promoted if needed to an independent DB instance, but with a Db snapshot it becomes mandatory to launch a separate DB Instance.

**Q36. Can I run more than one DB instance for Amazon RDS for free?**

**Ans:** Yes. You can run more than one Single-AZ Micro database instance, that too for free! However, any use exceeding 750 instance hours, across all Amazon RDS Single-AZ Micro DB instances, across all eligible database engines and regions, will be billed at standard Amazon RDS prices. For example: if you run two Single-AZ Micro DB instances for 400 hours each in a single month, you will accumulate 800 instance hours of usage, of which 750 hours will be free. You will be billed for the remaining 50 hours at the standard Amazon RDS price.

**Q37. Which AWS services will you use to collect and process e-commerce data for near real-time analysis?**

- A. Amazon ElastiCache
- B. Amazon DynamoDB
- C. Amazon Redshift
- D. Amazon Elastic MapReduce

**Ans: B, C**

**Explanation:** DynamoDB is a fully managed NoSQL database service. DynamoDB, therefore can be fed any type of unstructured data, which can be data from e-commerce websites as well, and later, an analysis can be done on them using Amazon Redshift. We are not using Elastic MapReduce, since a near real time analyses is needed.

**Q38. Can I retrieve only a specific element of the data, if I have a nested JSON data in DynamoDB?**

**Ans:** Yes. When using the GetItem, BatchGetItem, Query or Scan APIs, you can define a Projection Expression to determine which attributes should be retrieved from the table. Those attributes can include scalars, sets, or elements of a JSON document.

**Q39. A company is deploying a new two-tier web application in AWS. The company has limited staff and requires high availability, and the application requires complex queries and table joins. Which configuration provides the solution for the company's requirements?**

- A. MySQL Installed on two Amazon EC2 Instances in a single Availability Zone
- B. Amazon RDS for MySQL with Multi-AZ
- C. Amazon ElastiCache
- D. Amazon DynamoDB

**Ans: D**

**Explanation:** DynamoDB has the ability to scale more than RDS or any other relational database service, therefore DynamoDB would be the apt choice.

**Q40. What happens to my backups and DB Snapshots if I delete my DB Instance?**

**Ans:** When you delete a DB instance, you have an option of creating a final DB snapshot, if you do that you can restore your database from that snapshot. RDS retains this user-created DB snapshot along with all other manually created DB snapshots after the instance is deleted, also automated backups are deleted and only manually created DB Snapshots are retained.

**Q41. Which of the following use cases are suitable for Amazon DynamoDB?  
Choose 2 answers**

- A. Managing web sessions.
- B. Storing JSON documents.
- C. Storing metadata for Amazon S3 objects.
- D. Running relational joins and complex updates.

**Ans: C, D.**

**Explanation:** If all your JSON data have the same fields eg [id,name,age] then it would be better to store it in a relational database, the metadata on the other hand is unstructured, also running relational joins or complex updates would work on DynamoDB as well.

**Q42. How can I load my data to Amazon Redshift from different data sources like Amazon RDS, Amazon DynamoDB and Amazon EC2?**

**Ans:** You can load the data in the following two ways:

- You can use the COPY command to load data in parallel directly to Amazon Redshift from Amazon EMR, Amazon DynamoDB, or any SSH-enabled host.
- AWS Data Pipeline provides a high performance, reliable, fault tolerant solution to load data from a variety of AWS data sources. You can use AWS Data Pipeline to specify the data source, desired data transformations, and then execute a pre-written import script to load your data into Amazon Redshift.

**Q43. When an instance is unhealthy, it is terminated and replaced with a new one, which of the following services does that?**

- A. Sticky Sessions
- B. Fault Tolerance
- C. Connection Draining
- D. Monitoring

**Ans: B**

**Explanation:** When ELB detects that an instance is unhealthy, it starts routing incoming traffic to other healthy instances in the region. If all the instances in a region becomes unhealthy, and if you have instances in some other availability zone/region, your traffic is directed to them. Once your instances become healthy again, they are rerouted back to the original instances.

**Q44. You are running a website on EC2 instances deployed across multiple Availability Zones with a Multi-AZ RDS MySQL Extra Large DB Instance. The site performs a high number of small reads and writes per second and relies on an eventual consistency model. After comprehensive tests you discover that there is read contention on RDS MySQL. Which are the best approaches to meet these requirements? (Choose 2 answers)**

- A. Deploy ElastiCache in-memory cache running in each availability zone
- B. Implement sharding to distribute load to multiple RDS MySQL instances
- C. Increase the RDS MySQL Instance size and Implement provisioned IOPS
- D. Add an RDS MySQL read replica in each availability zone

**Ans: A, C**

**Explanation:** Since it does a lot of read writes, provisioned IO may become expensive. But we need high performance as well, therefore the data can be cached using ElastiCache which can be used for frequently reading the data. As for RDS since read contention is happening, the instance size should be

increased and provisioned IO should be introduced to increase the performance.

**Q45. A startup is running a pilot deployment of around 100 sensors to measure street noise and air quality in urban areas for 3 months. It was noted that every month around 4GB of sensor data is generated. The company uses a load balanced auto scaled layer of EC2 instances and a RDS database with 500 GB standard storage. The pilot was a success and now they want to deploy at least 100K sensors which need to be supported by the backend. You need to store the data for at least 2 years to analyze it. Which setup of the following would you prefer?**

- A. Add an SQS queue to the ingestion layer to buffer writes to the RDS instance
- B. Ingest data into a DynamoDB table and move old data to a Redshift cluster
- C. Replace the RDS instance with a 6 node Redshift cluster with 96TB of storage
- D. Keep the current architecture but upgrade RDS storage to 3TB and 10K provisioned IOPS

**Ans: C**

**Explanation:** A Redshift cluster would be preferred because it is easy to scale, also the work would be done in parallel through the nodes, therefore is perfect for a bigger workload like our use case. Since each month 4 GB of data is generated, therefore in 2 year, it should be around 96 GB. And since the servers will be increased to 100K in number, 96 GB will approximately become 96TB. Hence option C is the right answer.

**Q46. Suppose you have an application where you have to render images and also do some general computing. From the following services which service will best fit your need?**

- A. Classic Load Balancer
- B. Application Load Balancer
- C. Both of them
- D. None of these

**Ans: B**

**Explanation:** You will choose an application load balancer, since it supports path based routing, which means it can take decisions based on the URL, therefore if your task needs image rendering it will route it to a different instance, and for general computing it will route it to a different instance.

**Q47. What is the difference between Scalability and Elasticity?**

**Ans:** Scalability is the ability of a system to increase its hardware resources to handle the increase in demand. It can be done by increasing the hardware specifications or increasing the processing nodes.

Elasticity is the ability of a system to handle increase in the workload by adding additional hardware resources when the demand increases(same as scaling) but also rolling back the scaled resources, when the resources are no longer needed. This is particularly helpful in Cloud environments, where a pay per use model is followed.

**Q48. How will you change the instance type for instances which are running in your application tier and are using Auto Scaling. Where will you change it from the following areas?**

- A. Auto Scaling policy configuration
- B. Auto Scaling group



- C. Auto Scaling tags configuration
- D. Auto Scaling launch configuration

**Ans: D**

**Explanation:** Auto scaling tags configuration, is used to attach metadata to your instances, to change the instance type you have to use auto scaling launch configuration.

**Q49. You have a content management system running on an Amazon EC2 instance that is approaching 100% CPU utilization. Which option will reduce load on the Amazon EC2 instance?**

- A. Create a load balancer, and register the Amazon EC2 instance with it
- B. Create a CloudFront distribution, and configure the Amazon EC2 instance as the origin
- C. Create an Auto Scaling group from the instance using the CreateAutoScalingGroup action
- D. Create a launch configuration from the instance using the CreateLaunchConfigurationAction

**Ans: A**

**Explanation:** Creating alone an autoscaling group will not solve the issue, until you attach a load balancer to it. Once you attach a load balancer to an autoscaling group, it will efficiently distribute the load among all the instances. Option B – CloudFront is a CDN, it is a data transfer tool therefore will not help reduce load on the EC2 instance. Similarly the other option – Launch configuration is a template for configuration which has no connection with reducing loads.

**Q50. When should I use a Classic Load Balancer, and when should I use an Application load balancer?**

**Ans:** A Classic Load Balancer is ideal for simple load balancing of traffic across multiple EC2 instances, while an Application Load Balancer is ideal for microservices or container-based architectures where there is a need to route traffic to multiple services or load balance across multiple ports on the same EC2 instance.

