|  |
| --- |
| **What are the different components used in AWS?** The components that are used in AWS are: **1. Amazon S3:** it is used to retrieve input data sets that are involved in making a cloud architecture and also used to store the output data sets that is the result of the input. **2. Amazon SQS:** it is used for buffering requests that is received by the controller of the Amazon. It is the component that is used for communication between different controllers. **3. Amazon SimpleDB:** it is used to store intermediate status log and the tasks that are performed by the user/ **4. Amazon EC2:** it is used to run a large distributed processing on the Hadoop cluster. It provides automatic parallelization and job scheduling. |
| **What are the uses of Amazon web services?** Amazon web services consist of a component called as Amazon S3 that acts as a input as well used as an output data store. It is used in checking the input and according to that gives the output. The input consists of the web that is stored on Amazon S3 as object and it is update frequently to make the changes in the whole architecture. It is required due to the on demand growing of the data set and to provide persistent storage. |
| How to use Amazon SQS?  Amazon SQS is a message passing mechanism that is used for communication between different connectors that are connected with each other. It also acts as a communicator between various components of Amazon. It keeps all the different functional components together. This functionality helps different components to be loosely coupled, and provide an architecture that is more failure resilient system. |
| How buffer is used in Amazon web services?  Buffer is used to make the system more resilient to burst of traffic or load by synchronizing different component. The components always receive and process the requests in unbalanced way. Buffer keeps the balance between different components and makes them work at the same speed to provide faster services. |
| What is the need of the feature isolation in Amazon web services?  Isolation provides a way to hide the architecture and gives an easy and convenient way to the user to use the services without any difficulty. When a message is passed between two controllers then a queue is maintained to keep the message. No controller calls any other controller directly. The communication takes place between the controllers by storing their messages in a queue. It is a service that provides a uniform way to transfer the messages between different application components. This way all the controllers are kept isolated from each other. |
| What is the function of a Amazon controller?  The functions that are involved with an Amazon controller are: - Controllers are used to control the flow in which the messages between the other system components has to be passed. - It controls the overall structure of the Amazon and all to retrieve the message, process the message, execute a function and store the message in other queue that are completely isolated from other controllers. - It manages and monitors the messages passed between the systems. |
| What is the function of Amazon Elastic Compute Cloud?  Amazon Elastic compute cloud is also known as Amazon EC2 is an Amazon web service that provides scalable resources and makes the computing easier for developers. The main functions of Amazon EC2 are: - It provides easy configurable options and allow user to configure the capacity. - It provides the complete control of computing resources and let the user run the computing environment according to his requirements. - It provides a fast way to run the instances and quickly book the system hence reducing the overall time. - It provides scalability to the resources and changes its environment according to the requirement of the user. - It provides varieties of tools to the developers to build failure resilient applications. |
| What are the different types of instances used in Amazon EC2?The instances that can be used in Amazon EC2 are: **1. Standard instances:** it provides small instances, large instances, extra large instances that give various configuration options from low range to very high range like Computing power unit, memory, processor, etc. **2. Micro Instances** It provides small consistent resources like CPU, memory and computing unit. It provides the resources to the applications that consume less amount of computing unit. **3. High memory instances** It provides large memory sizes for high end application and it includes memory caching applications as well. |
| What are cluster compute instances?  The cluster compute instances consist of the high CPU with network performance and are suited with high end applications. It provides network bound application and provide extra large computing resources like 23 GB memory, 33.5 EC2 compute units. It provide general purpose graphics unit to allow user with high end configuration. It also provide highly parallelized processing application that user can use and modify the server accordingly. |
| How to use SimpleDB with Amazon?  Every architecture rely on a database that is easy to maintain and gets easily configured Amazon uses the database by the name SimpleDB. This is the database that is used for cloud architecture to track the statuses of the components. The component of the system are asynchronous and discrete, it requires capturing the state of the system so that in any failure the user can easily revert back to its normal configuration. SimpleDB is schema-less database and there is no need to define the structure before the creation of any data. Every controller in the database defines their own structure and link the data to a job. |
| How does component services used for Amazon SimpleDB?  Component services allow the controllers to independently store the states of the virtual machines and the database that is in use. It creates asynchronous highly available services. It stores active requests according to the unique ID that are associated with each system. It stores the status of the entire database that is having different states for different components in a log database file. |
| How to upload files in Amazon S3?  Amazon S3 provides uploading of large files and retrieve small offsets for end-to-end transfer data rates. The large file gets stored into small files that are smaller in size. Amazon S3 stores multiple of files together in a bundle or in a compressed form for example in .gzip or .gz format and then convert them into Amazon S3 objects. The files get uploaded on the Amazon server by the use of FTP or another protocol and then retrieved through the HTTP GET request. The request includes the defined parameters like URL, offset (byte-range) and size (length). |
| What is the use of multi-threaded fetching in Amazon S3?  - Multi-threading fetching in Amazon S3 is used to fetch the objects concurrently using the multiple threads and map the task so that fetching can be made simpler. - It is not a good practice to increase the threading for a particular object as every node on the server has some bandwidth constraints. - It provides user the ease with which they can upload the files and upload the threads in parallel. - It provides high speed of data transfer and easy maintenance of the sever as well. |
| What is the difference between on demand and reserved instances?  - On demand instance allow user to pay for the computing capacity according to their use every hour, whereas reserved instances provide user to pay for every instance which they use and they want to reserve.  - On demand instance provide user a free working environment in which there is no need for too much of planning related to complexities, whereas reserved instances provide user with discounts on the hourly charge of an instance and provide a easy way to manage the instances as well.  - On demand instance provide maintenance of hardware and transforms fixed cost into much smaller variable costs, whereas reserved instance provide easy way to balance the pay package. |
| What are the provisions provided by Amazon Virtual Private cloud?  Amazon private cloud provides a provision to create a private and isolated networking infrastructure to give easily the Amazon web services. - Virtual network topologies define the traditional data-center approach to control and mange the files from one place. - It provides complete control over IP address range, creation of sub-nets and configuring the network gateways and route tables. - It provides easy to customize network configuration like creation of public sub-net to access the Internet easily. - It allow to create multiple security layers and provide network control list by which you can control the access to Amazon EC2 instances. |

|  |
| --- |
| What do you understand by MapReduce?  MapReduce is a software framework that was created by Google. It`s prime focus was to aid in distributed computing, specifically large sets of data on a group of many computers. The frameworks took its inspiration from the map and reduce functions from functional programming. |
| Explain how mapreduce works.  The processing can occur on data which are in a file system (unstructured ) or in a database ( structured ). The mapreduce framework primarily works on two steps: 1. Map step 2. Reduce step  **Map step:** During this step the master node accepts an input (problem) and splits it into smaller problems. Now the node distributes the small sub problems to the worker node so that they can solve the problem.  **Reduce step:** Once the sub problem is solved by the worker node, the node returns a solution to the master node which accepts all the solutions of the worker node and re-compiles them into a solution. This solution is for the input that was provided to the master node. |
| What is an input reader in reference to mapreduce?  The input reader as the name suggests primarily has two functions: 1. Reading the Input 2. Splitting it into sub-parts  The input reader accepts a user entered problem and then it divides/splits the problem into parts which then each are assigned a map function. Also an input reader will always read data from a stable storage source only to avoid problems. |
| Define the purpose of the Partition function in mapreduce framework  In mapreduce framework each map function generates key values. The partition function accepts these key values and in return provides the index for a reduce. Generally the key is hashed and a modulo is done to the number of reducers. |
| Explain combiners.  Combiners codes are used to increase the efficiency of a mapreduce process. They basically help by reducing the amount of data that needs to be shifted across to reducers. As a safe practice the mapreduce jobs should never depend upon combiners execution. |
| Explain what you understand by speculative execution  Mapreduce works on the basis of large number of computers connected via a network also known as node. In a large network there is always a possibility that a system may not perform as quickly as others. This results in a task being delayed. By speculative execution this can be avoided as multiple instances of the same map are run on different systems. |
| When do reducers play their role in a mapreduce task?  The reducers in a mapreduce job do not begin before all the map jobs are completed. Once all the map jobs are completed the reducers begin copying the intermediate key-value pairs from the mappers. Overall reducers start working as soon as the mappers are ready with key-value pairs. |
| How is mapreduce related to cloud computing?  The mapreduce framework contains most of the key architecture principles of cloud computing such as: **- Scale:** The framework is able to expand itself in direct proportion to the number of machines available. **- Reliable:** The framework is able to compensate for a lost node and restart the task on a different node. **- Affordable:** A user can start small and over time can add more hardware.  Due to the above features the mapreduce framework has become the platform of choice for the development of cloud applications. |
| How does fault tolerance work in mapreduce?  In a mapreduce job the master pings each worker periodically. In case a worker does not respond to that system then the system is marked as failed. Even completed tasks are rescheduled because the output was stored in a in a local disk of a worker which failed. Hence mapreduce is able to handle large-scale failures easily by simply restarting a task. The master node always saves itself at checkpoints and in case of any failure it simply restarts from that checkpoint. |
| In mapreduce what is a scarce system resource? Explain?  A scarce resource is one which is available in limited quantities for the system. In mapreduce the network band-with is a scarce resource. It is conserved by making use of local disks and memory in cluster to store data during tasks. The function uses the location of the input files into account and aims to schedule a task on a system which has the input files. |
| What are the various input and output types supported by mapreduce?  Mapreduce framework provides a user with many different output and input types. **Ex.** Each line is a key/value pair. The key is the offset of the line from the beginning of the file and the value are contents of the line. It is up-to the will of the user. Also a user can add functionality at his will to support new input and output types. |
| Explain task granularity  In mapreduce the map phase if subdivided into M pieces and the reduce phase into R pieces. Each worker is assigned a group of tasks this improves dynamic load balancing and also speeds up the recovery of a worker in case of failures. |
| With the help of two examples name the map and reduce function purpose  **Distributed grep:** A line is emitted by the map function if it matches a pattern. The reduce function is an identity function that copies supplied intermediate data for output.  **Term-vector per host:** In this the map function emits a hostname, vector pair for every document (input). The reduce function adds all the term vectors pairs generated and discards any infrequent terms. |
| Explain the general mapreduce algorithm:  The mapreduce algorithm has 4 main phases: 1. Map, 2. Combine, 3. Shuttle and sort 4. Phase output  Mappers simply execute on unsorted key/values pairs.They create the intermediate keys. Once these keys are ready the combiners pair the key/value pairs with the right key. The shuttle/sort is done by the framework their role being to group data and transfer it. Once completed, it will proceed for the output via the phase output process. |
| Write a short note on the disadvantages of Mapreduce  Some of the shortcomings of mapreduce are: - One-input two-phase data flow is rigid i.e. it does not allow for multiple step processing of records. - Being based on a procedural programming model this framework requires code for simple operations. - The map and reduce functions being opaque does not allow for optimization easily. |

### **1. Compare between Cloud and On-premise Computing.**

|  |  |  |
| --- | --- | --- |
| **Criteria** | **Cloud** | **On-premise** |
| Initial cost | Low | High |
| Maintenance and support | Hassle-free | Needs attention |
| Upgrade | Automatic | Manual |
| Scalability | Excellent | Good |
| Pay as you go | Yes | No |

### **2. What is a Cloud?**

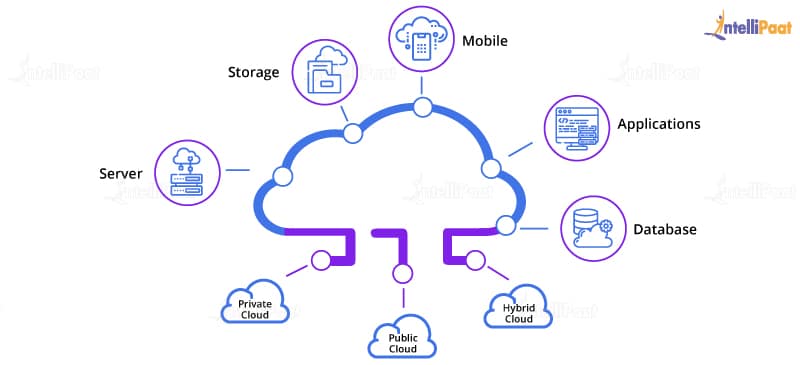
A cloud is a collaboration of networks, hardware, services, storage, and interfaces that help in delivering computing as a service. It has three users:

1. End users
2. Business management users
3. Cloud service providers

***Learn more about AWS from this insightful***[***AWS blog***](https://intellipaat.com/blog/what-is-amazon-web-services-aws/)***!***

### **3. What is Cloud Computing?**

It is an advanced-stage technology implemented so that the cloud provides its services globally as per the user requirements. It provides a method to access several servers worldwide.



### **4. What are some of the key features of Cloud Computing?**

* Reliable
* Scalable
* Agile
* Location Independent
* Multi-tenant

### **5. What are the benefits of Cloud Computing?**

The main benefits of Cloud Computing are:

* Data backup and storage of data
* Powerful server capabilities
* Incremented productivity
* Cost-effective and time-saving

### **6. Mention the layers of PaaS architecture.**

**Cloud Controller**

* Automatically creates virtual machines and controllers
* Deploys applications
* Connects to services
* Automatically scales up and down

**Storage Services**

* Object
* NoSQL
* Relational
* Block storage

**Applications Stored in Storage Services**

* Simple-to-scale applications
* Easier recovery from failure

### **7. What are the cloud service models?**

Infrastructure as a Service (IaaS)  
Platform as a Service (PaaS)  
Software as a Service (SaaS)

### **8. Write about Function as a Service.**

FaaS provides users with a fully functional platform where they can create, manage and run their applications without having to worry about maintaining the infrastructure.

Bottom of Form

### **9. What are the components of Windows Azure?**

Windows Azure Platform Services

* Cloud
* SQL Azure
* App Fabric: Allows fabric cloud

### **10. Which are the main constituents of the cloud ecosystem?**

* Cloud service providers
* Cloud consumers
* Direct consumers

### **11. Who are the cloud consumers in a cloud ecosystem?**

People and teams who use different types of cloud services, within your organization.

### **12. Who are the direct consumers in a cloud ecosystem?**

The individuals who utilize the service provided by your company, build within a cloud environment.

### **13. Who are the Cloud service providers in a cloud ecosystem?**

Cloud service providers are the companies that sell their cloud services to others. Sometimes these companies also provide cloud services internally to their partners, employees, etc.

### **14. What are the differences occurred in distributed operations?**

**FC**: Master–Slave operations  
**Nova**: Parallel process and its shared database

### **15. Which agent is equivalent of Nova Compute?**

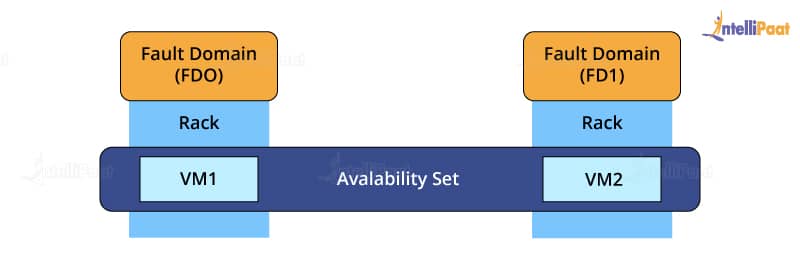
Azure Agent

**Learn more about AWS in this**[**AWS Training in New York**](https://intellipaat.com/aws-certification-training-online-new-york/)**to get ahead in your career!**

### **16. Mention the reliability and availability of Cloud Computing.**

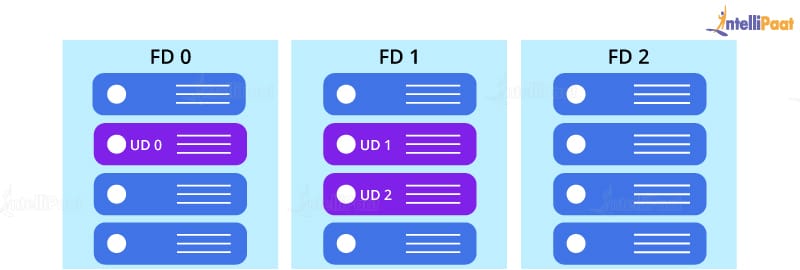
**Use of Fault Domains:**

* Two virtual machines are in a single fault domain if a single hardware can bring down both the virtual machines.
* Azure automatically distributes instances of a role across fault domains.



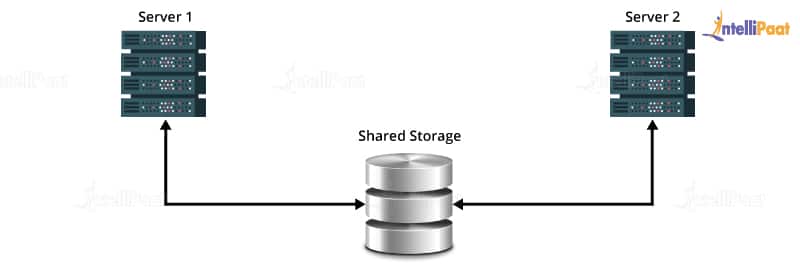
**Use of Upgrade Domains:**

* When a new version of the software is rolled out, then only one up-gradation of the domain is done at a time.
* It ensures that any instance of the service is always available.
* There is an availability of the applications in multiple instances.



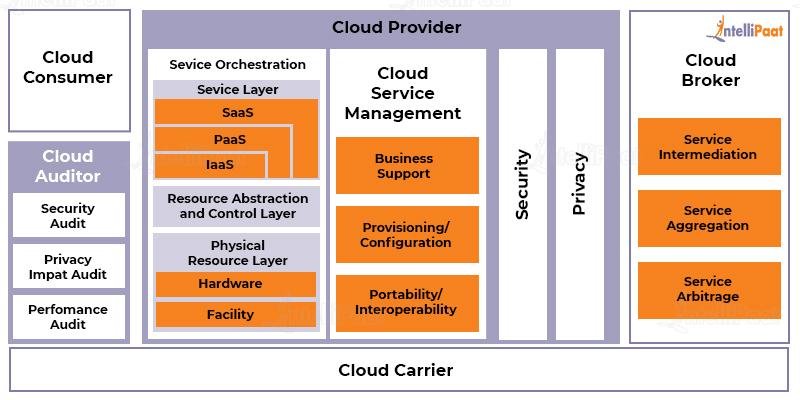
**Storage and Network Availability:**

* Copies of data are stored in different domains.
* it is a mechanism to guard against DoS and DDoS attacks.



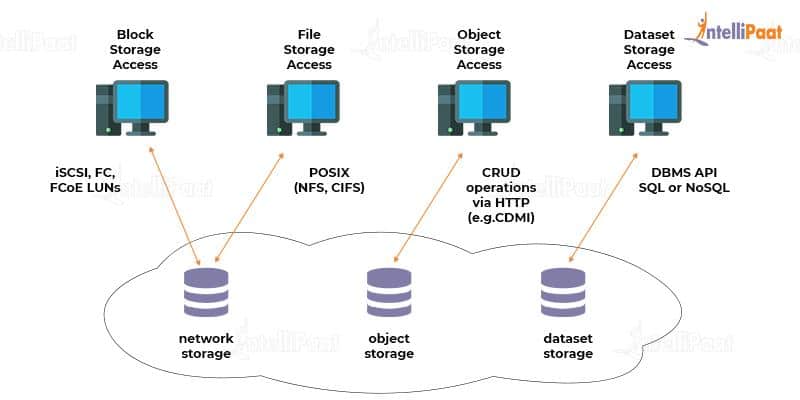
### **17. Explain the Cloud Computing Architecture.**

Cloud Computing Architecture brings together two components of cloud computing – the front-end and the back-end. It is important to bring the correct services together for the benefit of both internal and external people. If need be, the cloud management should be able to quickly make required changes.



### **18. What are the Cloud Storage Levels?**

* Files
* Blocks
* Datasets
* Objects



## Intermediate Interview Questions

### **19. What are the serverless components in cloud computing with their advantages and disadvantages?**

These components allow you to create apps without the stress of managing the infrastructure.

|  |  |
| --- | --- |
| **Advantages** | **Disadvantages** |
| Cost-effective | Can cause late responses |
| Increases productivity | Not ideal for high-computing operations |
| Scalable | More vulnerable when it comes to security |
| No server management | Debugging is challenging |

***To know more pros and cons of cloud computing check our blog on the***[***Advantages and disadvantages of cloud computing***](https://intellipaat.com/blog/tutorial/amazon-web-services-aws-tutorial/advantages-and-disadvantages-of-cloud-computing/)***now!***

### **20. Give the best example for the open-source Cloud Computing.**

OpenStack

### **21. What are microservices and their importance in the cloud?**

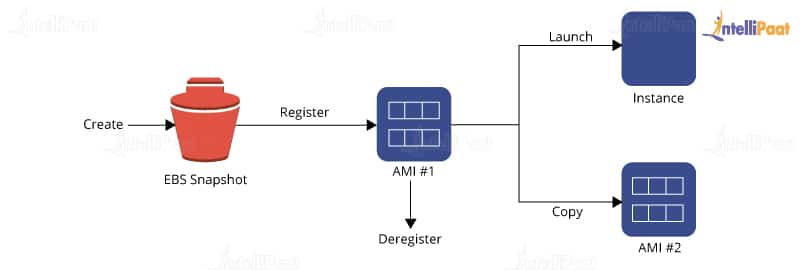
Microservices help create apps that consist of codes that are independent of one another and the platform they were developed on.  Microservices are important in the cloud because of the following reasons:

* Each of them is built for a particular purpose. This makes app development simpler.
* They make changes easier and quicker.
* Their scalability makes it easier to adapt the service as needed.

### **22. What is an AMI? How do we implement it?**

AMI is [Amazon Machine Image](https://intellipaat.com/blog/what-is-amazon-machine-image/), which basically is a copy of your root file system. It feeds the information required to launch an instance.

We implement AMI by specifying an AMI whenever we want to launch an instance. Multiple instances can be launched from a single AMI with the same configuration.  
In the case of launching instances with different configurations, we would need to launch different AMIs.



AMI includes one or more snapshots of your EBS volumes, in the case of instance-store backed AMIs, along with a template for the root volume of your instance (like an operating system, an application server, and applications).

It launches the permissions that decide which AWS accounts can use the AMI for launching instances. It also needs a block device mapping for specifying the volumes in order to attach them to the instances whenever they are launched.

**Interested in learning AWS? Check out our**[***AWS Training in Sydney***](https://intellipaat.com/aws-certification-training-online-sydney/)**!**

### **23. Why Hybrid Clouds are so important?**

**Cloud Bursting:**  
Access capacity and specialized software are available in the public cloud and not in the private cloud.  
Examples: Virtual Amazon and [Dynamo](https://intellipaat.com/blog/amazon-aws-dynamodb-tutorial/)

Leveraging the best of both worlds, there are hybrid clouds.

**vCloud**:

* It is a VMware cloud.
* It is expensive.
* It gives enterprise quality.

**OpenStack:**

* It has commodity servers and storage.
* It is less reliable.
* We can run web servers on OpenStack.
* the database is built on vCloud.

### **24. List the platforms that are used for large-scale Cloud Computing.**

The platforms that are used for large-scale Cloud Computing are:

* [Apache Hadoop](https://intellipaat.com/blog/tutorial/hadoop-tutorial/)
* [MapReduce](https://intellipaat.com/blog/what-is-mapreduce/)

### **25. Mention the different types of models used for the deployment in Cloud Computing.**

The different deployment models in Cloud Computing are:

* Private Cloud
* Public Cloud
* Community Cloud
* Hybrid Cloud

***Become a master of AWS by enrolling in this online***[***AWS Course in Toronto***](https://intellipaat.com/aws-certification-training-online-toronto/)***!***

### **26. Explain security management in terms of Cloud Computing.**

* **Identity management** access provides the authorization of application services.
* **Access control** permission is given to users to have complete controlling access of another user who is entering into the cloud environment.
* **Authentication and authorization** provide permission to only the authorized and authenticated users to access the data and applications.

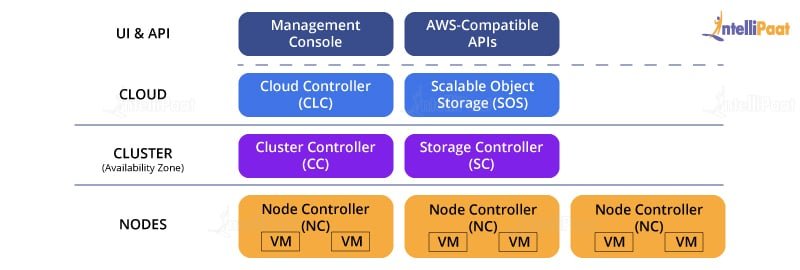
### **27. Which are the layers of Cloud Computing?**

The different layers used by [cloud architecture](https://intellipaat.com/blog/cloud-computing-architecture/) are as follows:

* CLC or Cloud Controller
* Walrus
* Cluster Controller
* Storage Controller (SC)
* Node Controller (NC)

### **28. Explain the full form and the usage of 'Eucalyptus' in Cloud Computing.**

The full form of ‘Eucalyptus’ is ‘Elastic Utility Computing Architecture for Linking Your Programs to Useful Systems.



Eucalyptus is an open-source software infrastructure in Cloud Computing, which enables us to implement clusters in the Cloud Computing platform. It is mainly used to build public, hybrid, and private clouds. It has capabilities to convert our own data center into a private cloud and provides its functionalities for various other organizations.

***Get certified from the top***[***AWS Course in Singapore***](https://intellipaat.com/aws-certification-training-online-singapore/)***now!***

### **29. Mention the names of some large cloud providers and databases.**

* Google Bigtable
* Amazon Simple Database
* Cloud-based SQL (Sequential Query Language)

**30. Explain the difference between cloud and traditional data centers.**

* The traditional data center is expensive due to heating and hardware/software issues. Mostly, the expenditure is on the maintenance of the data centers.
* Cloud is scaled up when there is an increase in demand, hence such expenditure issues are not faced in Cloud Computing.

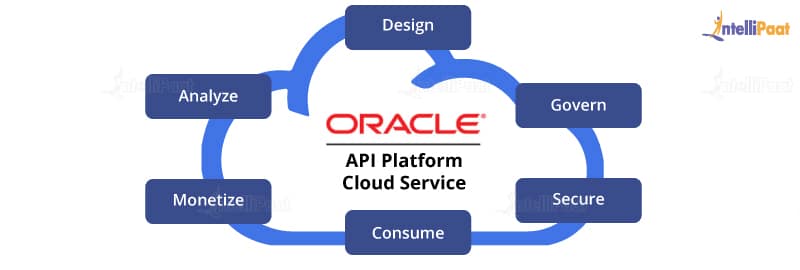
***Learn more about AWS in***[***AWS Tutorial***](https://intellipaat.com/blog/tutorial/amazon-web-services-aws-tutorial/)***!***

### **31. What is meant by Edge Computing?**

Edge computing is a part of the distributed computing structure. It brings companies closer to the sources of data. This benefits businesses by giving them better insights, good response time and better bandwidth.

## Advanced Interview Questions

### **32. What are the uses of APIs in cloud services?**



* APIs (Application Programming Interfaces) are used to eliminate the necessity to write complete programs.
* Here, instructions are provided to make communication between one or more applications.
* Creation of applications is made easy and accessible for the link of cloud services with other systems.

### **33. Mention different data center deployments of Cloud Computing.**

Cloud Computing consists of different data centers as follows:

* **Containerized data centers:**Containerized data centers are the packages that contain a consistent set of servers, network components, and storage delivered to large warehouse kind of facilities. Here, each deployment is relatively unique.
* **Low-density data centers:** Containerized data centers promote heavy density which in turn causes much heat and significant engineering troubles. **Low-density data centers** are the solution to this problem. Here, the equipment is established far apart so that it cools down the generated heat.

***Learn the complete concepts of***[***AWS at Hyderabad***](https://intellipaat.com/aws-certification-training-online-hyderabad/)***in 26 hours!***

### **34. List down the three basic functioning clouds in Cloud Computing.**

* Professional cloud
* Personal cloud
* Performance cloud

### **35. What are the characteristics of cloud architecture that differ from traditional cloud architecture?**

The characteristics are:

* In cloud, the hardware requirement is fulfilled as per the demand created for cloud architecture.
* Cloud architecture is capable of scaling up resources when there is a demand.
* Cloud architecture is capable of managing and handling dynamic workloads without any point of failure.

### **36. What are the building blocks of cloud architecture?**

* Reference architecture
* Technical architecture
* Deployment operation architecture

### **37. Explain AWS.**

AWS stands for[Amazon Web Services](https://intellipaat.com/blog/tutorial/amazon-web-services-aws-tutorial/aws-introduction/) which is a collection of remote computing services also known as Cloud Computing. This technology is also known as IaaS or Infrastructure as a Service.

### **38. Mention the key components of AWS.**

The key components of AWS are as follows:

* **AWS Route 53:** [AWS Route 53](https://intellipaat.com/blog/what-is-aws-route53/) is a DNS (Domain Name Server) web-based service platform.
* **Simple E-mail Service:** Sending of e-mail is done by using a RESTFUL API call or via regular SMTP (Simple Mail Transfer Protocol).
* **Identity and Access Management:** Improvised security and identity management are provided for an AWS account.
* **Simple Storage Device (S3):** It is a huge storage medium, widely used for AWS services.
* **Elastic Compute Cloud (EC2):** It allows on-demand computing resources for hosting applications and is essentially useful for unpredictable workloads.
* **Elastic Block Stores (EBS):** They are storage volumes attached to EC2 and allow the data lifespan of a single EC2.
* **CloudWatch:** [**Amazon CloudWatch**](https://intellipaat.com/blog/what-is-cloudwatch-in-aws/) is used to monitor AWS resources, and it allows administrators to view and collect the keys required. Access is provided so that one can set a notification alarm in the case of trouble.

### **39. Explain how you can vertically scale an Amazon instance.**

This is one of the essential features of AWS and cloud virtualization. We spinup a newly developed large instance where we pause that instance and detach the root EBS volume from the server and discard. Later, we stop our live instance and detach its root volume connected. here, we note down the unique device ID and attach the same root volume to the new server, and we restart it again. This results in a vertically scaled Amazon instance.

### **40. Explain the security usage in the Amazon Web Services model.**

* AWS supports security groups.
* Access is provided to create a security group for a jump box with SSH access only for port 22 open. Later, a webserver group and a database group are created. The webserver group provides 80 and 443 from around the world, but only port 22 will be vital among the jump box group. The database group allows port 3306 from the webserver group and port 22 from the jump box group. The addition of any machines to the webserver group can store in the database. No one can directly SSH to any of our boxes.

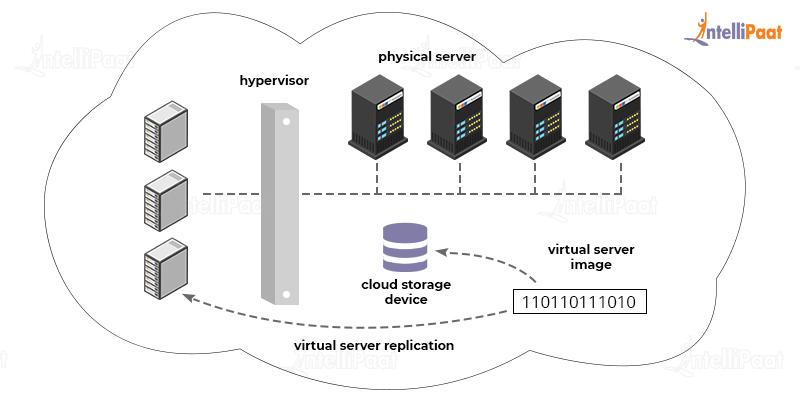
### **41. What are reasons that make Amazon so big?**

* In Amazon, the backup storage of EBS volumes is maintained by inserting the snapshot facility via an API call or via a GUI interface like Elasticfox.
* Performance is improved by using Linux software RAID and striping across four volumes.

Hope that you will find these interviews questions for the role of Cloud Engineer useful. Do let us know in the comments section below!

### **42. How does Resource Replication take place in Cloud Computing?**

Resource Replication creates duplicates of the same resource. Replication is employed when a resource is needed more and more. The resource is virtualized to replicate cloud-based resources.



### **43. What is Containers as a Service (CaaS)?**

CaaS is a system that allows developers to run, scale, manage, upload, and organize containers by using virtualization.

A container is a software pack. It allows teams to scale their apps to highly available cloud infrastructures.

Hope that you will find these interview questions for the role of Cloud Engineer useful. Do let us know in the comments section below!

[**Next**](https://intellipaat.com/blog/interview-question/aws-lambda-interview-questions/)