

By OnlineInterviewQuestions.com

Amazon DevOps Engineer Interview Questions

Amazon is an electronic commerce and cloud computing company. Also, this is a huge online retailer platform which can offer any kind of products to the customers. The goods and products can be ordered via online where the dealer will deliver it to the customers. If you are looking for making the career in **DevOps** then you can apply for the position of **Amazon DevOps Engineer**. Being one of the giants in the e-com industry every developer wishes to be a part of it, therefore, an engineer must explore all the possible **Amazon DevOps Interview Questions** before facing the panelists.

Here are a few Amazon [DevOps Interview Questions](#) that you must know before you face the interview panel at Amazon:

Q1. What is iNode on Linux and more details on that?

The iNode in Linux is an entry table containing information about the regular file and directory. It can be viewed as a data structure that contains the metadata about the files. The following are the contents of the iNode.

- **User ID** - Owner of the file.
- **Group ID** - Owner of the group.
- **Size of File** - a major or minor number in some files.
- **Timestamp** - access time, and modification time.
- **Attributes** - some properties of the file.
- **Access control list** - permission for users.
- **Link count** - The number of hard links relative to the inode.
- **File type** - Type of the file such as regular, directory, or pipe.

Link to the location of the file and other metadata.

Q2. How do you check physical memory on Linux?

There are a number of ways to check for physical memory size on Linux. The popular ways are,

Free command: Type the free command to check the physical memory size.

```
free -b //gives the size in bytes
free -k //gives the size in kilobytes
free -m //gives the size in megabytes
free -g //gives the size in gigabytes
```

Top command: The top command lists the physical memory information in a clear way.

Vmstat command: The vmstat (virtual memory stats) command with -s switch lists the memory in detail.

Q3. How do you copy a local file to HDFS?

The fs put command is used to copy or upload a file from the local filesystem to the specific HDFS.

Syntax

```
fs put --from source_path_and_file --to dest_path_and_file
```

Q4. What steps happen when a copy of the local file to HDFS run?

The steps in copying a local file to the HDFS.

1. Create a new file in the local filesystem with any name such as test.txt in the folder /home/user/.
2. Add some content to the file test.txt: echo "hello" > /home/user/test.txt.
3. Create a new directory in the HDFS using the command: `hadoop fs -mkdir /user/in`
4. Finally, copy the local file to the HDFS using the command: `hadoop fs -copyFromLocal /home/user/test.txt /user/in/test.txt`

Q5. Explain the steps/channels When the user accesses a URL via a browser?

The steps in accessing a URL through the browser are,

1. Enter the URL into a web browser
2. The browser looks at the IP address for the entered domain using the DNS.
3. The browser then sends the HTTP request to the server.
4. The server sends back an HTTP response.
5. With the request, the browser begins rendering the HTML.
6. Then the browser sends the requests for additional objects embedded in the HTML such as images, CSS, JavaScript, etc.
7. Finally, the page gets loaded. The browser may send additional async requests if needed.

Q6. What is distributed cache in Hadoop?

A **distributed cache** is a facility in the Hadoop MapReduce framework that is used to cache files. These cached files are frequently needed by the applications. It is used to cache read-only files such as text files, archives, jar files, etc. Hadoop will use the cached file for the job on each data node where the map/reduce tasks are running. When the task is completed, these files are deleted from the data nodes. A distributed cache is used because we

can send a copy of the file to all data nodes instead of reading the file every time from the HDFS.

Q7. How do you measure network packet?

The network performance of a packet is measured using various factors such as,

- **Latency** - It is the amount of time that takes for the data to travel from one location to another.
- **Packet Loss** - It is the number of packets transmitted from one location to another that fails to transmit.
- **Throughput** - It is the number of items passing through a particular system.
- **Bandwidth** - It is the amount of data that can be transferred over a given period of time.
- **Jitter** - It is defined as the variation in time delay for the data packets that are sent over a network.

Q8. If there are no cookies how do you make your application work?

The application can make use of the session ID tag to be used for creating sessions in the applications without the need for the cookies. Using the session ID, the application can create individual sessions for users without using cookies.

Q9. How did you make your application work without cookies?

The application can work and create sessions for users without the need for cookies. It has to use the session ID tag to create individual sessions for the users.

Q10. How do you make your database work for a high number of requests at peak hours?

To make the database perform higher, here are the following things that we can do.

- **CPU** - Increase the number of cores of the CPU to keep the host responsive.
- **Memory** - Look at the page faults per second in the memory and keep it low.
- **Disk space** - Make sure that you have a high amount of disk space.
- **Database connections** - Make sure that you have enough database connections.

Q11. How do you make a background process run?

To run a background process, use the following command.

\$command & - It will make your command run in the background but will be killed if the session is closed.

nohup <command><args> ><filename> 2>&1 & - It will make your command running in the background even if the session is closed.

Q12. How do you optimize database?

To fine-tune the performance and optimize the database, execute the following steps.

- **Use Indexing** - Index is a data structure that increases the speed of the data retrieval operations.
- **Execution plans** - Execution plan tool in the SQL server is useful in creating indexes.
- **Avoid coding loops** - When possible avoid the loops in your code to increase the performance of the database.
- **Avoid correlated SQL subqueries** - A correlated subquery gets values from the parent query. It decreases the performance of the database operations. So try to avoid it. Finally, Use or avoid temporary tables according to your specific requirements.

Q13. How do you know how much memory your java application is taking on Linux?

You can know how much memory your java application is taking on Linux using

- The **jstat -gccapacity [pid]** will give you information about the memory pool generation and space capabilities.
- The **jstat -gcutil [pid]** will give you the information about the utilization of each generation as a percentage of its capacity.

The PID is the unique process identifier that is associated with each java process.

Q14. Tell me about garbage collection programming

Garbage collection is the collection or gaining the memory back from the objects. The memory collected are not in use at the moment in any part of the program where the object is used. This process frees up the memory space that is no longer used by the objects and such. This process is implemented differently in different languages. Most of the high-level programming languages have garbage collection process built into it. Low-level programming languages add garbage collection processes through external libraries. For eg: In C programming language, the garbage collection is taken care of by the user by using the malloc() and dealloc() functions. In C# programming language, the garbage collection is taken care of automatically. Users don't need to do anything.

Please Visit OnlineInterviewquestions.com to download more pdfs