**L1**

1. Which command is used to create a new container from an image and start it?

Ans: docker run –it –name Rohith ubuntu /bin/bash

1. What is the purpose of the docker pull command in docker?

Ans: To download docker images from internet

1. Which docker command is used to list all running containers

Ans: docker ps or docker ls

1. What is the purpose of docker stop command in docker?

Ans: to stop the docker container gracefully

1. What is the purpose of docker compose up command in docker compose?

Ans: To start and run an entire app on a standalone host that contains multiple services

1. What is ansible primarily used for in the context of it and Devops?

Ans: Automation and Configuration management

1. What is role of ansible in IT?

Ans: Reuse of collections of tasks and configuration management

1. What is the file used in ansible to define tasks, place and configuration?

Ans: playbook.yaml,or ansible.cfg

1. What is the purpose of playbook.yaml?

Ans: to define series of tasks and configuration file

1. Which module in ansible is used to install and manage the packages?

Ans: Package

1. What is the primary role of CI/CD?

Ans: Automation,**to continuous integration and testing,** CI/CD allows organizations to ship software quickly and efficiently

1. What is the key advantage of IAAS in CI/CD?

Ans: The key advantage of IaaS in CI/CD lies in its ability to provide a scalable, flexible, and automated infrastructure that supports the dynamic and rapid nature of modern software development and deployment processes.

1. What is the purpose of /bin/bash

Ans: tells your terminal that when you run the script it should use bash to execute it and bash contains in bin folder , **to indicate the interprator to run the script.**

1. Which command is used to give permission to execute in linux?

Ans: chmod + x

1. What is purpose of if statements in unix scripting?

Ans: to perform conditional statements in our code

1. What is command is used to retrieve data from database?

Ans: select statement

1. What is select \* from employees?

Ans: to retrieve data from all columns and rows

1. Which sql command is used to filter on specific condition?

Ans: where command

1. What is group by command in sql?

Ans: to group same rows that have Same values

1. What is the purpose of primary key?

Ans: To ensure data in the specific column is unique

1. Sql command to change existing data?

Ans: Update

1. Which is not the v Of 3v’s

Ans: **Validity**, Volume,velocity,variety

1. Which type of type of data is structured data?

Ans: Structured data is information that has been organized into a specific format and stored in an easily retrievable form.

1. Which python library is used for large dataset and manipulating the dataset?

Ans: Pandas

1. In python what is the most common way to schedule and automate the data processing tasks?

Ans: Crone Jobs

------------

1. Every organization needs?

Ans: faster delivery,higher quality,lesser spending,availability

1. POM ful form?

Ans: project object model

1. What is data warehouse?

Ans: A data warehouse is a central repository of information that can be analysed to make more informed decisions.

1. Datawarehouse is?

Ans: subject-oriented,integrated,time variant,nonvolatile

1. What are the source data components in data warehouse?

Ans: production data,archive data,external data,internal data.

1. ETL stands for?

Ans: extract ,transform and load

1. Red hat family?

Ans: centos, amazon linux, open SUSE,mandrake,fedora

1. Debian family?

Ans: ubuntu,lindows,knoppix

1. YUM stands for?

Ans: yellowdog updated modified,it is warehouse of linux softwares

1. Modules of Big data?

Ans: HDFS(Hadoop distributed file system),YARN(Yet another resource negotiater),Map Reduce,Hadoop Common

1. HDFS ?

Ans: HDFS is the distributed file system that stores data across multiple machines. It breaks large files into smaller blocks and distributes them across the cluster for parallel processing.

1. YARN?

Ans: YARN is a resource management layer that manages and schedules resources in a Hadoop cluster. It allows multiple applications to share resources efficiently.

1. Hadoop common?

Ans: Hadoop Common includes libraries and utilities that support the other Hadoop modules. It provides common functionality like input/output formats, serialization, and various other utilities used by Hadoop components.

1. Map Reduce?

Ans: This is a framework which helps Java programs to do the parallel computation on data using key value pair. This is where the processing takes place under key value pairs

1. Hadoop Common?

Ans: it is a library which contains all the resources for the processing of hadoop cluster

1. What is Apache spark?

Ans: Apache Spark is an open-source, distributed computing system designed for fast and flexible large-scale data processing. It provides an interface for programming entire clusters with implicit data parallelism and fault tolerance. Spark's key features include in-memory processing, support for a wide range of data processing tasks, and compatibility with various programming languages.

1. What is Apache Pyspark?

Ans: Processing of big data using python libraries for fast data processing and minimum data tolerance

1. RDD?

Ans: Resilient Distributed Dataset,it provides fault tolerance using lineage mechanism,this linege mechanism keeps track of transformation and allows us to replicate the data of lost nodes and contains a copy of it so there will be least amount data will be lost,it is the main component in apache pyspark,

1. What is linux?

Ans: linux is an operating system which acts as an interface between user and the hardware

1. Unix family?

Ans: linux,Mac Os,Solaris,HP-UX,AIX

1. Linux Family?

Ans: Centos,rad hat,ubuntu,fedora,amazon linux,linuxmint,open suse

1. What is mean by la in ls -la in unix?

Ans: list information of every file (l)and also which are hidden(a)

1. Touch in linux?

Ans: to create blank file

1. Pwd full form and use?

Ans: Print Working Directory,Shows/prints the current working directory

1. Diff between nano and VI?

Ans: both are text editors but nano is more user friendly and easy to use for beginners but vi is most enhanced version with powerful editing capabilities

1. Mv command in linux?

Ans: move (mv) command is used to move the file and also to rename the file

1. Rm -rf in linux?

Ans: rm stands for remove ,-r means recursive which means delete all the directories and its sub directories and -f mean forcefully without asking permissions

1. Grep command?

Ans: global search for regular expression and print output,it is command which is used to search for pattern of characters in the files and print it

grep -c = count in the files

grep -h = display the matched files but not the file names

grep -i = ignores case sensitive for matching

grep -l = displays list of filenames only

grep -n = display the matched lines and their line numbers

grep -v = printout all the lines that do not match the pattern

1. Less command ?

Ans: it outputs the portion of file from the large files

1. Diff between yum and wget?

Ans: yum is used to install and wget is used to download

1. Ifconfig?

Ans: to get ipaddress

1. find / -type d -name <dir\_name> ?

ans: to find a specific directory under top level root directory

1. Adding single/multiple users to group?

gpasswd -a/-M <user1><user2> <groupname>

1. cat,cat>,cat>>?

Ans: cat is to display the file, cat > is create a file , cat >> is used to edit the file

1. To copy all files with extension txt to mydir directory \_\_\_ command is used, if mydir is parent directory of current directory.?

Ans: cp \*.txt.. /mydir

1. What is Webserver?

Ans: A Web server is a program that uses HTTP (Hypertext Transfer Protocol) to serve the files that from Web pages to users, in response to their request.

1. What is httpd?

Ans: HTTP Daemon is a software program that runs in the background as a web server.

It waits for the incoming server requests. The daemon answers the request and serves the hypertext and multimedia documents over the Internet using HTTP.

1. What is a load balancer?

Ans: A load balancer accepts incoming traffic from clients and routes requests to EC2 instances. The load balancer also monitors the health of its registered targets and ensures that it routes traffic only to healthy targets.

When the load balancer detects an unhealthy target, it stops routing traffic to that target. It then resumes routing traffic to that target when it detects that the target is healthy again.

1. What is auto scaling?

Ans: Autoscaling, also spelled auto scaling or auto-scaling, and sometimes also called automatic scaling, is a method used in cloud computing that dynamically adjusts the amount of computational resources in a server farm - typically measured by the number of active servers - automatically based on the load on the farm.

1. What is s3?

Ans: Simple Storage Service,it is secure, durable and highly available storage service for objects storing,we can store and retrieve any amount of data from anywhere in the web

1. S3 Buckets?

Ans: A bucket is a container for storing objects (files) in Amazon S3. Each object is contained in a bucket

1. What is Roles and Policies in S3?

Ans: Policies are like rules that define what actions someone or something is allowed to do in AWS. Roles are like predefined sets of permissions that can be assumed by users, services, or resources when they need specific access

1. What is Application Deployment?

Ans: the process of deploying an application over an executable environment is known application deployment.

1. What is Virtulization?

Ans: Virtulization is technology where you make the virtual representations of the servers,storage,database and other physical machines,it mimics the physical machine and creates multiple virtual machines on a single physical machine

1. Advantages of virtulization?

Ans: Multiple os in same machine,Easy maintenance and recovery,lower cost to build

1. Disadvantages of virtulization?

Ans: multiple vms leads to unstable performance,hypervisor is not so efficient,long boot up process,each VM requires cpu allocation,RAM. More Vms means more resources, Application portability not guaranteed

1. What is microservices?

Ans: breaks the application into single operation,makes the app independently scalable,highly available and stateless

1. What is Container?

Ans: A container is standard unit of software that packages the code and all its dependencies to run application quickly and reliably from one computing environment to another

**Docker enables you to separate your applications from your infrastructure so you can deliver software quickly.**

1. What is Docker image?

Ans: A docker image is a lightweight, standalone, executable package of software that includes everything to run an application like tools, libraries, code, runtime and settings

1. Docker containers that run on Docker Engine?

Standard: Docker created the industry standard for containers, so they could be portable anywhere

Lightweight: Containers share the machine’s OS system kernel and therefore do not require an OS per application, driving higher server efficiencies and reducing server and licensing costs

Secure: Applications are safer in containers and Docker provides the strongest default isolation capabilities in the industry

1. What is Docker Daemon?

Ans: The Docker daemon is a background process that manages Docker containers on a system. It is responsible for building, running, and managing containers. docker build ,docker ps,docker run

1. What is docker client?

Ans: The Docker client (docker) is the primary way that many Docker users interact with Docker. When you use commands such as docker run, the client sends these commands to dockerdaemon, which carries them out. The docker command uses the Docker API. The Docker client can communicate with more than one daemon.

1. Docker ps -a?

Ans: all containers present in the docker

1. Docker create -it –name sample ubudntu /bin/bash?

Ans: to create a docker

1. Docker images?

Ans: shows all images.

1. What is docker log?

Ans: Docker provides a command called docker logs that allows you to retrieve and view the logs of a specific container. You can use this command to see what's happening inside a container, identify any issues, or monitor the application's activity.

1. What is -it in docker?

Ans: interactive terminal

1. What is docker attach command in docker?

Ans: to get into the docker we use docker attach command

1. To push the repo into the sonartype nexus artifactory what is command is used?

Ans: mvn clean deploy

1. To check the issues in sonarcube from maven what command is used?

Ans: mvn clean verify sonar:sonar

1. To build the package in maven what command is used ?

Ans: mvn clean package

1. What is docker file?

Ans: Docker can build images automatically by reading the instructions from a Dockerfile. A Dockerfile is a text document that contains all the commands a user could call on the command line to assemble an image. This page describes the commands you can use in a Dockerfile

1. Where does the volumes are stored outside the union file system?

Ans: /var/lib/docker/volumes

1. What is etcd?

Ans: . All the cluster, pods, replicasets, deployments information would be

stored in distributed database

1. What is kube api server?

Ans: API server is used to carry the commands from kubectl (command

prompt) to kubernetes cluster

1. What is scheduler?

Ans: 1. Scheduler is responsible scheduling the pods unto the nodes

2. When a new pod is getting created, on which node this pod has to be

created would be done by scheduler

3. How much CPU is available? Are there any port conflicts? Scheduler

considers all these facts before it schedules a pod on a node

1. What is control manager?

Ans: 1. kubectl manager is responsible for overall health of the cluster

2. It make sures the actual state of kubernetes cluster matches to the

desired state

1. What is kubelet?

Ans: 1. The job of kubelet is to make sure what is happening on the node and

communicates it back to the master node

2. If a pod goes down, kubelet communicates to master nod

1. What is kube-proxy?

Ans: 1. Creating deployments and exposing the deployments as service is the job of networking component 2. It helps you in exposing services around your nodes and pods

1. What is container runtime?

Ans: docker is nothing but container runtime, the container runtime is responsible for running containers within a Pod. A container runtime is the software that is responsible for pulling container images from a registry, creating container instances, and managing their lifecycle.

1. kubectl get pods -o wide is used for?

Ans: The kubectl get pods -o wide command is used to retrieve information about pods in a Kubernetes cluster, and the -o wide option is used to display additional details in a wide format

1. kubectl run podname –image ubuntu –port=80?

Ans: creating a pod with image of ubuntu with porty number 80

1. kubectl create deployment name –image=ubuntu

Ans: creating Kubernetes deployment with image of ubuntu

1. kubectl scale deployment name –replicas=5?

Ans: creates replicas of the deployment

1. kubectl expose deployment name –type=loadbalancer –port=8080

Ans: creating a load balancer so that we can connect externally

1. kubectl get events --sort-by=.metadata.creationTimestamp is used for?

Ans: to know how scaling is happened