Subject: Big Data Engineering (DJ19DSL604)

AY: 2022-23

Experiment 2

(Infrastructure Implementation)

Name: Sarvagya Singh

SAPID: 60009200030 BATCH: K1

Aim: Hands on experience of Dockers and Kubernetes.

Theory:

Docker Overview

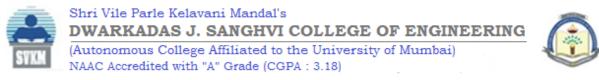
Docker is an open platform for developing, shipping, and running applications. Docker enables you to separate your applications from your infrastructure so you can deliver software quickly. With Docker, you can manage your infrastructure in the same ways you manage your applications. By taking advantage of Docker's methodologies for shipping, testing, and deploying code quickly, you can significantly reduce the delay between writing code and running it in production.

The Docker Platform

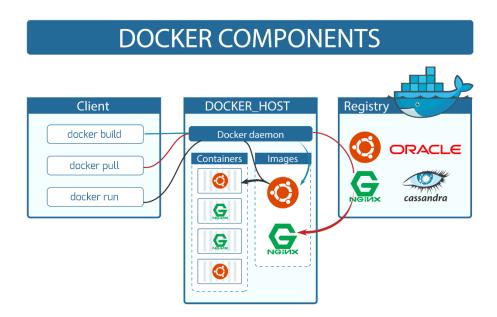
Docker provides the ability to package and run an application in a loosely isolated environment called a container. The isolation and security allows you to run many containers simultaneously on a given host. Containers are lightweight and contain everything needed to run the application, so you do not need to rely on what is currently installed on the host. You can easily share containers while you work, and be sure that everyone you share with gets the same container that works in the same way.

Docker provides tooling and a platform to manage the lifecycle of your containers:

• Develop your application and its supporting components using containers.



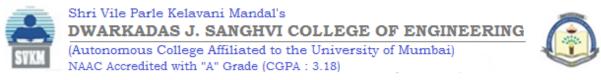
- The container becomes the unit for distributing and testing your application.
- When you're ready, deploy your application into your production environment, as a container or an orchestrated service. This works the same whether your production environment is a local data center, a cloud provider, or a hybrid of the two.

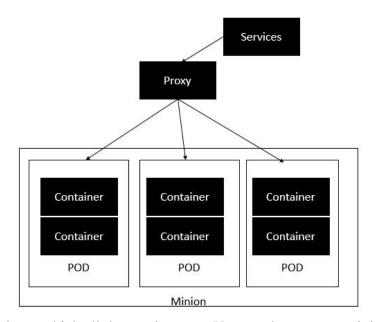


Kubernetes

Kubernetes is an orchestration framework for Docker containers which helps expose containers as services to the outside world. For example, you can have two services — One service would contain nginx and mongoDB, and another service would contain nginx and redis. Each service can have an IP or service point which can be connected by other applications. Kubernetes is then used to manage these services.

The following diagram shows in a simplistic format how Kubernetes works from an architecture point of view.





The minion is the node on which all the services run. You can have many minions running at one point in time. Each minion will host one or more POD. Each POD is like hosting a service. Each POD then contains the Docker containers. Each POD can host a different set of Docker containers. The proxy is then used to control the exposing of these services to the outside world.

Kubernetes has several components in its architecture. The role of each component is explained below &mius;

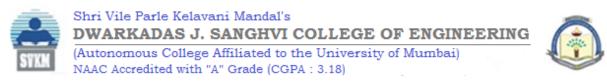
etcd – This component is a highly available key-value store that is used for storing shared configuration and service discovery. Here the various applications will be able to connect to the services via the discovery service.

Flannel – This is a backend network which is required for the containers.

kube-apiserver – This is an API which can be used to orchestrate the Docker containers.

kube-controller-manager – This is used to control the Kubernetes services.

kube-scheduler – This is used to schedule the containers on hosts.



Kubelet – This is used to control the launching of containers via manifest files.

kube-proxy – This is used to provide network proxy services to the outside world.

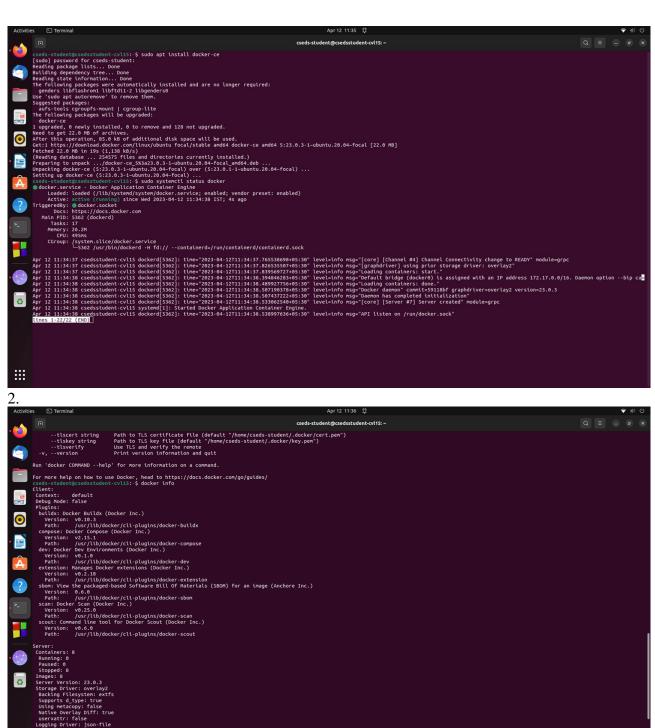
Lab Assignment:

- 1. Installation of Docker and Kubernetes on Ubuntu 20.4.5.
- 2. Group Activity (3 students): Create two Docker containers. One with an OS image and another with Node.js and React.
- 3. Develop a website in one Docker Container, share is with your team mates and allow parallel corrections from all team members.
- 4. Connect both the Dockers through Kubernetes and show how new request is handled.

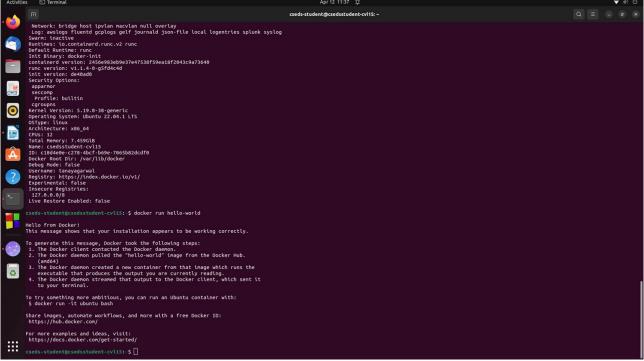
EXPERIMENT: 2

Name: Sarvagya Singh SAPID: 60009200030

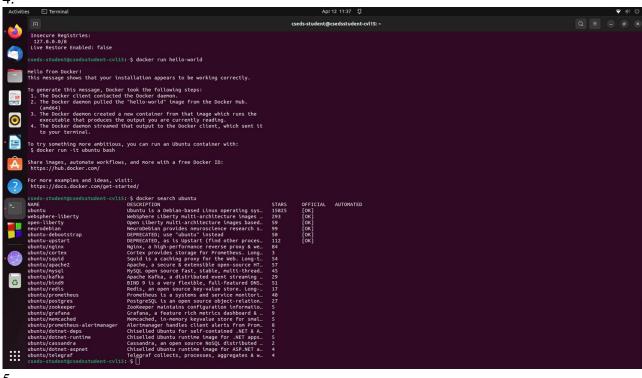
Dockers:

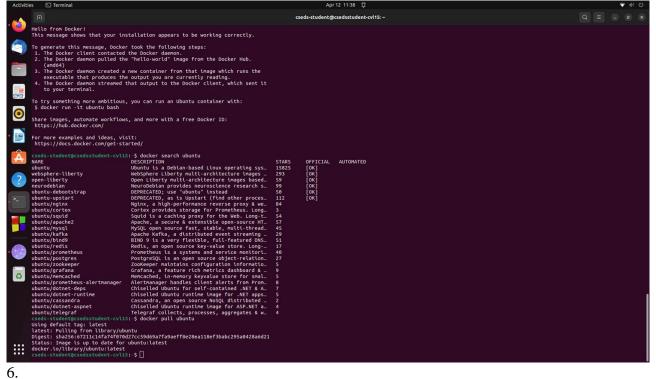


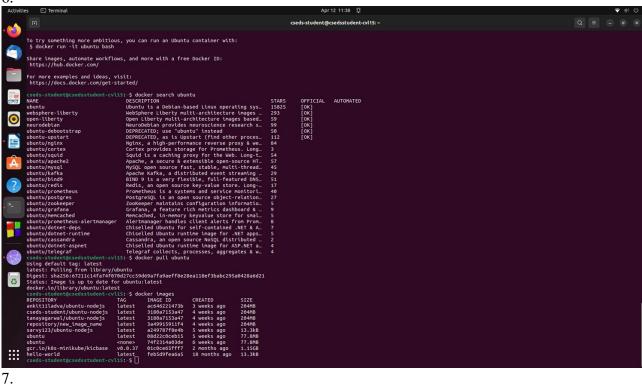
s: er local rk: bridge host ipvlan macvlan null overlay amslogs fluentd gcplogs gelf journald json-file local logentries splunk syslog

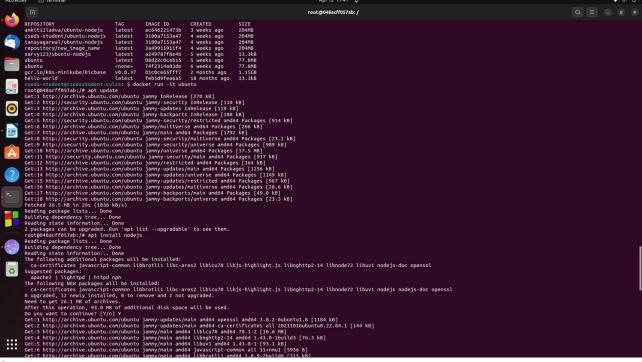


4.









8.

