sudo apt update

sudo apt install cpu-checker

sudo apt install qemu-kvm libvirt-daemon-system libvirt-clients bridge-utils -y

sudo apt install qemu-kvm libvirt-daemon-system libvirt-clients bridge-utils virt-manager

sudo adduser $(whoami) libvirt

sudo adduser $(whoami) kvm

sudo apt install virt-manager

virt-manager

**VirtualBox** is an open-source virtualization software developed by Oracle that allows you to run multiple operating systems (like Windows, Linux, or macOS) simultaneously on a single physical computer. It creates a virtual environment where you can install and use an operating system as if it were running on a separate machine.

**Google App Engine (GAE)** is a cloud computing platform provided by Google that lets developers build and run applications without managing the underlying infrastructure. It automatically handles things like servers, scaling, and load balancing, so you can focus only on writing your code.

In short, **App Engine** is a **Platform as a Service (PaaS)** — you just upload your application, and Google takes care of everything else (like scaling it up if more users come or scaling it down when it's quiet).

It supports popular programming languages like **Python, Java, Node.js, Go**, and more.

Ip add show

Ping ip\_address -c 5

**Network Virtualization Definition:**

👉 **Network virtualization** is the process of **combining hardware (like switches, routers, and servers)** and **software resources** into a **single virtual network**.

It allows multiple **virtual networks** to run on **top of one physical network** — just like running multiple apps on one computer.

**In simple words:**

"Network virtualization makes one physical network behave like **many separate networks**."

**✨ Key Points:**

* It **improves network efficiency** and **reduces hardware costs**.
* Networks can be **created, changed, and managed easily** through **software**, without touching the physical devices.
* It is a key technology behind **cloud computing** and **data centers**.

**Types of Network Virtualization**

**1. Internal Network Virtualization**

* Combines resources inside a single server (like CPU, storage, network cards) to create **virtual machines (VMs)**.
* **Each VM** has its own **virtual network**.
* Common in **data centers**.

✅ Example: VMware, Hyper-V

**2. External Network Virtualization**

* Combines multiple **physical networks** into **one virtual network** or splits **one physical network** into **multiple isolated virtual networks**.
* Used when you want **many networks** over **shared infrastructure**.

**KVM lets you run multiple operating systems (like Windows, Linux) at the same time on a Linux computer.**

sudo apt update

sudo apt install cpu-checker

sudo apt install qemu-kvm libvirt-daemon-system libvirt-clients bridge-utils -y

sudo apt install qemu-kvm libvirt-daemon-system libvirt-clients bridge-utils virt-manager

sudo adduser $(whoami) libvirt

sudo adduser $(whoami) kvm

sudo apt install virt-manager

virt-manager

**What is Hadoop?**

**Hadoop** is an **open-source framework** that allows you to **store** and **process** **huge amounts of data** (big data) across **many computers** (called a cluster) in a **fast, reliable, and scalable** way.

**Hadoop has two main components:**

* HDFS (Hadoop Distributed File System): This is the storage component of Hadoop, which allows for the storage of large amounts of data across multiple machines. It is designed to work with commodity hardware, which makes it cost-effective.
* YARN (Yet Another Resource Negotiator): This is the resource management component of Hadoop, which manages the allocation of resources (such as CPU and memory) for processing the data stored in HDFS.

 **Distributed Storage**:  
Hadoop **splits** large files into smaller parts and **stores them across multiple computers** (called nodes).

 **Parallel Processing**:  
Hadoop can **process data on all the nodes at the same time** — making it very fast for big

datasets.

 **Fault Tolerance**:  
If one computer (node) crashes, Hadoop **automatically recovers** using data copies stored elsewhere.

 **Scalability**:  
You can **add more machines** easily if you need more storage or computing power.

**What is MapReduce?**

**MapReduce** is a **programming model** (method) for **processing big data** by:

* **Splitting** tasks into smaller parts
* **Processing** them in **parallel** across many machines
* Then **combining** the results at the end

It was invented by **Google** to handle **huge datasets**.

In Hadoop, **MapReduce** is one of the main engines to **analyze and transform big data** stored in **HDFS**.

MapReduce is a programming model and associated implementation for processing and generating large datasets in a parallel and distributed manner, particularly within the Hadoop ecosystem. It works by dividing a large task into two phases: the map phase, which transforms data into key-value pairs, and the reduce phase, which aggregates these pairs based on their keys.

Here's a more detailed explanation:

1. Map Phase:

* The input data is split into smaller chunks and distributed across multiple nodes in a cluster.
* A map function is applied to each chunk independently, transforming the data into key-value pairs.
* For example, in a word count task, the map function might take a line of text and emit each word as a key and a count of 1 as the value.

2. Reduce Phase:

* The reduce phase takes the key-value pairs produced by the map phase and groups them based on their keys.
* A reduce function is applied to each group, summarizing the values associated with the same key.
* In the word count example, the reduce function would add up the counts for each word to produce a final count.

 **KVM** is a **virtualization technology** built **inside the Linux kernel**.

 It allows you to **turn your Linux system into a hypervisor** — meaning you can run **multiple virtual machines (VMs)** on your computer.

**What is a Hypervisor?**

A **hypervisor** is **software** (or sometimes firmware) that **creates and manages virtual machines (VMs)**.

**It allows multiple operating systems to share a single physical computer.**  
Each virtual machine acts like a complete independent computer (with its own CPU, memory, storage, etc.).

**Two Types of Hypervisors:**

| **Type** | **Description** | **Examples** |
| --- | --- | --- |
| **Type 1 (Bare-metal)** | Runs **directly on hardware**, no main OS needed | KVM, VMware ESXi, Microsoft Hyper-V (Server), Xen |
| **Type 2 (Hosted)** | Runs **inside an existing OS** (like an app) | VirtualBox, VMware Workstation, Parallels |