MARVEL

**CLOUD COMPUTING**

https://lh6.googleusercontent.com/yFyjrHbN7zCNrEq3AWxJxDHW8wtNMeR3wElAif2loPz3kYB3Qstz5DKCiStnh57nLpNtlvUCBn2NTsJIs53B9fyl4qZN_WYAbKHf7C-Gya5R3_brkUiG861rMMoCWic7nVLgGDuI

**TOPICS NAME: Map Reduce                                                                    Date: 28-04-2022**

MapReduce is a programming paradigm that enables massive scalability across hundreds or thousands of servers in a Hadoop cluster. 

There are two phases in map reduce. One is the map phase and the other is the reduce phase. In the map phase the whole data is reduced into various chunks. These chunks are worked at the same time in parallel.

In the reduce phase the data that is worked on is collected and organized together the get the desired output.

https://www.youtube.com/watch?v=PhdRyrmbRYQ

**Adeesh Padwalkar**

**TOPIC NAME: Gossip**

Gossip is a communication protocol. It is a process computer to computer communication that works in the same way as how information is shared on the social media networks.

The gossip protocol is used to solve the problems caused by multicasting (Multicasting: Sending data across a computer network to several users at the same time). In this type of communication the data is sent from one or more nodes to a set of other nodes in a network. This is useful when a group of clients in the network need the data at the same time.

https://www.geeksforgeeks.org/the-gossip-protocol-in-cloud-computing/

**Adeesh Padwalkar**

**TOPIC NAME: P2P system**

P2P stands for peer to peer. A P2P network is a network where computer systems are connected to each other via the internet. In a P2P network files can be directly shared between systems on the network without the need of central system. Hence every system on a P2P network acts like a file server.

https://techterms.com/definition/p2p#:~:text=Stands%20for%20%22Peer%20to%20Peer,as%20well%20as%20a%20client.

**Adeesh Padwalkar**

**TOPIC NAME: Key Value Stores**

The key value database is the simplest form of NoSQL database. Every single item in the database is stored as an attribute name (basically the key) together with its value. Therefore the data is stored as a collection of key-value pairs. The values can be integer, strings or complex data types.

To retrieve the data we must have the exact key. The advantage of key store is the new values can be easily added to the database by assigning new pairs.

https://www.youtube.com/watch?v=XroDCGGED04

**Adeesh Padwalkar**

**TOPIC NAME: Concurrency and Replication Control**

In a given database if a given data is accessed by multiple users at the same time then it is called as concurrent access. It means that the same database is executed simultaneously on a multi-user system by different users.   
 The problem with concurrent control is that if the two main operation i.e read and write, in a concurrent execution are not performed in an interleaved manner then the data may become inconsistent.

Data replication is the process in which the data is copied at multiple locations to improve the availability of data. The goal of replication is to first increase the availability of the data and second is to speed up the query evaluation. Types of replication control are

* Synchronous Replication Control
* Asynchronous Replication Control

javatpoint.com/dbms-concurrency-control

https://www.tutorialspoint.com/distributed\_dbms/distributed\_dbms\_replication\_control.htm

**Adeesh Padwalkar**

**TOPIC NAME: Containers**

Containers are effectively a form of virtualization. They let us do app isolation unlike the OS isolation in the virtual machine. Basically containers are packages of software that contain all of the necessary elements to run in any environment. Containers virtualize the operating system and run anywhere, from a private data center to the public cloud or even on a developer’s personal laptop

<https://cloud.google.com/learn/what-are-containers>

https://www.youtube.com/watch?v=GOuVeZmMee0

**Adeesh Padwalkar**

**TOPIC NAME: Virtual Machines**

Virtual Machines, commonly also known as VM’s are basically virtual computers or software defined computers within physical servers existing only as code. A virtual machine is no different from any other physical computer such as laptop or a smart phone. IT contains a CPU, memory to store your files and can connect to the internet if needed.

https://azure.microsoft.com/en-in/overview/what-is-a-virtual-machine/#how-do-work

**Adeesh Padwalkar**

**TOPIC NAME: JVM**

JVM stands for Java Virtual Machine. It is a virtual machine that enables a computer to run Java programs as well as programs written in other languages that are also compiled to java bytecode.

https://en.wikipedia.org/wiki/Java\_virtual\_machine#:~:text=A%20Java%20virtual%20machine%20(JVM)%20is%20a%20virtual%20machine%20that,required%20in%20a%20JVM%20implementation.

**Adeesh Padwalkar**

**TOPIC NAME: MAAS**

Maas stands for Mobility as a Service. It is an open source software that offers cloud style provisioning for physical servers.

**Adeesh Padwalkar**

**TOPIC NAME: PAAS**

Maas stands for Platform as a Service. It is a cloud computing model where the third party provider delivers hardware and software tools to the user over the internet. Usually these are the tools needed for application development.

https://www.techtarget.com/searchcloudcomputing/definition/Platform-as-a-Service-PaaS#:~:text=Platform%20as%20a%20service%20(PaaS,software%20on%20its%20own%20infrastructure.

**Adeesh Padwalkar**

**TOPIC NAME: Web Services**

A web service is a standardized method for propagating messages between client and server applications on the World Wide Web. A web service is a set of open protocols and standards that allow data exchange between different applications or systems.

**Adeesh Padwalkar**

**TOPIC NAME: Storage**

**Ceph:** Ceph is an open source software defined storage platform that implements object storage. Ceph replicates data and makes it fault tolerant.

**Swift:** Swift is also known as OpenStack Object store project. It offers cloud storage so that you can store and retrieve lots of data using a simple API.

**HDFS:**  
 The Hadoop Distributed File System (HDFS) is a distributed file system designed to run on commodity hardware. HDFS is highly fault-tolerant and is designed to be deployed on low-cost hardware.

**Adeesh Padwalkar**