

# Technologies Enabling IoT



## Team Members

- Shreedhar Bargaje
- Shivanjali Bhosale
- Chaitanya Deore
- Varad Deshmukh

PRN-123B1F006

PRN-123B1F009

PRN-123B1F020

PRN-123B1F022

GUIDED BY - SONAL SHIRKE MAM



# COMPUTING



# COMPUTING

1.

## Definition

Computing refers to the process of using computer technology to complete a given goal, such as calculations, data processing, or automation.

## Examples:

Personal computers, servers, embedded systems.

# DISTRIBUTED COMPUTING

1.

## Definition

A computing model where multiple computers (nodes) work together to solve a task by sharing processing power, memory, and resources.

2.

## Features

- Tasks are divided among multiple machines.
- Improved performance and fault tolerance.
- Used in large-scale systems like scientific simulations and big data processing

## Examples:

Blockchain, Google Search Engine, Distributed Databases

# GOOGLE SEARCH ENGINE - DISTRIBUTED COMPUTING USE CASE

Google Search is one of the best real-world examples of distributed computing. Here's how Google leverages distributed systems to process billions of search queries efficiently.

## 1. Distributed Web Crawling

- Googlebot, a distributed web crawler, collects data from millions of web pages.
- Works across thousands of machines in parallel to fetch and store web data.
- Uses Google File System (GFS) to store and manage large-scale web data efficiently.

## 2. Indexing (Data Organization)

- The collected data is processed and indexed using MapReduce, a distributed computing model.
- Indexing breaks down web pages into keywords, metadata, and links, which are stored in multiple distributed databases.
- Google's Bigtable, a distributed database, is used for storing indexed data.

## 3. Query Processing (Search Execution)

- When a user searches for something, the request is sent to Google's nearest data center using Content Delivery Networks (CDNs) for fast response.
- The query is processed in parallel across multiple servers to find relevant results.

# GOOGLE SEARCH ENGINE - DISTRIBUTED COMPUTING USE CASE

4.

## Ranking and Result Display

- Google's PageRank Algorithm, which runs on thousands of machines, determines the best results based on factors like:
  - Relevance
  - Authority (backlinks, domain ranking)
  - User behavior (click-through rates)
- Results are ranked and displayed within milliseconds.

## Why Google Uses Distributed Computing?

- Scalability: Can handle billions of searches daily.
- Fault Tolerance: If one machine fails, others take over.
- Speed: Parallel processing ensures fast response times.
- Cost-Effectiveness: Uses commodity hardware instead of expensive supercomputers.

# CLOUD COMPUTING

1.

## Definition

A computing model that provides on-demand access to computing resources (like servers, storage, and databases) over the internet.

2.

## Features

- Scalable and flexible.
- Cost-effective (pay-as-you-go model).
- No need for local infrastructure.

## Examples:

Netflix, Google Drive, AWS.

**On-Demand Self-Service** – Users can access resources whenever they need them, without . human intervention



## KEY CHARACTERISTICS OF CLOUD COMPUTING

**Resource Pooling** – Cloud providers allocate resources dynamically among multiple users

**Broad Network Access** – Services are available over the internet from any location.

**Scalability & Elasticity** – Users can scale resources up or down depending on demand.

# NETFLIX - CLOUD COMPUTING USE CASE

Netflix is one of the best real-world examples of cloud computing. Here's how Netflix leverages cloud computing to deliver a seamless experience to millions of users worldwide

## 1. AWS-Based Infrastructure

- Netflix runs almost entirely on Amazon Web Services (AWS), eliminating the need for on-premise servers.
- This allows Netflix to scale up or down based on demand, reducing costs and improving efficiency

## 2. Scalability & High Availability

- Netflix has over 230+ million subscribers globally.
- To handle high traffic, cloud computing enables automatic scaling during peak hours (e.g., weekends, new show releases).
- AWS data centers ensure low-latency streaming, regardless of location

## 3. Content Delivery Network (CDN) – Open Connect

- To provide buffer-free streaming, Netflix has its own CDN called Open Connect.
- This system caches content closer to users by storing copies of popular shows/movies in regional data centers.
- Reduces the load on central servers and improves speed

# NETFLIX - CLOUD COMPUTING USE CASE

4.

## Data Analytics & AI on the Cloud

Netflix uses cloud-based Big Data analytics and AI algorithms to:

- Personalize recommendations based on viewing history.
- Optimize video quality based on the user's internet speed.
- Prevent fraud and unauthorized account sharing

5.

## Disaster Recovery & Security

- Cloud computing ensures automatic backups of Netflix's content and user data.
- Security features like encryption, firewalls, and multi-factor authentication protect user accounts and transactions

# FEATURE

## DISTRIBUTED COMPUTING

## CLOUD COMPUTING

Definition

Multiple connected devices working together

Uses remote data centers to provide services

Infrastructure

Uses multiple physical machines

Hosted on cloud providers like AWS, Azure

Scalability

Limited by hardware setup

Highly scalable & elastic

Fault Tolerance

Managed across multiple nodes

Cloud providers handle redundancy

Example

Blockchain, Google Engine

Netflix, Google Cloud



# HADOOP

# WHAT IS HADOOP ?



1.

## Definition

- Open-source framework for distributed data storage & processing.
- Handles large-scale data using a cluster of computers.
- Part of the Apache Software Foundation.

2.

## Why Hadoop ?

- Manages Big Data efficiently.
- Provides scalability, fault tolerance, and cost-effectiveness.
- Works on commodity hardware, reducing costs.





# CORE COMPONENTS OF HADOOP

HDFS (Hadoop  
Distributed File  
System)



MapReduce

YARN (Yet  
Another Resource  
Negotiator)

Hadoop  
Common

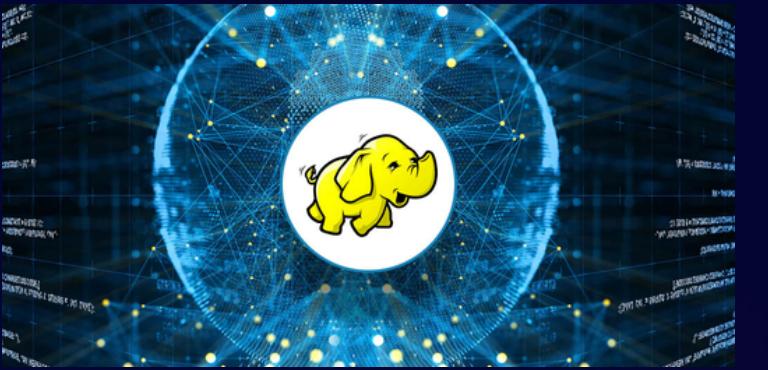
# USE CASES:

- 
1. Search Engines
  2. Social Media Analytics
  3. Internet of Things (IoT)  
Data Processing
  4. Data analytics & machine learning.
  5. Fraud detection in banking.
  6. Log processing & recommendation systems.



# HADOOP IN U.S. NATIONAL SECURITY





# HADOOP'S ROLE: USED BY

- Big Data Framework
- Scalability, Fault Tolerance, efficiency
- Intelligence gathering, surveillance, and cybersecurity.

- NSA (National Security Agency)
- DoD (Department of Defense)
- DHS (Department of Homeland Security)

## DATA TYPES

- Structured Data
- Semi-structured
- Unstructured Data

1.

Mass Data  
Storage &  
Processing

2.

Real-Time  
Threat  
Detection &  
Cybersecurity

3.

Intelligence  
Correlation &  
Counterterrorism

4.

Predictive  
Analytics for  
Risk  
Assessment

# KEY APPLICATIONS OF HADOOP IN NATIONAL SECURITY

# WHY HADOOP IS IDEAL FOR NATIONAL SECURITY

1. Scalability
2. Cost-Effectiveness
3. Fault Tolerance & High Availability
4. Flexibility in Handling Diverse Data

# MORE USE CASES OF HADOOP

Facebook

Twitter

LinkedIn

NASA



# THANK YOU!

