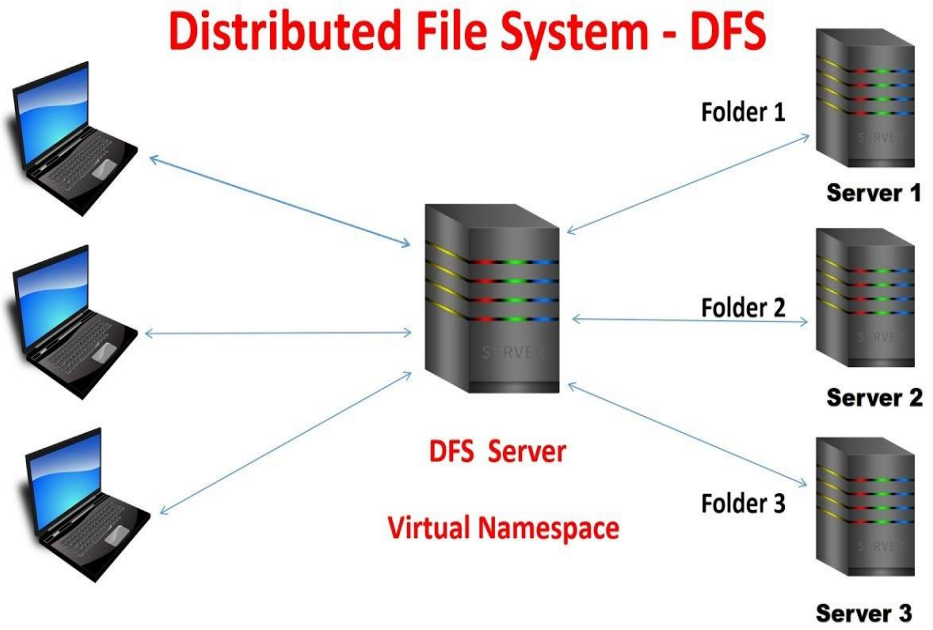
Distributed File Systems & Emerging Storage Technologies

PRESENT BY: BOM BAHADUR BK
LEVEL: MASTER
NEPAL COLLEGE OF INFORMATION TECHNOLOGY(NCIT)
BALKUMARI, LALITPUR

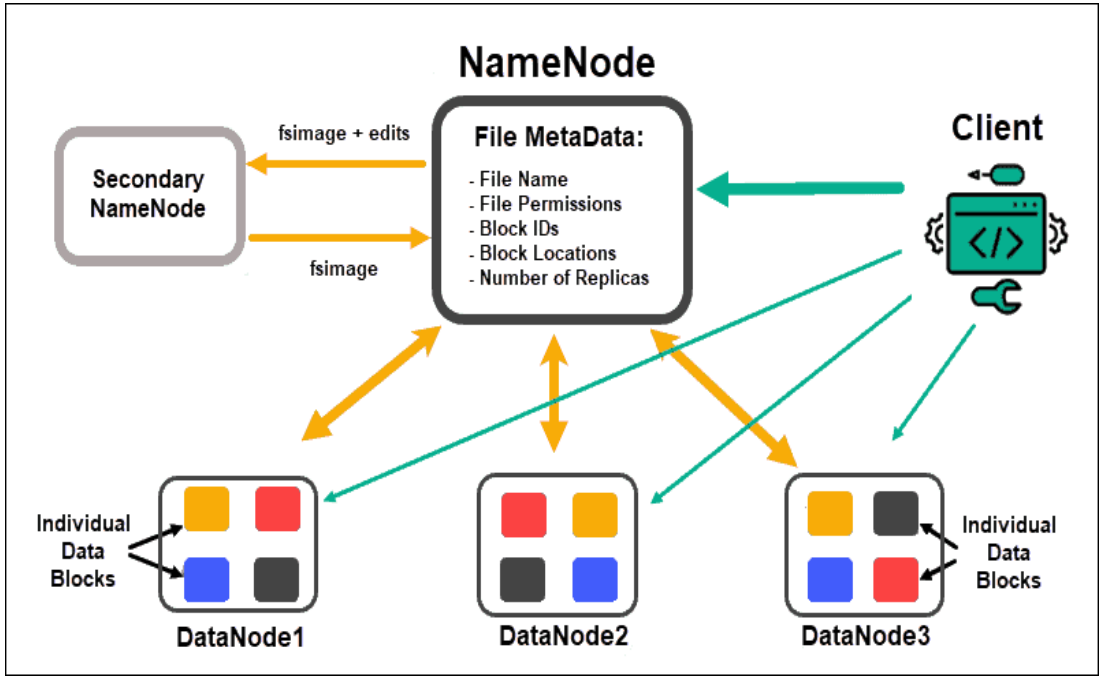
Introduction

- **What is a Distributed File System?**
  - A file system that allows storing and accessing files across multiple servers.
  - Designed for scalability, fault tolerance, and high availability.
- **Emerging Storage Technologies:**
  - Innovations like DNA, holographic, and quantum storage aim to revolutionize data storage for the future.



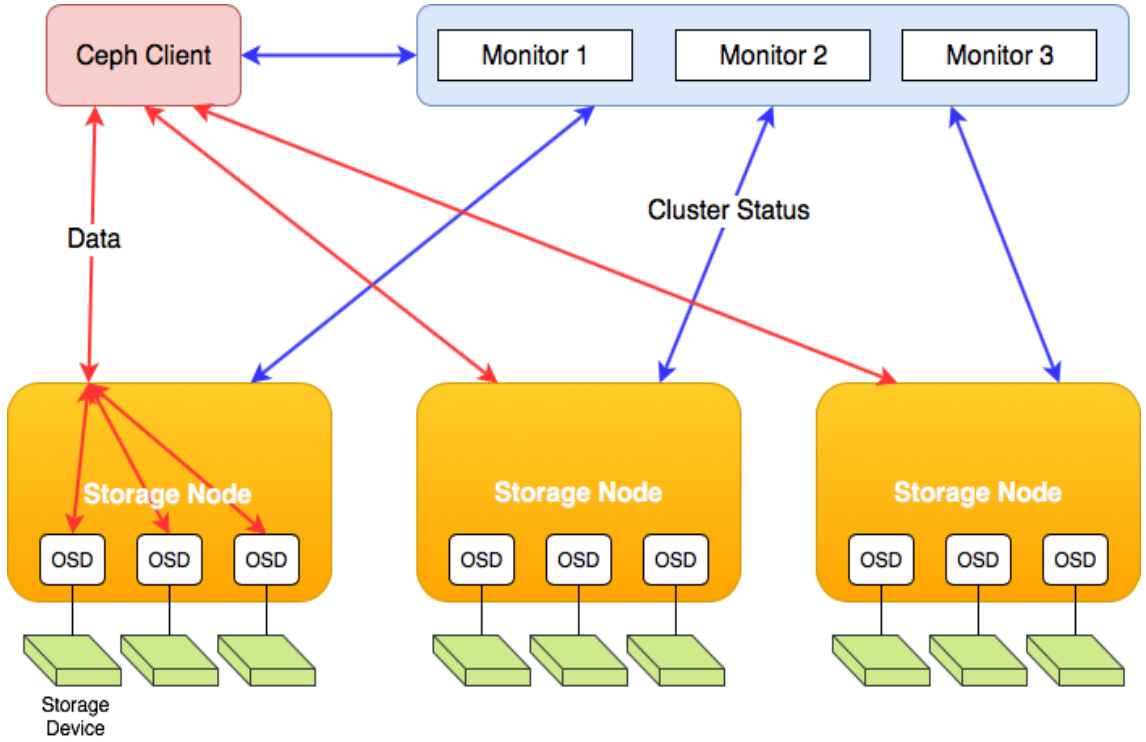
Hadoop Distributed File System (HDFS)

- **Overview:**
  - Designed for big data processing using Hadoop.
  - Stores large files across multiple machines.
- **Key Features:**
  - Fault tolerance (replication of data blocks).
  - Scalability (handles petabytes of data).
  - High throughput for large datasets.
- **Use Cases:**
  - Big data analytics, log processing, and data warehousing.



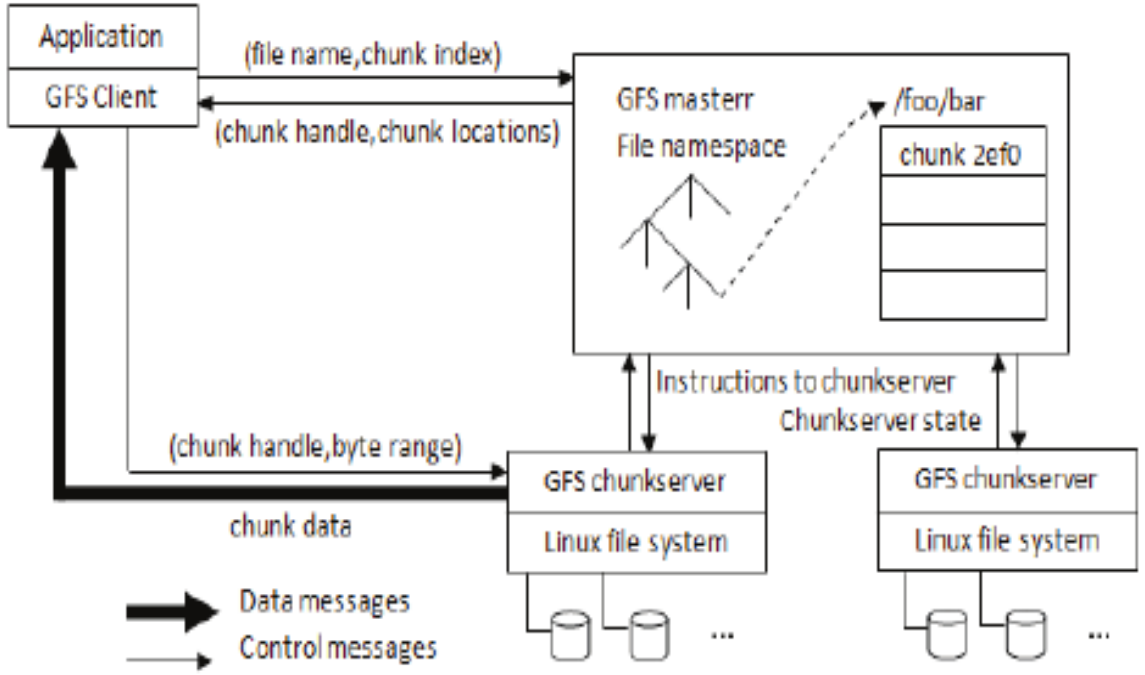
Ceph

- **Overview:**
  - A unified distributed storage system providing object, block, and file storage.
  - Open-source and highly scalable.
- **Key Features:**
  - Decentralized architecture (no single point of failure).
  - Self-healing and self-managing.
  - Supports petabytes of data.
- **Use Cases:**
  - Cloud storage, virtualization, and backup systems.



Google File System (GFS)

- **Overview:**
  - Proprietary distributed file system developed by Google.
  - Designed for large-scale data processing.
- **Key Features:**
  - Master-slave architecture.
  - High fault tolerance through replication.
  - Optimized for large files and sequential reads/writes.
- **Use Cases:**
  - Google Search, YouTube, and other Google services.

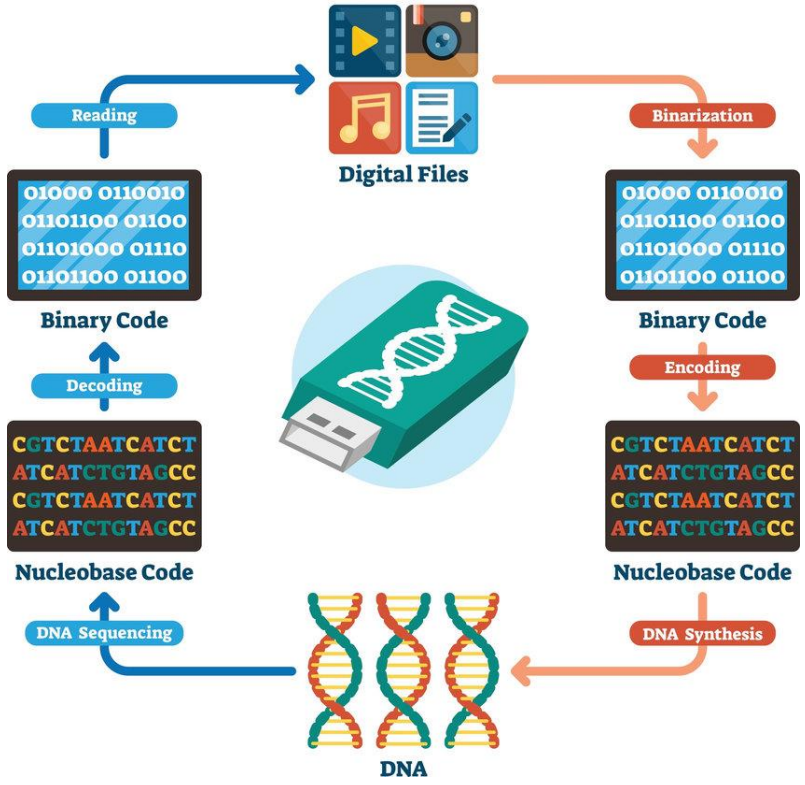


Comparison of HDFS, Ceph, and GFS

Feature	HDFS	Ceph	GFS
Architecture	Master-Slave	Decentralized	Master-Slave
Scalability	High	Very High	Very High
Fault Tolerance	Replication	Self-Healing	Replication
Use Case	Big Data Analytics	Cloud Storage	Google Services

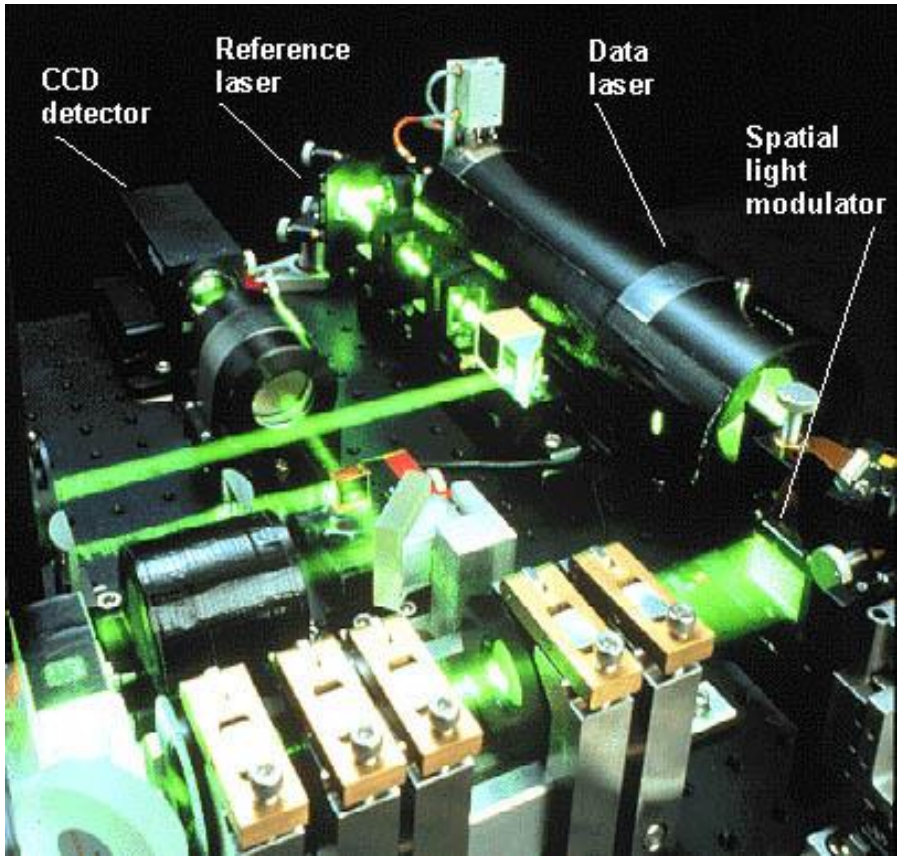
DNA Storage

- **Overview:**
  - Uses synthetic DNA molecules to store digital data.
  - Extremely high density and longevity (thousands of years).
- **Key Features:**
  - 1 gram of DNA can store ~215 petabytes of data.
  - Energy-efficient and compact.
- **Challenges:**
  - High cost and slow read/write speeds.
- **Use Cases:**
  - Long-term archival storage, scientific research.



Holographic Storage

- **Overview:**
  - Uses lasers to store data in 3D holograms within a photosensitive medium.
- **Key Features:**
  - High storage density and fast data access.
  - Potential for terabytes of data in a single disc.
- **Challenges:**
  - Expensive and complex technology.
- **Use Cases:**
  - Data archiving, medical imaging, and multimedia storage.



Quantum Storage

- **Overview:**
  - Uses quantum mechanics principles to store and process data.
  - Quantum bits (qubits) enable superposition and entanglement.
- **Key Features:**
  - Potential for exponentially higher storage capacity.
  - Enables quantum computing and secure data transmission.
- **Challenges:**
  - Requires extremely low temperatures and is still in experimental stages.
- **Use Cases:**
  - Quantum computing, cryptography, and advanced research.

Comparison of Emerging Storage Technologies

Feature	DNA Storage	Holographic Storage	Quantum Storage
Capacity	Extremely High	High	Exponentially High
Speed	Slow	Fast	Ultra-Fast (theoretical)
Cost	High	High	Very High
Maturity	Experimental	Developing	Experimental

Thank You So Much!

Have a great time...