



SNS COLLEGE OF TECHNOLOGY
(An Autonomous Institution)
COIMBATORE-35



DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

BIG DATA ANALYTICS


NIMMY PRABHA
AP/AIML

RDBMS Versus Hadoop

FEATURES	RDBMS	HADOOP
Data structure	Structured data in fixed schemas (tables)	Unstructured, semi-structured, and structured data (schema-on-read)
Storage	Centralized storage on a single server/cluster	Distributed storage across multiple nodes
Query Language	SQL (Structured Query Language)	Tools like Hive (SQL-like), MapReduce for processing
Transaction Support	Strong ACID compliance	Generally lacks full ACID compliance; focuses on eventual consistency
Performance	High performance with small/medium datasets; may struggle with scalability	High scalability for large datasets; individual queries may be slower

Distributed Computing Challenges

Features	Challenges
Network Reliability	Network failures can disrupt communication, leading to data loss or inconsistent states.
Data Consistency	Maintaining consistency is difficult due to the CAP theorem; only two of consistency, availability, and partition tolerance can be guaranteed
Latency	Increased network latency can affect performance, especially in real-time processing scenarios
Fault Tolerance	Designing systems to recover from failures without data loss or significant downtime is complex.
Scalability	Ensuring efficient scaling as demand grows poses challenges in data storage and processing power
Security	Distributed systems are more vulnerable to security threats; effective security measures across nodes are essential.
Complexity of Development	Building distributed applications involves increased complexity in architecture and debugging.



Resource Management	Efficiently managing CPU, memory, and bandwidth across nodes requires sophisticated algorithms.
Heterogeneity	Diverse hardware and software environments lead to compatibility issues and integration challenges.
Testing and Debugging	Testing and debugging are complicated due to non-deterministic behavior and interactions between components.