① Types of Schedulers in Hadoop.

Schedulers in Hadoop determine how tasks are assigned to resources (CPU, memory) in a cluster.

- FIFO (First-In-First-Out) Scheduler:-

Oldest jobs are scheduled first. Simple but not resource-efficient.

- Capacity Schedular:-

Allow multiple orgaizations to share a cluster. Resources are divided into queues, each with guaranteed capacity.

- Fair Scheduler:

Distributes resources so all jobs get an equal share over time. Supports pooling & preemption.

- Resource Manager Scheduler in YARN:-

In Hadoop YARN, the resource manager manages scheduling using Capacity or Fair scheduler

② Limitations of Hadoop

- High latency:- Not suitable for real-time processing (batch-oriented).
- Small file problem:- HDFS performs poorly with many small files.
- No transactional support:- Limited ACID compliance.
- Complex programming model:- Writing MAPReduce jobs requires deep technical knowledges.
- Resource intensive:- Needs a lot of memory & disk space.
- Data security:- Basic by default; advanced features (Kerberos, Ranger) need configurations.

pool SQL support:- Compared to traditional databases or newer engines like Hive or Spark SQL.

Features	DBMS	Data Warehouse (DW)	Hadoop
Purpose	OLTP (transactional)	OLAP (analysis)	large scale batch data processing.
Data Type	Structured	Structured	Structured, Semi structured. Unstructured.
Schema	Fixed (schema on write)	Fixed (schema on write)	Flexible (schema on read).
Scalability	Vertical	Moderate (up to Petabytes)	Horizontal (to petabytes/ exabytes)
Cost	Expensive	Very Expensive	Cost-effective (commodity h/w)
Performance	Fast for transactions.	Optimized for analytical queries	Slower for real-time, great for batch.
Ex:-	MYSQL, Oracle	Amazon Redshift, Snowflake, Teradata	Hadoop, HDFS + MapReduce/ YARN

Features	Map Reduce	YARN (Yet Another Resource Negotiator
Role	Programming model & execution engine	Resource management layer.
Job Management	Bundled with MapReduce	Decoupled from processing engine.
Resource Manager	Limited	Centralized (Resource Manager & Node Manager)
Multi framework	NO	Yes (Spark, Tez, Flink can run on YARN)
Scalability	Limited	Highly scalable

MapReduce is a data processing model; YARN is a more flexible & scalable resource management ~~fla~~ plateform introduced in Hapoop 2.X.