CBTU presents a course on **Big data and Hadoop**

Module 2: hadoop

Section 2.3: Hadoop Architecture
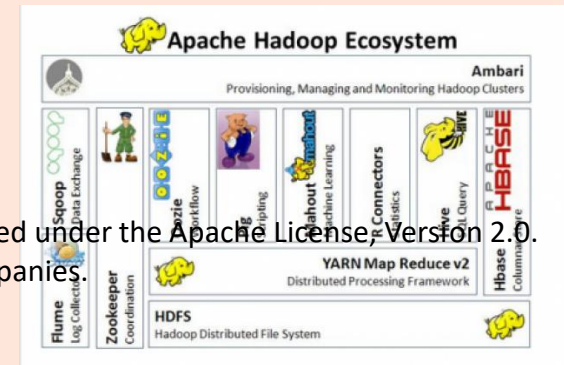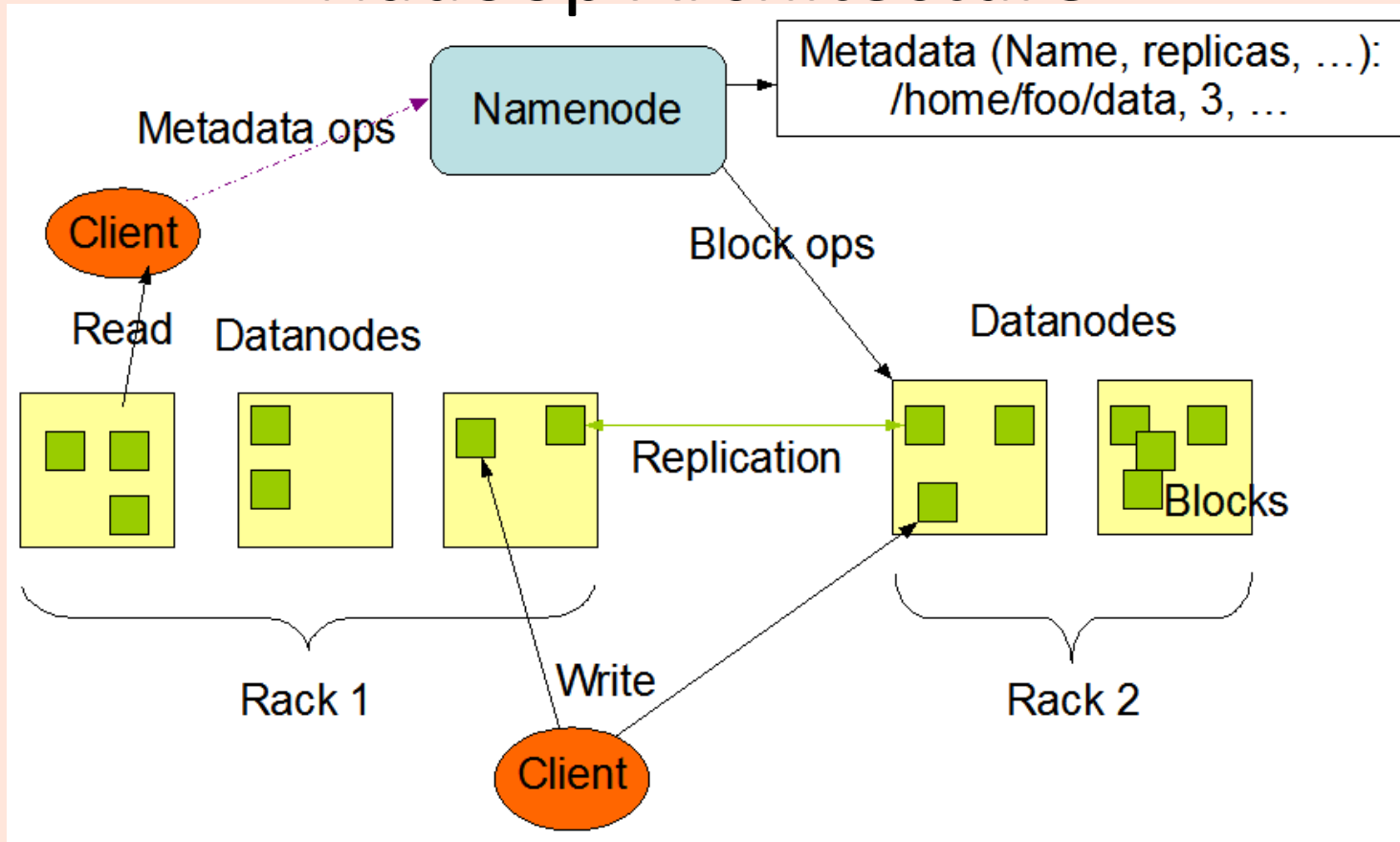
Hadoop Architecture

http://hadoop.apache.org/docs/current/hadoop-project-dist/hadoop-hdfs/HdfsDesign.html

Apache Hadoop framework

Apache Hadoop framework is composed of the following modules:

- Hadoop Common – contains libraries and utilities needed by other Hadoop modules.
- Hadoop Distributed File System (HDFS) – a distributed file-system that stores data on commodity machines.
- Hadoop YARN – manages computing resources in clusters .
- Hadoop MapReduce – an implementation of the MapReduce programming model for large scale data processing.

Hadoop Common

Hadoop Common package:

- Contains libraries and utilities needed by other Hadoop modules.

- Provides file system and OS level abstractions, a MR engine (either MapReduce/MR1 or YARN/MR2) and the HDFS.

- Contains the necessary Java ARchive (JAR) files and scripts needed to start Hadoop.

Architecture - redundancy

- HDFS maintains location/rack info/ where worker node is.

- Hadoop applications can use this information to execute code on the node where the data is.

- HDFS replicates data for data redundancy across multiple racks.

- Impact of a rack power outage or hardware failure is none as data remain available.

High Level Architecture of Hadoop
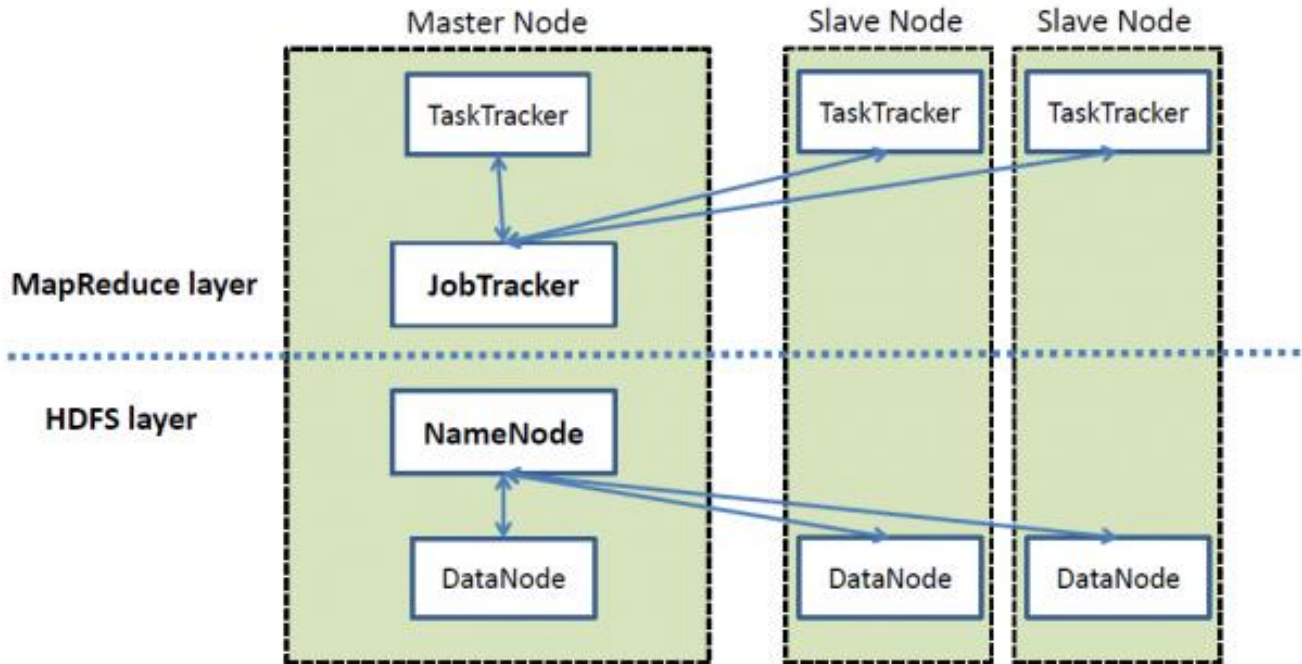
Source: opensource.com

Hadoop – Large cluster

In a larger cluster,

- HDFS nodes are managed through

  - a primary NameNode server to host the file system index and metadata.

  - a secondary NameNode that can generate snapshots of the namenode's memory structures, acting as backup.

- A standalone JobTracker server can manage job scheduling across nodes.

Hadoop Daemons

– NameNode

– DataNode

– ResourceManager

– NodeManager

```
[root@localhost hadoop-3.1.0]# jps
7634 Jps
5963 SecondaryNameNode
5582 NameNode
```

Hadoop Nodes

- The NameNode and DataNode are pieces of software designed to run on commodity machines. These machines typically run a GNU/Linux operating system (OS).

- HDFS can be deployed on a wide range of machines.

- A dedicated machine may run NameNode software and other nodes in the cluster runs an instance of the DataNode software.

- The NameNode is the controller and repository for all HDFS metadata.

Data Disk Failure, Heartbeats and Re-Replication

- Each DataNode sends a Heartbeat message to the NameNode periodically.

- The NameNode marks DataNodes without recent Heartbeats as dead and does not forward any new IO requests to them.

- NameNode constantly tracks which blocks need to be replicated.

- The time-out to mark DataNodes dead is ~10 minutes by default and configurable.

Data Integrity

- HDFS implements data integrity for files by implementing checksum checking and eliminates corrupt block of data.

- When a client creates a HDFS file, it computes a checksum of each block of the file and stores these checksums in a separate hidden file in the same HDFS namespace.

- When a client retrieves file contents it verifies that the data it received from each DataNode matches the checksum stored in the associated checksum file. If not, then the client can opt to retrieve that block from another DataNode that has a replica of that block.

FsImage and the EditLog

- The FsImage and the EditLog are central data structures of HDFS. Corruption of these files can cause the HDFS instance to be non-functional.

- NameNode is configured to support maintaining multiple copies of the FsImage and EditLog.

- Any update to either the FsImage or EditLog causes each of the FsImages and EditLogs to get updated synchronously.

- When a NameNode restarts, it selects the latest consistent FsImage and EditLog to use.

Snapshots

- Snapshots support storing a copy of data at a particular instant of time.

- Snapshots are useful to roll back a corrupted HDFS instance to a previously known good point in time.

Replication Pipelining

- Default replication factor is 3 data nodes for writing data. Namenode maintains list of datanodes where blocks reside.

- The client writes to the first DataNode which will replicate to the other nodes in the list.

- The data is pipelined from one DataNode to the next.

Accessibility

- FileSystem Java API and REST API.

- HTTP browser and can also be used to browse the files of a HDFS instance.

- HDFS can be mounted as part of the client's local file system.

- File System (FS) Shell CLI to interact with the data in HDFS.

Thanks for watching

CBTUniversity.com

learnq@CBTUniversity.com

+91 963 246 5599

/CBTUniversity