

Quick Revision on Hadoop 1.x Architecture

Hadoop Ecosystem

► Why to understand Hadoop 1.x which is history now??

- Yes, Hadoop 1.x is history now and most of the applications & Tech giants are using Hadoop 2.x But understanding Hadoop 1.x provides insights of main Hadoop components and to rectify its drawbacks, how it's been evolved over the time.
- Hadoop ecosystem has two types of Daemons:

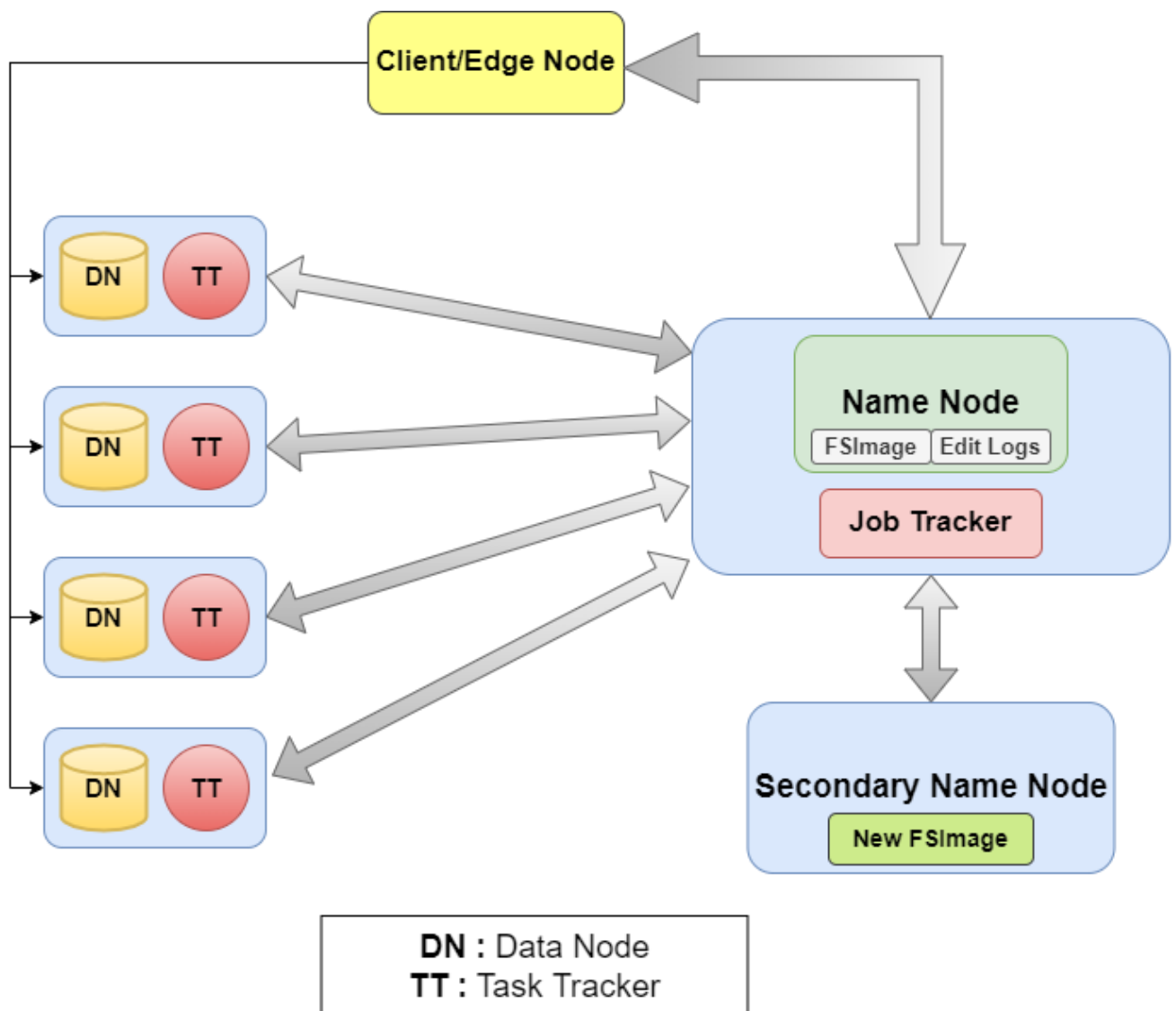
HDFS Daemons

- Name Node
- Secondary Name Node
- Data Node

Map-Reduce Daemons

- Job Tracker
- Task Tracker

Hadoop Architecture Diagram



Daemon & Client/Edge Node

What is Daemon?

It's a computer program that runs as a background process, rather than being under the direct control of an interactive user.



Client or Edge Node:

1. It's a computer which acts as end user to communicate with master and worker node
2. Edge Node is the access point for the external applications, tools, and users that need to utilize the Hadoop environment
3. The Edge Node sits between the Hadoop cluster and the corporate network to provide access control, policy enforcement, logging, and gateway services to the Hadoop environment

Name Node (NN)

- ▶ Its the master node in Hadoop ecosystem
- ▶ The main purpose of name node is cluster and Metadata management
- ▶ It keeps track of all the files in file system
- ▶ It also Tracks where across the cluster data file is kept
- ▶ It doesn't store the data of any files
- ▶ Name Node creates a FSImage of complete directory where it's stored on the OS file system
- ▶ It also has details about the files location which are in data node
- ▶ Edit log is the transactional log that records the change in HDFS file system. It includes logs like updates, deletes of files, adding new block. Basically it records all the changes since the last FSImage was created
- ▶ Name node also initiates replication of blocks in data node whenever necessary
- ▶ When the name node goes down, the cluster will go in safe mode

Secondary Name Node (SNN)

- ▶ It's not the backup name node
- ▶ Name node uploads FSImage and edit logs from SNN to create a new fs image in SNN at regular intervals like after every one hour or after every 1 million transactions in edit logs (Whatever happens 1st). This process is called checkpointing
- ▶ After checkpointing the edit log in NN will be truncated
- ▶ It helps NN in it's wrong time.
 - When name nodes run for a long period of time without restart, edit logs grows unwisely in size
 - NN restart can take longer as too many changes are to be merged which can cause to restart failure i.e., crashes then there will be significant amount of data loss as the FSImage used at a time of restart is very old

Data Node (DN)

- ▶ It's the worker node in HDFS
- ▶ The actual data is stored in Data Node
- ▶ A file system has more than one data node where data is replicated across them
- ▶ DN sends the heartbeat signal to name node in every 3 secs. Threshold of this is 10 mins and after that its marked dead by NN
- ▶ Client or edge node can directly communicate with data nodes once the file location is shared with them
- ▶ It performs block verification to identify corrupt blocks periodically

Job & Task Tracker

Job Tracker:

- ▶ (Acts like a master) It's responsible for complete execution of submitted map-reduce jobs to task trackers
- ▶ For every job submitted for execution in file system, there is one job tracker which resides on name node and there are multiple task trackers which reside in multiple data nodes
- ▶ In the event of the failure of task tracker, the job tracker will reschedule it to a different Task Tracker

Task Tracker:

- ▶ Task Tracker is coordinated by job tracker
- ▶ It's responsible to execute a portion of a job i.e., task assigned by job tracker in data nodes
- ▶ Task tracker sends the periodic update of task which it's executing to job tracker
- ▶ It also sends the heartbeat periodically to job tracker to notify the current state of the system

Limitation of Hadoop 1.x

- ▶ SPOF (Single Point of Failure). There is no backup node in Hadoop 1.x
- ▶ 2. Due to single name node there is no concept of namespace in 1.x
- ▶ It supports up to 4k node cluster. Adding more than that will cause an issue
- ▶ It's not suitable for Data Streaming

Examples on how data is processed and stored using Hadoop 1.x

- ▶ Client or Edge Node sends a File of 200Mb to process
- ▶ The 200MB File is divided into 4 Blocks, 3 Blocks of 64MB & 1 Block of 8MB
- ▶ Before assigning work to DN, NN checks of the resource availability. It Creates the FSImage which has metadata and makes an entry in edit logs
- ▶ Once the resources are available, NN creates a Job Tracker for the particular job to get the updates of job execution from Task Tracker
- ▶ For 1 job there is one job tracker and multiple task trackers and data nodes
- ▶ The 4 blocks which were splitted from the file of 200MB will reside in various Data Nodes. Blocks are written parallely to Data Nodes
- ▶ As soon as blocks are written, Replication job triggers to various other data nodes. Default Replication Factor is 3 so, the blocks will replicate 3 times. This process of replication happens sequentially

Examples on how data is processed and stored using Hadoop 1.x

- ▶ After Every 1 Hour or after 1 Million Transaction in Edit logs, Name Node uploads the data of FSImage and Edit logs to Secondary Name Node to create a new FSImage out of it. This New FSImage can be used if Name Node goes down (i.e., Crashes) or when name node scheduled restart takes place

Single Point of Failure:

- ▶ When the name node is down, the cluster will go in safe mode
- ▶ Hadoop admin team will try to bring name node up and will fetch the FSImage from secondary name node to name node
- ▶ When cluster is in safe mode client cannot send a request to compute something, they can only send the request to view or read any files which are present in data nodes

Data Node Failure:

- ▶ Now Imagine that client node sends one job to execute same as of above and when the computation is in progress in data node suddenly one of the data node goes down 😞
- ▶ As we know that Data node sends its “Heartbeat” as a signal to name node after every 3 secs but as data node is down, Name node is not receiving any signal from data node

Examples on how data is processed and stored using Hadoop 1.x

- ▶ Name node waits for 10 mins to receive its heartbeat and after that it will declare that particular data node as dead node
- ▶ Job tracker then assigns the work of dead node to other data nodes
- ▶ As you know replication factor is 3 and after one data node is down this doesn't satisfy the replication factor. This is called under replica
- ▶ Name node again creates a new replica to satisfy the condition
- ▶ Once the dead data node is brought back to life again, name node realizes that the replication of the particular FSImage is being done more than the replication factor (i.e., 3). This is called as over replica
- ▶ As the data node which was brought active is now seating ideal. So, name node takes some work from other node and assigns to the ideal node to maintain load balancing
- ▶ This is the basic working of Hadoop 1.x. ***This process also includes many different operations while computing which are not mentioned in this document***