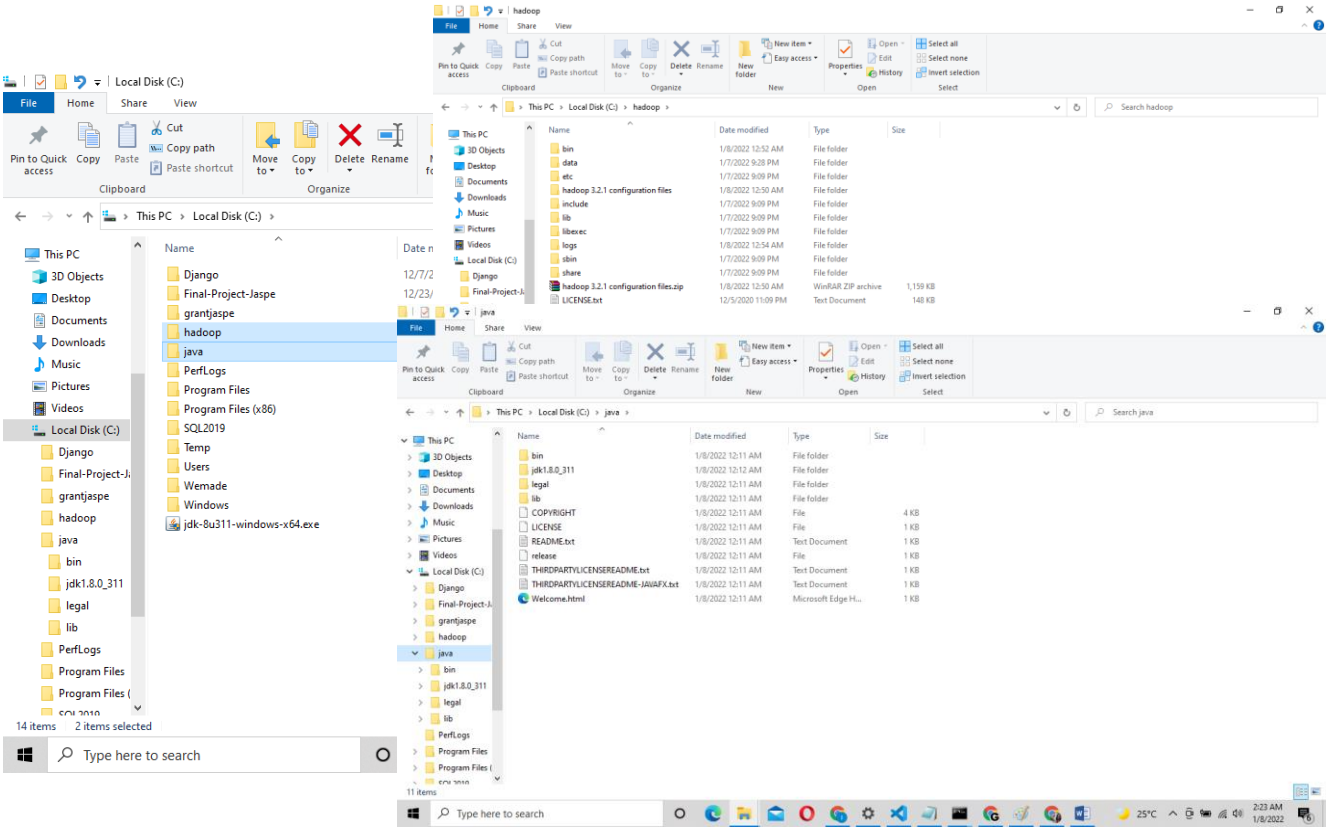
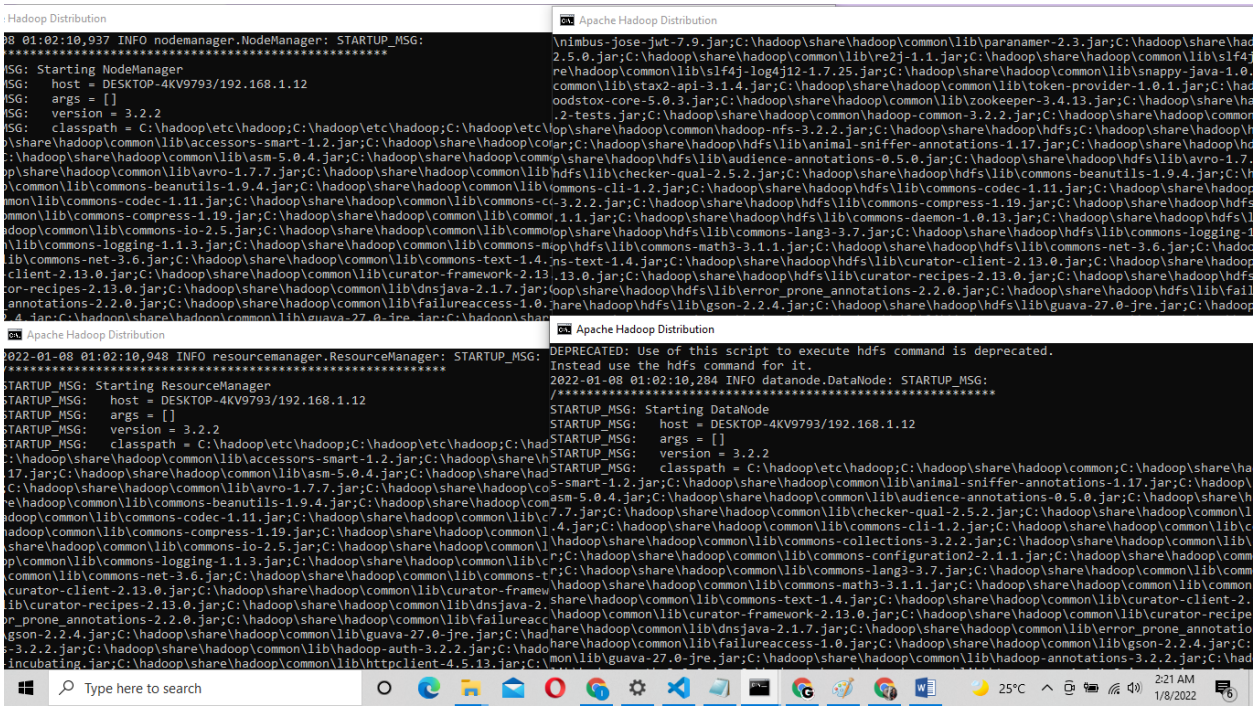
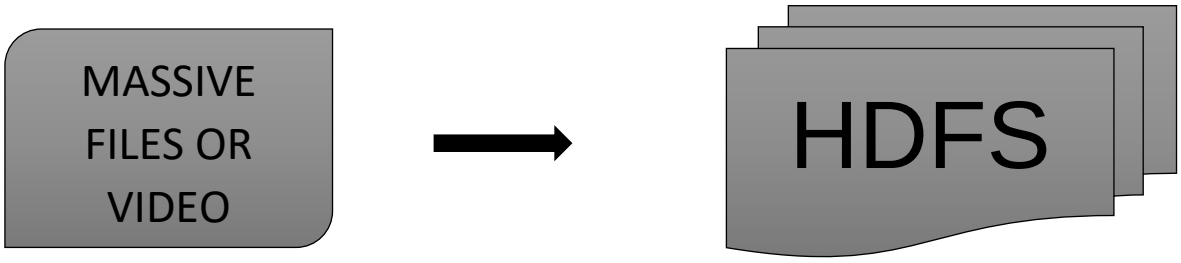
On commodity hardware clusters, Hadoop is an open-source software platform for storing data and executing applications. It has immense storage for any type of data, a lot of processing power, and the capacity to handle nearly infinite concurrent processes or jobs at the same time.



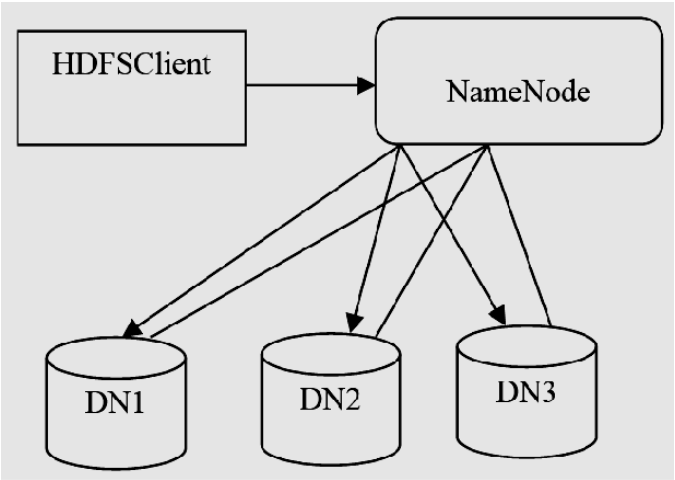


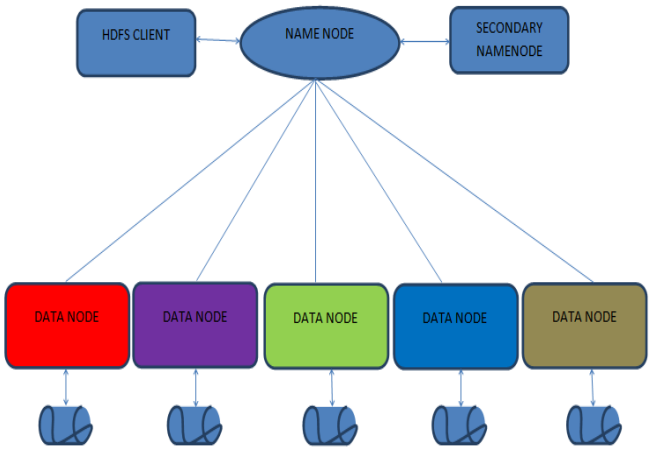
What I Learned in HDFS?

HDFS is a distributed file system that runs on commodity hardware and manages massive data collections.



HDFS is designed for applications with data sets ranging from gigabytes to terabytes in size. HDFS can expand to hundreds of nodes in a single cluster and delivers high aggregate data bandwidth.





Advantage of HDFS

It has the ability to store a big quantity of data, enables simple and quick access to the data, and allows for streaming reading. By dispersing data, access and recovery will be applied fast and efficiently.

Disadvantage of HDFS

Just managing a complex applications such as Hadoop can be challenging. A simple example can be seen in the Hadoop security model, which is disabled by default due to sheer complexity. The framework is written almost entirely in Java, one of the most widely used yet controversial programming languages in existence.