

# HADOOP (HDFS)

Topología basada en nodos maestro-esclavo, cada nodo tiene sus demonios ejecutándose internamente. Con su sistema de ficheros, se pueden escribir datos en el servidor una vez, leerlos y usarlos varias veces.

Provee de almacenamiento redundante de grandes volúmenes de datos transparente al usuario.

Veamos sus demonios...

- **Nodo Maestro**

- NameNode

- + Almacena y gestiona los metadatos, informa de permisos y usuarios, que bloques componen un archivo y donde están esos bloques.
    - + Los metadatos son guardados en disco y cargados en memoria cuando el clúster arranca en un fichero llamado **FSImage**.
    - + Los cambios realizados en metadatos se almacenan en un fichero llamado **edits** en memoria

- Secondary NameNode (SNN)

Se encarga de realizar las labores de mantenimiento cuenta con un checkpoint (copia del archivo edits y fusionarlo con FSImage) que cada 1 hora o cada 1M de transacciones se activa mediante el SNN.

Asegura que los metadatos estén regularmente respaldados y listos para facilitar la recuperación del sistema en caso de fallo del NameNode principal.

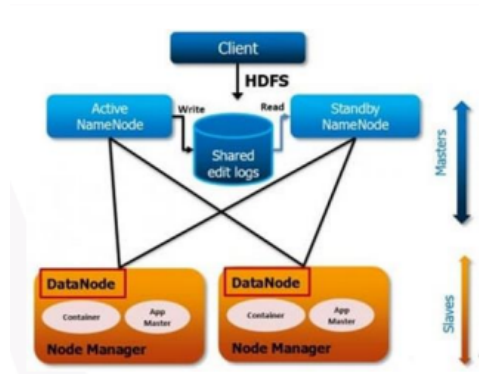
- Standby NameNode

Actúa como una copia de seguridad en vivo del NameNode activo, asegurando que los metadatos estén siempre respaldados y preparados para facilitar una recuperación rápida en caso de fallo del sistema principal.

- **Nodo Esclavo**

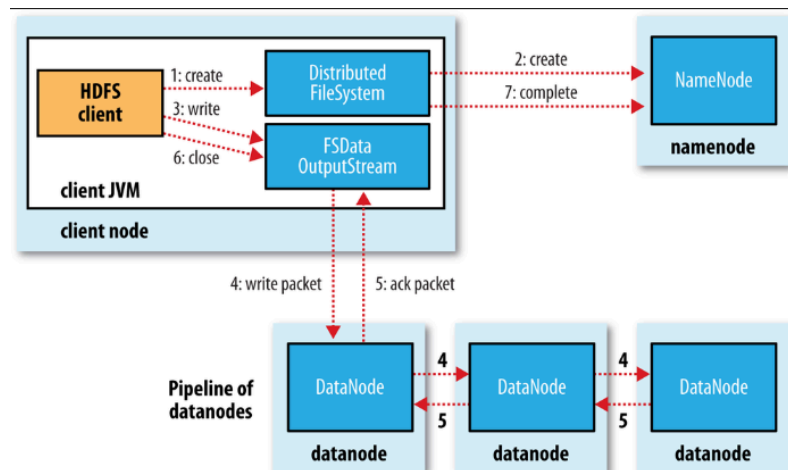
- DataNode

- + Los nodos que son DataNode, almacenan los bloques de datos (NO información referente al bloque en sí)
    - + Se encarga de acceder a dichos bloques de datos. Cuando un fichero entra en el clúster, se divide en bloques de datos, cada nodo esclavo se encarga de hacer una copia al siguiente nodo esclavo, hasta llegar al **factor de replicación**.
    - + Se encarga de gestionar las **tasks** que componen un job.



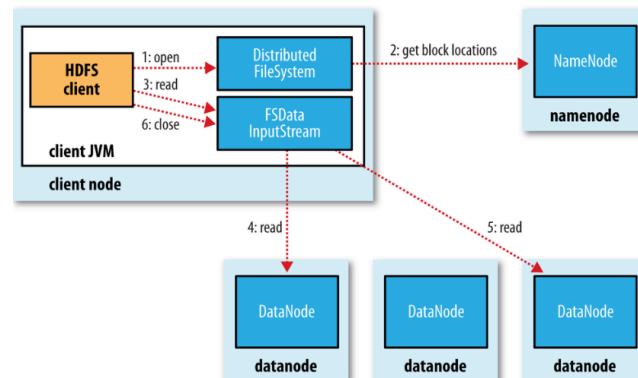
## - Escritura

1. El cliente se conecta al Namenode.
2. El NameNode busca sus metadatos para obtener información sobre el archivo y devuelve al cliente el nombre del bloque y la lista de los DataNodes donde el bloque será almacenado.
3. El cliente se conecta con el primer DataNode de la lista y empieza el envío de los datos.
4. El cliente se conecta con el segundo DataNode para realizar el envío y lo mismo con el tercero (si se está utilizando una replicación de tres).
5. Finaliza el proceso de envío.
6. El cliente indica al NameNode donde ha realizado la escritura (la ubicación de escritura).



## - Lectura

1. El cliente se conecta al Namenode y solicita la ubicación del archivo deseado.
2. El NameNode verifica sus permisos y si está autorizado devuelve una lista con los DataNodes que contienen el bloque buscado.
3. El cliente conecta con el primero de los DataNode y comienza la lectura del bloque



## - Fiabilidad y recuperación de datos en HDFS

- Los DataNode envían heartbeats al NameNode cada 3 segundos para indicar su estado.
- Si pasado un tiempo el NameNode no recibe el heartbeat de alguno de los nodos se da por perdido ese nodo.
- En caso de recuperar el nodo perdido, el sistema automáticamente decidirá qué bloques eliminar para mantener el número de replicación de bloques.
- Speculative execution

- **Web UI ->** A través de la web de UI del NameNode seremos capaces de ver el estado de nuestro cluster (Espacio total del HDFS, espacio ocupado del HDFS, espacio disponible del HDFS, estado de los nodos, navegación por el sistema de ficheros), se encuentra en el **puerto 50070**.