

Hadoop Distributed File System (HDFS)

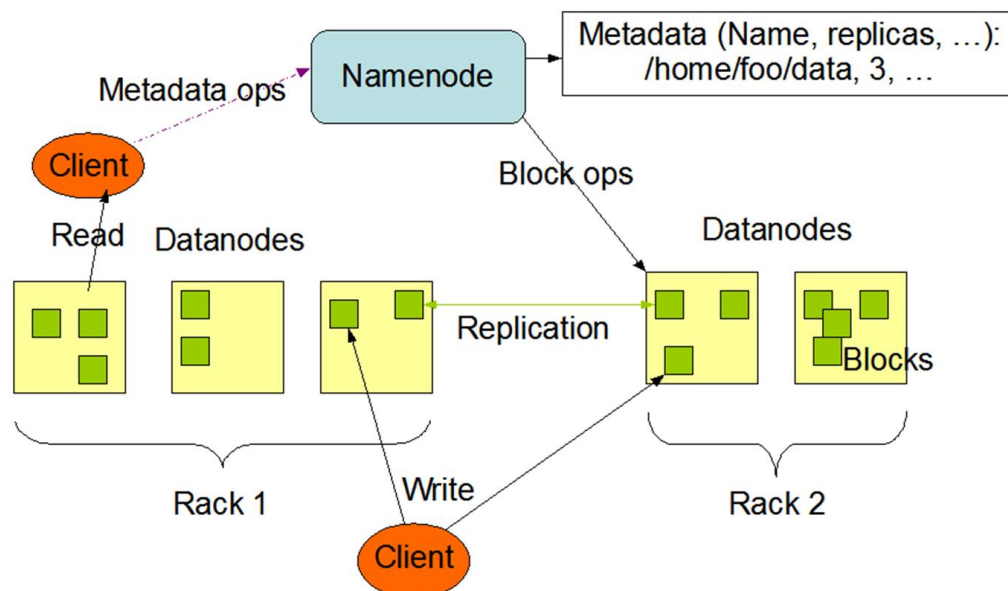
El Sistema de archivos HDFS es un sistema de archivos distribuido altamente tolerante a fallos, escalable y extremadamente fácil de ampliar. Fue diseñado para implementarse en hardware de bajo costo, está basado en GFS, se creó originalmente como infraestructura para el proyecto del motor de búsqueda web *Apache Nutch* pero ahora es un subproyecto *Apache Hadoop*.

HDFS tiene una arquitectura maestro/esclavo. Un clúster HDFS consta de:

- **NameNode:** servidor maestro que administra el espacio de nombres del sistema de archivos y regula el acceso de los clientes a los archivos, ejecuta operaciones de espacio de nombres del sistema de archivos como abrir, cerrar y renombrar archivos y directorios. También determina la asignación de bloques a DataNodes
- **DataNodes:** administran el almacenamiento adjunto a los nodos en los que se ejecutan, son responsables de atender las solicitudes de lectura y escritura de los clientes del sistema de archivos. repositorio de todos los metadatos de HDFS.

HDFS expone un espacio de nombres del sistema de archivos y permite que los datos del usuario se almacenen en archivos. Internamente, un archivo se divide en uno o más bloques y estos bloques se almacenan en un conjunto de DataNodes.

HDFS Architecture



NameNode y DataNode son piezas de software diseñadas para ejecutarse en máquinas de productos básicos. Estas máquinas normalmente ejecutan un sistema operativo GNU / Linux (OS). HDFS se construye utilizando el lenguaje Java; cualquier máquina que admita Java puede ejecutar el software NameNode o DataNode. Una implementación típica tiene una máquina dedicada que ejecuta solo el software NameNode. Cada una de las otras máquinas en el clúster ejecuta una instancia del software DataNode. La arquitectura no impide ejecutar múltiples DataNodes en la misma máquina, pero en una implementación real que rara vez es el caso.

HDFS también tiene un montón de características únicas que lo hacen ideal para sistemas distribuidos:

- **Tolerante a los fallos:** los datos se duplican en múltiples DataNodes para proteger contra fallas de la máquina. El valor predeterminado es un factor de replicación de 3 (cada bloque se almacena en tres máquinas). La cantidad de copias de un archivo se denomina factor de replicación de ese archivo. Esta información es almacenada por el NameNode.
- **Escalabilidad:** las transferencias de datos se realizan directamente con el DataNodes para que su capacidad de lectura / escritura se ajuste bastante bien con el número de DataNodes.
- **Espacio:** si se necesita más espacio en el disco solo se agregan más DataNodes y se aplica un re-balance.

HDFS es compatible con una organización de archivos jerárquica tradicional. HDFS aún no implementa cuotas de usuario. HDFS no soporta enlaces duros o enlaces blandos. Sin embargo, la arquitectura HDFS no impide la implementación de estas características.

HDFS está diseñado para almacenar de manera confiable archivos muy grandes en máquinas en un grupo grande. Almacena cada archivo como una secuencia de bloques; Todos los bloques de un archivo, excepto el último bloque, tienen el mismo tamaño. El tamaño de bloque y el factor de replicación son configurables por archivo. Una aplicación puede especificar el número de réplicas de un archivo. El factor de replicación se puede especificar en el momento de la creación del archivo y se puede cambiar más adelante. Los archivos en HDFS se escriben una vez y tienen estrictamente un escritor en cualquier momento.

El NameNode toma todas las decisiones con respecto a la replicación de bloques. Periódicamente recibe un **Heartbeat** y un **Blockreport** de cada uno de los DataNodes en el cluster. La recepción de un latido del corazón implica que el DataNode está funcionando correctamente. Un **Blockreport** contiene una lista de todos los bloques en un DataNode.

Referencias:

- https://www.tamps.cinvestav.mx/~vjsosa/clases/sd/DistributedFileSystems_gfs_hdfs.pdf
- https://hadoop.apache.org/docs/r1.2.1/hdfs_design.html