

Introducción al Almacenamiento Distribuido

Juan Mario Haut



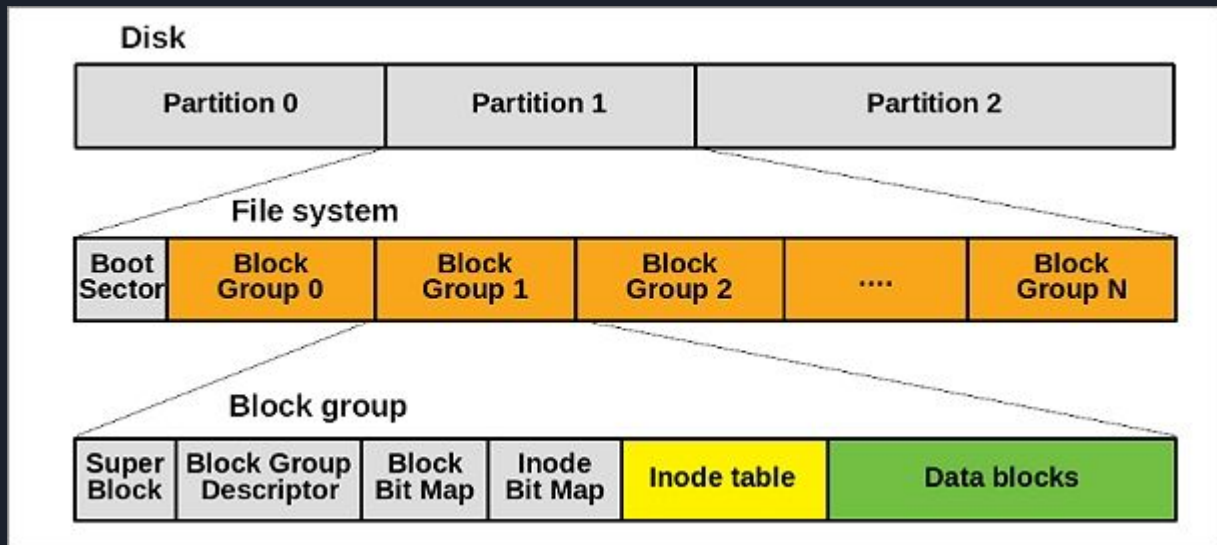
QUÉ ES UN SISTEMA DE FICHEROS



SISTEMA DE FICHEROS

Un sistema de archivos (file system o FS) en informática es el componente que gestiona cómo se almacenan y recuperan los datos en un dispositivo. Sin este sistema, los datos guardados en un medio de almacenamiento serían simplemente un conjunto desorganizado, sin una forma clara de identificar dónde termina un dato y comienza otro. Su función principal es administrar y optimizar el uso de las memorias, tanto internas como periféricas, ya sean secundarias o terciarias.

Las principales funciones de un sistema de archivos incluyen la asignación de espacio para los archivos, la gestión del espacio disponible y el control del acceso a los datos almacenados. Organizan la información en un dispositivo o unidad de almacenamiento, como un disco duro, para que luego pueda ser representada de forma textual o gráfica mediante un gestor de archivos.





Sistema de archivo distribuido: HDFS

- HDFS es el sistema de archivo distribuido de Hadoop.
- Diseñado para almacenamiento multinodo.
- Optimizado para almacenar ficheros de gran tamaño.
 - Write-Ones, Read many times.
- HDFS no necesita ser montado en hardware costoso. Está diseñado para funcionar en clústeres de bajo coste, ya que incorpora una estrategia robusta para tolerar fallos en cualquiera de los nodos. A medida que el clúster crece, aumenta la probabilidad de fallos, pero HDFS está preparado para gestionarlos eficientemente.



Sistema de archivo distribuido: HDFS

Aplicaciones NO recomendadas para HDFS:



Sistema de archivo distribuido: HDFS

Aplicaciones NO recomendadas para HDFS:

- Aplicaciones que requieran un acceso a datos con baja latencia.
- Aplicaciones que manejen una gran cantidad de archivos pequeños.
- Aplicaciones con múltiples usuarios escribiendo y realizando modificaciones arbitrarias en los archivos.

Apache Ozone <https://ozone.apache.org/>



Arquitectura HDFS

La arquitectura de HDFS es un modelo maestro-trabajador.

- MAESTRO (Namenode): Gestiona los metadatos del sistema de archivos. El NameNode mantiene un registro de dónde están almacenados los bloques de cada archivo en el sistema y coordina el acceso a ellos. Sin el NameNode, no se puede acceder a los datos almacenados.
- TRABAJADOR (Datanodes): Encargados de almacenar los bloques de datos. Los DataNodes almacenan y recuperan bloques según las solicitudes del NameNode. Cada archivo en HDFS se divide en bloques, y estos bloques se distribuyen entre los DataNodes para maximizar la tolerancia a fallos.



Conceptos fundamentales. Bloque

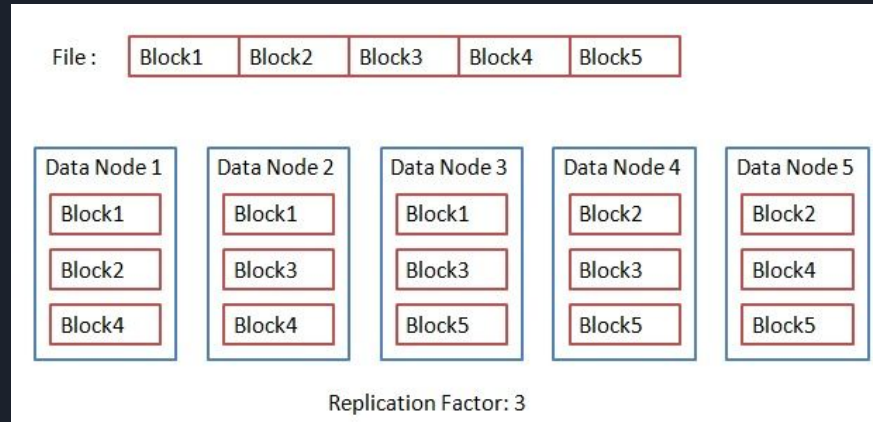
Los archivos en HDFS se fragmentan en bloques, que se almacenan en el sistema de archivos como unidades independientes.

- El bloque es la unidad mínima de lectura y escritura.
- En HDFS, los bloques tienen un tamaño considerable, siendo 128 MB por defecto (configurable).
- ¿De qué depende el tamaño del bloque?

Conceptos fundamentales. Bloque

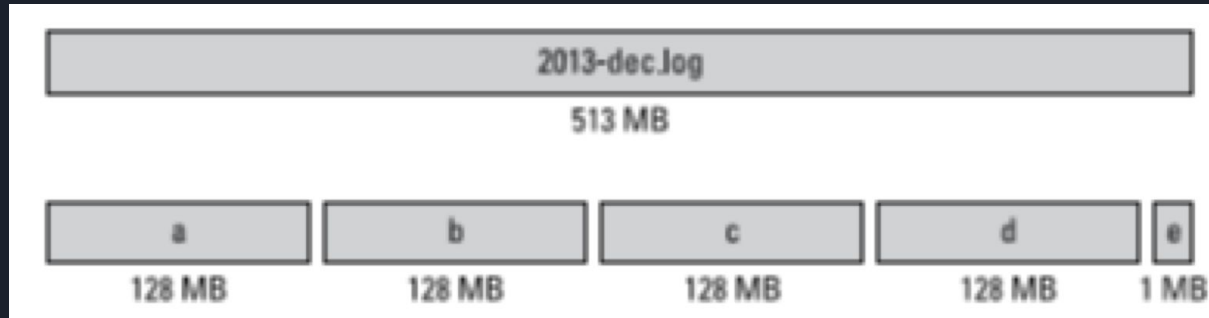
El uso de bloques facilita la gestión del almacenamiento y optimiza la tolerancia a fallos.

- Cada bloque de un archivo se replica en función de un factor de replicación, distribuyéndose entre las diferentes máquinas del clúster. Esto garantiza que, en caso de que alguna máquina falle, siempre habrá una copia disponible de ese bloque.

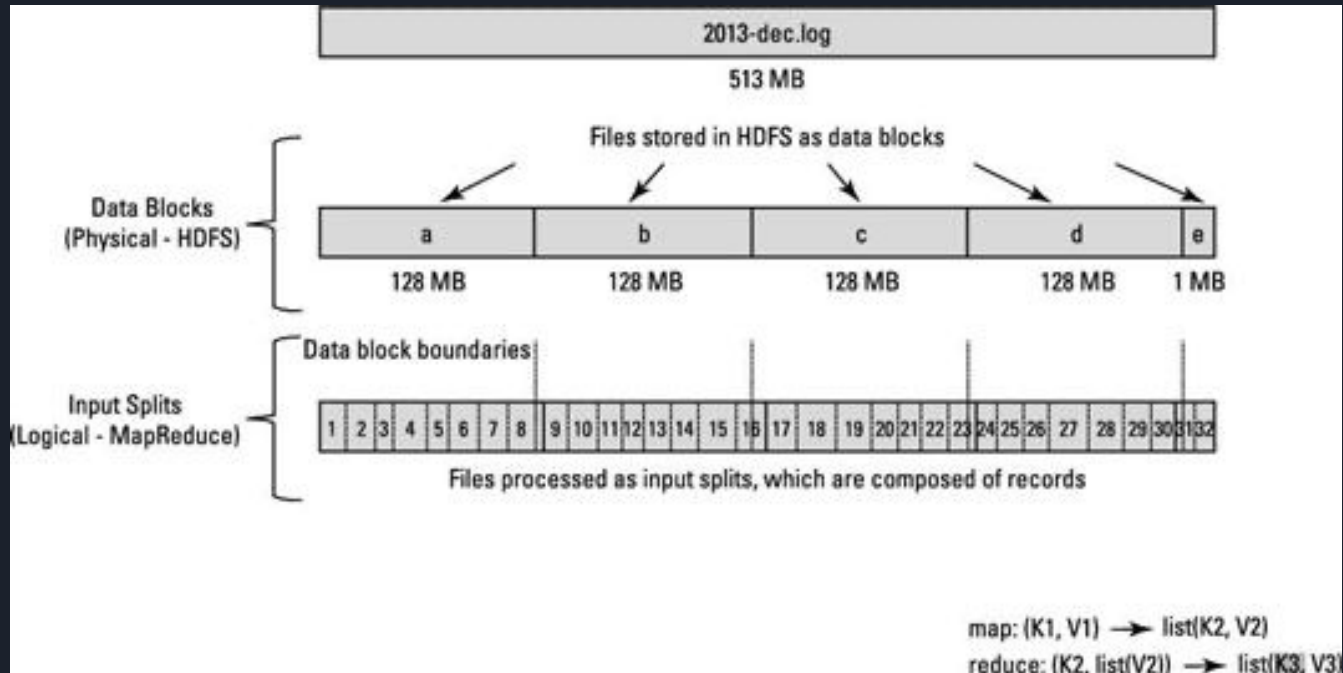


Conceptos fundamentales. Bloque

- Todos los bloques tienen el mismo tamaño excepto el último

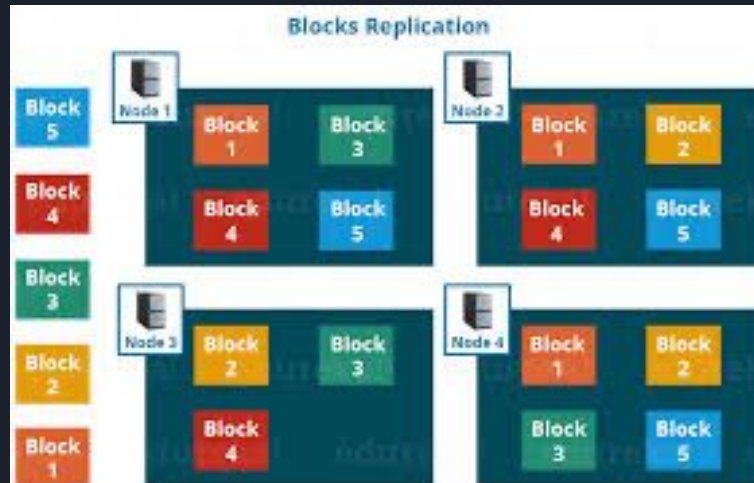


Conceptos fundamentales. Bloque



Conceptos fundamentales. Replicación

La **replicación** en el contexto de HDFS (Hadoop Distributed File System) es el proceso mediante el cual se crean múltiples copias (réplicas) de los bloques de datos en diferentes nodos del clúster. Esto asegura que, si un nodo falla o un bloque de datos se corrompe, las copias almacenadas en otros nodos puedan ser usadas para recuperar los datos sin pérdida de información ni interrupciones en el servicio.





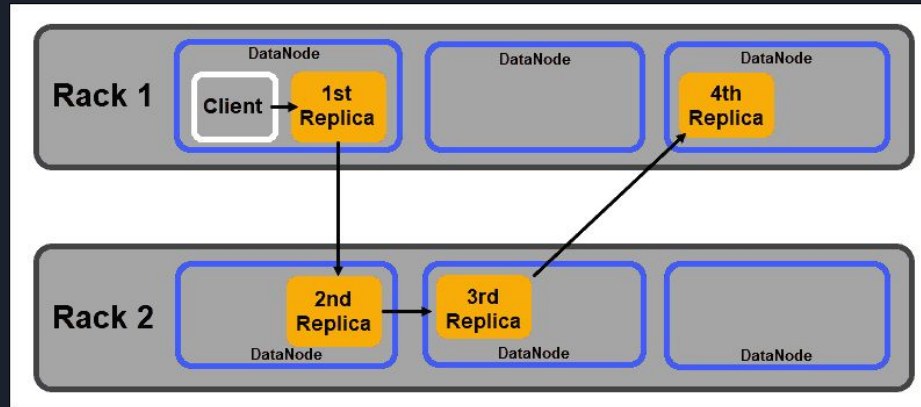
Conceptos fundamentales. Replicación

HDFS está diseñado para almacenar datos en hardware económico y con menor fiabilidad, por lo que implementa la replicación de bloques. En caso de que un nodo falle, se recupera una copia del bloque desde otro nodo del clúster.

- El factor de replicación es configurable; en arquitecturas de un solo nodo es 1, mientras que en clústeres por defecto es 3.

Conceptos fundamentales. Replicación

El primer bloque se almacena en uno de los nodos dentro de un rack, mientras que las otras dos copias se distribuyen en nodos de un rack diferente dentro del clúster. Esto asegura que siempre haya disponibilidad de los datos, incluso en caso de que todo un rack falle.





Conceptos fundamentales. Replicación

En HDFS, el sistema de replicación garantiza la disponibilidad y fiabilidad de los datos. Sin embargo, pueden surgir situaciones que afecten al número de copias de los bloques, como la sub-replicación o la sobre-replicación, las cuales deben ser gestionadas automáticamente para mantener el equilibrio en el sistema.

- **Sub-replicación:** Ocurre cuando el número de copias de un bloque es menor que el factor de replicación configurado. Cuando HDFS detecta esta situación, ordena la creación de una copia adicional del bloque para restaurar el número correcto de réplicas.

- **Sobre-replicación:** Ocurre cuando el número de copias de un bloque supera el factor de replicación establecido. En este caso, HDFS detecta el exceso de réplicas y ordena eliminar una o más copias adicionales del bloque para optimizar el uso del almacenamiento.

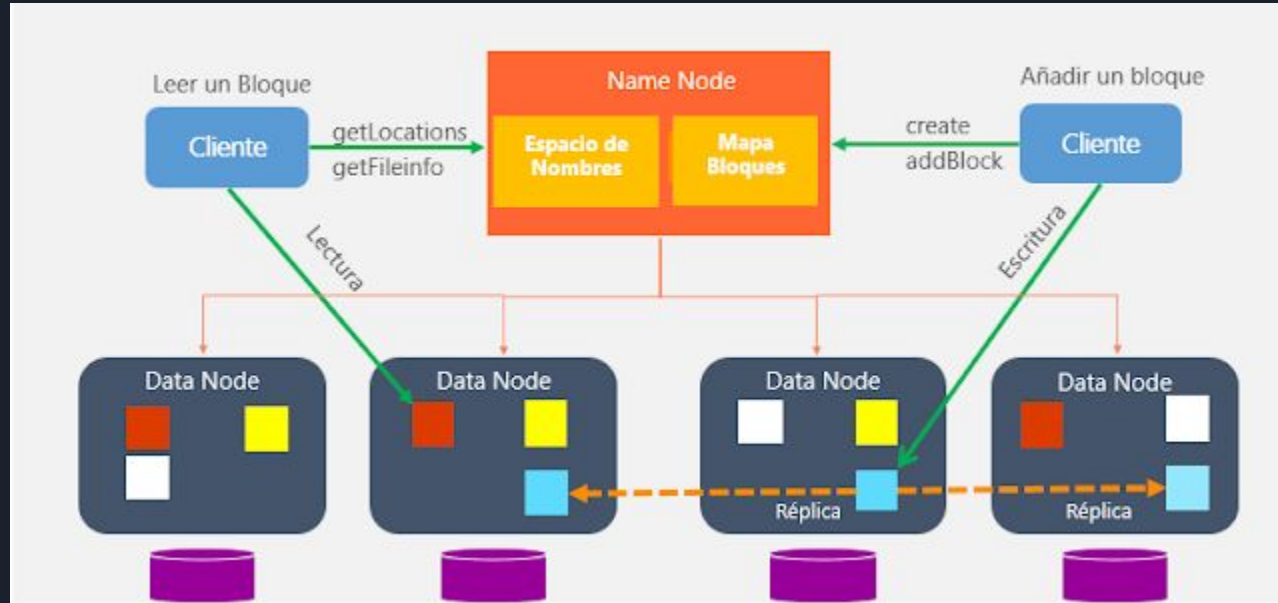



Conceptos fundamentales. Replicación

- **Bloques replicados:** Cada archivo en HDFS se divide en bloques, y cada bloque se replica en varios nodos (por defecto, el factor de replicación es 3).
- **Distribución de réplicas:** Las réplicas se distribuyen de forma que no estén en el mismo nodo ni en nodos cercanos, para garantizar la disponibilidad de los datos incluso si varios nodos fallan.
- **Recuperación ante fallos:** Si un nodo falla, HDFS accede a las réplicas de los bloques en otros nodos para continuar proporcionando acceso a los datos sin interrupciones.
- **Reequilibrio automático:** Cuando un nodo falla, HDFS automáticamente replica nuevamente los bloques que se perdieron a otros nodos, manteniendo siempre el número configurado de réplicas.

Arquitectura HDFS


La arquitectura de HDFS es un modelo maestro-trabajador.





Introducción al Almacenamiento Distribuido


Juan Mario Haut



Un archivo de 500 MB es cargado en HDFS, cuyo tamaño de bloque está configurado a 128 MB y el factor de replicación es 3.

- **Preguntas:**


1. ¿Cuántos bloques se crearán en HDFS para almacenar este archivo?
2. Si se configura el tamaño de bloque a 64 MB, ¿cuántos bloques se necesitarían ahora para almacenar el archivo?



El factor de replicación en un clúster HDFS está configurado a 3. Un archivo de 1 GB es almacenado en HDFS.

- **Preguntas:**


1. ¿Cuántos bloques totales se generarán en el clúster si el tamaño de bloque es 256 MB?
2. Si un DataNode falla y pierde uno de los bloques replicados, ¿cómo HDFS maneja la recuperación del bloque perdido?
3. ¿Qué sucede si se aumenta el factor de replicación a 4?



Un administrador de sistemas tiene un clúster HDFS con un tamaño de bloque de 256 MB y un factor de replicación de 3. En este clúster, está almacenando miles de archivos pequeños, cada uno de aproximadamente 1 MB de tamaño.

- **Preguntas:**

1. ¿Por qué es ineficiente almacenar muchos archivos pequeños en HDFS?
2. ¿Qué cambios recomendarías para optimizar el almacenamiento en este clúster para archivos pequeños?
3. ¿Qué efectos tendría el aumento del tamaño de bloque a 512 MB en el rendimiento del sistema?



Un clúster HDFS tiene un factor de replicación configurado a 3, pero tras varios fallos de nodos, algunos bloques del sistema están sub-replicados.


- **Preguntas:**

1. ¿Qué es la sub-replicación y cómo puede afectar al clúster?
2. Explica cómo HDFS detecta y maneja la sub-replicación.
3. ¿Qué es la sobre-replicación y qué impacto tiene en la eficiencia del sistema de archivos?
4. Proporciona un escenario en el que pueda producirse sobre-replicación.



Cuando un cliente intenta escribir un archivo en HDFS de 2 GB, HDFS lo divide en bloques de 128 MB, con un factor de replicación de 3.

- **Preguntas:**
 1. Describe el proceso completo que sigue HDFS desde que el cliente comienza a escribir el archivo hasta que los datos son replicados y almacenados en los DataNodes.
 2. Si se reduce el factor de replicación a 2, ¿cómo cambia este proceso?



Supón que tienes un clúster de HDFS con 20 DataNodes, cada uno con 2 TB de capacidad. El factor de replicación está configurado a 3 y el tamaño de bloque es 128 MB.

- **Preguntas:**

1. ¿Cuál es la capacidad total disponible en el clúster para almacenar datos, teniendo en cuenta la replicación?
2. Si se almacenan 5 TB de datos en el clúster, ¿cuánto espacio real en disco utilizarán los DataNodes debido a la replicación?
3. Si subo un fichero de 2TB. ¿Cuántas bloques hay en cada datanode?
4. ¿Cuántos bloques habría si el factor de replicación es 2?
5. ¿Cuantos nodos tienen un bloque de más?