

OpenCourseWare

Sistemas Paralelos y Distribuidos

Félix García Carballeira

Alejandro Calderón Mateos

Ejercicios de autoevaluación

Tema 7



1. ¿Cuál es la función principal del NameNode en la arquitectura de HDFS?
 - a. Comunicarse directamente con los clientes para transferir los datos de los ficheros.
 - b. Ejecutar las tareas de procesamiento de datos sobre los ficheros almacenados.
 - c. Gestionar el espacio de nombres del sistema de ficheros y regular el acceso a los ficheros.
 - d. Almacenar y replicar los bloques de datos de los ficheros.
2. ¿Cuál es el principal problema que introduce el uso de caché en los clientes de un sistema de ficheros distribuido visto en el curso?
 - a. Un aumento significativo de la latencia en el acceso a los datos.
 - b. La necesidad de emplear discos locales de gran capacidad en cada cliente.
 - c. El problema de la coherencia de caché cuando un fichero es compartido en escritura.
 - d. La sobrecarga del servidor debido a la gestión de múltiples cachés.
3. En el contexto de los Sistemas de Ficheros Distribuidos (SFD), ¿qué caracteriza al modelo de acceso de 'servicios remotos'?
 - a. El servidor proporciona todas las operaciones sobre el fichero, y el acceso se realiza por bloques.
 - b. El cliente accede directamente a los discos del servidor sin intermediarios.
 - c. El cliente descarga el fichero completo para trabajar con él localmente.
 - d. Combina el uso de una caché en el cliente con la descarga parcial de ficheros.
4. ¿Cuál es la estrategia del protocolo de coherencia de Sprite cuando se detecta un conflicto de acceso concurrente (por ejemplo, un escritor y varios lectores)?
 - a. Utiliza un sistema de tokens para otorgar permisos de escritura exclusivos.
 - b. Notifica a los clientes lectores para que actualicen sus copias en caché desde el escritor.
 - c. Desactiva la caché para ese fichero en todos los clientes implicados, forzando accesos remotos al servidor.
 - d. Permite que todos los clientes sigan usando su caché y sincroniza los cambios periódicamente.
5. ¿Qué ventaja principal ofrecen los Sistemas de Ficheros Paralelos (SFP) en comparación con los sistemas tradicionales?
 - a. Simplificar la gestión de ficheros al centralizar todos los metadatos en un único servidor.
 - b. Minimizar el espacio de almacenamiento requerido mediante compresión.
 - c. Incrementar el ancho de banda distribuyendo los ficheros y el acceso a través de múltiples discos y nodos de E/S.
 - d. Proporcionar baja latencia para un uso interactivo intensivo.
6. ¿Qué es la 'falsa contutilización' (*false sharing*) en el contexto de la granularidad de los protocolos de coherencia?
 - a. Cuando el sistema detecta un conflicto de escritura que en realidad no existe.
 - b. Cuando dos procesos creen que están modificando ficheros diferentes, pero en realidad es el mismo.
 - c. Cuando un cliente invalida su caché sin que los datos en el servidor hayan cambiado.
 - d. Dos procesos acceden a partes distintas de un fichero, pero esas partes caen dentro de la misma unidad de coherencia.

7. ¿Cuál es una característica distintiva de los sistemas de ficheros de discos compartidos?
 - a. Los datos de los ficheros siempre pasan a través de un servidor de E/S centralizado.
 - b. El sistema de ficheros se distribuye entre los discos locales de los nodos de cómputo.
 - c. Cada cliente tiene su propio almacenamiento local y lo comparte con otros a través de la red.
 - d. Los clientes acceden directamente al almacenamiento sin pasar por servidores de datos, aunque puede haber un servidor de metadatos (MDS).

8. Dentro de las políticas de actualización de caché, ¿qué ventaja ofrece la 'escritura diferida' (*write-back*) sobre la 'escritura inmediata' (*write-through*)?
 - a. Simplifica el protocolo de coherencia de caché al eliminar la necesidad de invalidaciones.
 - b. Ofrece mayor fiabilidad, ya que los datos se escriben en el servidor instantáneamente.
 - c. Garantiza que todos los clientes tengan siempre la versión más actualizada del fichero.
 - d. Las operaciones de escritura son más rápidas para la aplicación y se reduce el tráfico de red.

9. ¿Cuál es el objetivo principal de las optimizaciones de E/S en paralelo como 'Two-Phase I/O'?
 - a. Reorganizar los accesos a datos no contiguos para que coincidan con la disposición física en los discos y reducir el número de operaciones.
 - b. Cifrar los datos durante la transferencia entre los nodos de cómputo y los nodos de E/S.
 - c. Asegurar que las escrituras se realicen de forma secuencial por orden de rango del proceso.
 - d. Dividir cada fichero en múltiples subficheros, uno por proceso.

10. ¿Qué define a un sistema de ficheros paralelo ad-hoc, como Expand Ad-hoc?
 - a. Es una implementación hardware del sistema de ficheros para maximizar la velocidad.
 - b. Es un sistema de ficheros permanente diseñado para el almacenamiento a largo plazo en un supercomputador.
 - c. Es un sistema de almacenamiento intermedio y efímero que reduce la presión sobre el sistema de ficheros principal (back-end).
 - d. Es un sistema de ficheros que solo permite el acceso secuencial por un único proceso a la vez.