



TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

File systems

Instructor: Dr. Thanh-Chung Dao

Slides by Dr. Viet-Trung Tran

School of Information and Communication Technology

1

File systems



NTFS

Ext4
File System



NFS
NETWORK FILE SYSTEM

PVFS

The Google File System

By Sanjay Ghemawat, Howard Gobioff, and
Shun-Tak Leung
(Presented at SOSP 2003)

2

Overview

- Filenames
 - File Identity
- Directories (folders)
 - Group of files in separate collections
- Metadata
 - Creation time, last access time, last modification time
 - Security information (Owner, Group owner)
 - Mapping file to its physical location of file (e.g. location in storage devices)
- Computer file
 - A resource for storing information
 - Durable, remained available for access
 - Data: sequences of bits
- File system
 - Control how computer file are stored and retrieved
 - Main operators: READ, WRITE (offset, size), CREATE, DELETE

3

Local vs. distributed file systems

Local file systems



NTFS

Ext4
File System



NFS
NETWORK FILE SYSTEM

The Google File System

By Sanjay Ghemawat, Howard Gobioff, and
Shun-Tak Leung
(Presented at SOSP 2003)

PVFS

4

Distributed file system

- File system
 - Abstraction of storage devices
- Distributed file system
 - Available to remote processes in distributed systems
- Benefits
 - File sharing
 - Uniform view of system from different clients
 - Centralized administration

Hệ thống tập tin

- Tóm tắt các thiết bị lưu trữ

Hệ thống tệp phân tán

- Có sẵn cho các quy trình từ xa trong hệ thống phân tán

Những lợi ích

- Chia sẻ file
- Chế độ xem thống nhất của hệ thống từ các khách hàng khác nhau
- Quản trị tập trung

5

Mục tiêu: Mạng (Truy cập) Minh bạch

Goals: Network (Access) Transparency

- Network (Access) Transparency
 - Users should be able to access files over a network as easily as if the files were stored locally.
 - Users should not have to know the physical location of a file to access it.
- Transparency can be addressed through naming and file mounting mechanisms
 - Location Transparency: file name doesn't specify physical location
 - Location Independence: files can be moved to new physical location, no need to change references to them. (A name is independent of its addresses)
 - Location independence → location transparency, but the reverse is not necessarily true.

Tính minh bạch của Mạng (Truy cập)

- Người dùng có thể truy cập tệp qua mạng dễ dàng như thể tệp được lưu trữ cục bộ.
- Người dùng không cần phải biết vị trí thực của một tệp để truy cập nó.

Tính minh bạch có thể được giải quyết thông qua đặt tên và gắn tệp

- Tính minh bạch của vị trí: tên tệp không chỉ định vị trí thực tế
- Vị trí độc lập: các tệp có thể được di chuyển đến vị trí thực mới, không cần thay đổi các tham chiếu đến chúng. (Tên độc lập với địa chỉ của nó)
- Vị trí độc lập → tính minh bạch của vị trí, nhưng điều ngược lại không nhất thiết là đúng.

6

Goals: Availability

- Availability: files should be easily and quickly accessible.
- The number of users, system failures, or other consequences of distribution shouldn't compromise the availability.
- Addressed mainly through replication.

- Tính khả dụng: các tệp phải được truy cập dễ dàng và nhanh chóng.
- Số lượng người dùng, lỗi hệ thống hoặc các hậu quả khác của việc phân phối sẽ không ảnh hưởng đến tính khả dụng.
- Khắc phục chủ yếu thông qua bản sao

7

Architectures

- Client-Server
 - Sun Microsystem Network File System (NFS), Google File System (GFS)
 - Architecture
 - One or more machines (file servers) manage the file system.
 - Files are stored on disks at the servers
 - Requests for file operations are made from clients to the servers.
 - Client-server systems centralize storage and management; P2P systems decentralize it.
- Symmetric
 - Fully decentralized; based on peer-to-peer technology
 - e.g., Ivy (uses a Chord DHT approach)

- Đối xứng
- Phân cấp hoàn toàn; dựa trên công nghệ ngang hàng
- ví dụ: Ivy (sử dụng phương pháp DHT Chord)

- Ngành kiến trúc
- Một hoặc nhiều máy (máy chủ tập tin) quản lý hệ thống tập tin.
 - Các tệp được lưu trữ trên đĩa tại máy chủ
 - Các yêu cầu về hoạt động tệp được thực hiện từ máy khách đến máy chủ.
 - Hệ thống khách-máy chủ tập trung lưu trữ và quản lý; Hệ thống P2P phân quyền nó.

8

Design issues in distributed file systems

9

Design issues

- Naming and name resolution • Đặt tên và phân giải tên
- Semantics of file sharing • Ngữ nghĩa của chia sẻ tệp
- Caching
- Replication

10

Naming and name resolution

- A name space -- collection of names
- Name resolution -- mapping a name to an object
- 3 traditional ways
 - Concatenate the host name to the names of files stored on that host
 - Mount remote directories onto local directories
 - Provide a single global directory

Không gian tên - tập hợp các tên

Độ phân giải tên - ánh xạ tên với một đối tượng

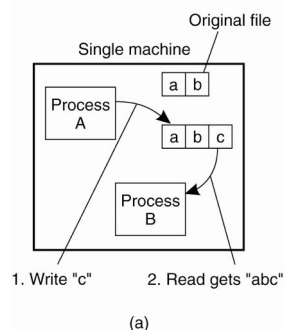
3 cách truyền thống

- Nối tên máy chủ với tên của các tệp được lưu trữ trên máy chủ đó
- Gắn các thư mục từ xa vào các thư mục cục bộ
- Cung cấp một thư mục chung duy nhất

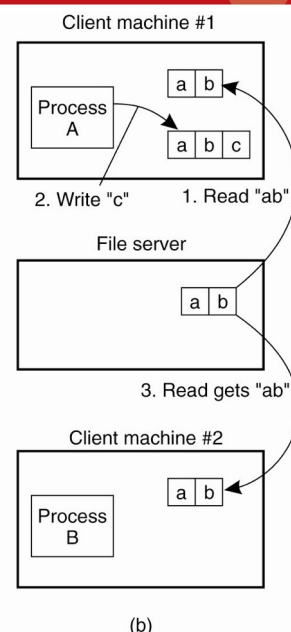
11

File Sharing Semantics (1/2)

- Problem: When dealing with distributed file systems, we need to take into account the ordering of concurrent read/write operations, and expected semantics (=consistency).



Vấn đề: Khi xử lý các hệ thống tệp phân tán, chúng ta cần tính đến thứ tự của các thao tác đọc / ghi đồng thời và ngữ nghĩa mong đợi (= nhất quán).



12

File Sharing Semantics (2/2)

• Assume open; reads/writes; close

- UNIX semantics: value read is the value stored by last write
Writes to an open file are visible immediately to others that have this file opened at the same time. Easy to implement if one server and no cache.

- Session semantics:
Writes to an open file by a user is not visible immediately by other users that have files opened already.
Once a file is closed, the changes made by it are visible by sessions started later.

- Immutable-Shared-Files semantics:
A sharable file cannot be modified.
File names cannot be reused and its contents may not be altered.
Simple to implement.

- Transactions: All changes have all-or-nothing property. $W1, R1, R2, W2$ not allowed where $P1 = W1; W2$ and $P2 = R1; R2$

• **Ngữ nghĩa phiên:**
Việc ghi vào tệp đang mở bởi người dùng không hiển thị ngay lập tức bởi những người dùng khác đã mở tệp.
Sau khi đóng tệp, các thay đổi do tệp thực hiện sẽ hiển thị trong các phiên bắt đầu sau đó.

• **Ngữ nghĩa UNIX:** giá trị đọc là giá trị được lưu trữ bởi lần ghi cuối cùng
Các ghi vào một tệp đang mở sẽ hiển thị ngay lập tức cho những người khác mở tệp này cùng lúc.
Dễ thực hiện nếu có một máy chủ và không có bộ nhớ cache.

• **Ngữ nghĩa của tệp tin chia sẻ bất biến:**
Không thể sửa đổi tệp có thể chia sẻ.
Không thể sử dụng lại tên tệp và nội dung của tệp không được thay đổi.
Đơn giản để thực hiện.

• **Giao dịch:** Tất cả các thay đổi đều có thuộc tính tất cả hoặc không có gì. $W1, R1, R2, W2$ không được phép khi $P1 = W1; W2$ và $P2 = R1; R2$

13

Caching

• Server caching: in main memory

- cache management issue, how much to cache, replacement strategy
- still slow due to network delay
- Used in high-performance web-search engine servers

• Client caching in main memory

- can be used by diskless workstation
- faster to access from main memory than disk
- compete with the virtual memory system for physical memory space

• Client-cache on a local disk

- large files can be cached
- the virtual memory management is simpler
- a workstation can function even when it is disconnected from the network

Bộ nhớ đệm máy khách trong bộ nhớ chính
• có thể được sử dụng bởi máy trạm không đĩa
• truy cập từ bộ nhớ chính nhanh hơn so với đĩa
• cạnh tranh với hệ thống bộ nhớ ảo về không gian bộ nhớ vật lý

Bộ nhớ đệm máy chủ: trong bộ nhớ chính
• vấn đề quản lý bộ nhớ cache, dung lượng bộ nhớ cache, chiến lược thay thế
• vẫn chậm do mạng chậm trễ
• Được sử dụng trong các máy chủ công cụ tìm kiếm web hiệu suất cao

Bộ đệm ứng dụng khách trên đĩa cục bộ
• các tệp lớn có thể được lưu vào bộ nhớ đệm
• quản lý bộ nhớ ảo đơn giản hơn
• một máy trạm có thể hoạt động ngay cả khi nó bị ngắt kết nối khỏi mạng

14

Cân bằng

Caching tradeoffs

- Reduces remote accesses => reduces network traffic and server load
Giảm truy cập từ xa => giảm lưu lượng mạng và tải máy chủ
- Total network overhead is lower for big chunks than a series of responses to specific requests.
Tổng chi phí mạng đối với khối lượng lớn dữ liệu (bộ nhớ đệm) thấp hơn so với một loạt phản hồi cho các yêu cầu cụ thể
- Disk access can be optimized better for large requests than random disk blocks
Quyền truy cập đĩa có thể được tối ưu hóa tốt hơn cho các yêu cầu lớn hơn so với các khối đĩa ngẫu nhiên
- Cache-consistency problem is the major drawback. If there are frequent writes, overhead due to the consistency problem is significant.
Vấn đề về tính nhất quán của bộ nhớ đệm là nhược điểm lớn. Nếu thường xuyên phải viết, chi phí do vấn đề nhất quán là đáng kể.

15

Replication

- File data is replicated to multiple storage servers
- Goals
 - Increase reliability
 - improve availability
 - balance the servers workload
- How to make replication transparent?
- How to keep the replicas consistent?
 - a replica is not updated due to its server failure
 - network partitioned

Dữ liệu tệp được sao chép sang nhiều máy chủ lưu trữ

Mục tiêu

- Tăng độ tin cậy
- cải thiện tính khả dụng
- cân bằng khối lượng công việc của máy chủ

Làm thế nào để nhân rộng minh bạch?

Làm thế nào để giữ cho các bản sao nhất quán?

- một bản sao không được cập nhật do lỗi máy chủ của nó
- mạng được phân vùng

16