

25 YEARS ANNIVERSARY
SOKT

ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG



ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

Chương 3

Hệ thống tập tin phân tán

Hadoop HDFS

1. Hệ thống tệp tin

Các hệ thống tệp tin



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)

Các khái niệm

- Tên tệp - Filename
 - Là định danh tệp tin - File Identity
- Thư mục - Directories (folders)
 - Nhóm các tệp tin trong 1 tập hợp riêng
- Siêu dữ liệu - Metadata
 - Thời gian khởi tạo, thời gian truy cập cuối cùng, thời gian chỉnh sửa cuối cùng
 - Các thông tin bảo mật (chủ sở hữu, nhóm sở hữu)
 - Thông tin ánh xạ tệp tin tới vị trí lưu trữ vật lý của tệp tin trên thiết bị lưu trữ (
- Tệp máy tính - Computer file
 - Là một tài nguyên để lưu trữ thông tin
 - Có tính bền vững, tồn tại lâu dài
 - Dữ liệu: là chuỗi các bit thông tin
- Hệ thống tệp tin - File system
 - Quản lý cách thức mà các tệp tin máy tính được lưu trữ và truy cập
 - Các thao tác chính: READ, WRITE (offset, size), CREATE, DELETE

Hệ thống tệp tin phân tán vs. cục bộ

Hệ thống tệp cục bộ



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

Hệ thống tệp tin phân tán

- Là hệ thống tệp tin cho phép truy cập trên môi trường phân tán
- Lợi ích
 - Chia sẻ tệp tin
 - Cung cấp góc nhìn hợp nhất của hệ thống cho các ứng dụng khách (clients) khác nhau
 - Quản lý tập chung

Mục tiêu: Truy cập “trong suốt”

- Truy cập “trong suốt”
 - Người cung cấp có thể truy cập tới tệp tin qua mạng dễ dàng như các tệp tin này được lưu trữ cục bộ
 - Người dùng không nhất thiết cần biết vị trí vật lý của tệp tin để truy cập tới tệp tin đó
- Tính “trong suốt” có thể được đảm bảo thông qua kỹ thuật định danh và “mount”
 - Location Transparency: Tên tệp không chứa thông tin về vị trí vật lý của tệp tin
 - Location Independence: Tệp tin có thể được di rời qua vị trí vật lý mới mà không phải thay đổi tham chiếu tới nó (Tên tệp độc lập với vị trí lưu trữ vật lý của tệp)
 - Location independence → location transparency, nhưng ngược lại không được đảm bảo

Mục tiêu: Tính sẵn sàng (Availability)

- Tính sẵn sàng: Tập tin cần luôn luôn được truy cập dễ dàng và nhanh chóng
- Số lượng người dùng, hỏng hóc hệ thống, hay các hệ quả của sự phân tán cần không ảnh hưởng tiêu cực tới tính sẵn sàng
- Thường được giải quyết thông qua cơ chế nhân bản

Các kiến trúc chung

- Chủ khách /Client-Server
 - Một hay nhiều máy chủ tham gia quản lý hệ thống tệp
 - Các tệp được lưu trữ trên thiết bị lưu trữ của máy chủ
 - Các thao tác tới tệp được gửi từ máy khách tới máy chủ
 - Ví dụ: Sun Microsystem Network File System (NFS), Google File System (GFS)
- Đối xứng / Symmetric
 - Phân tán hoàn toàn, dựa trên công nghệ mạng ngang hàng
 - Kiến trúc chủ – khách tập trung hoá sự lưu trữ và quản lý, kiến trúc ngang hàng phân tán hoá
 - Ví dụ Ivy (sử dụng Chord DHT)

Các vấn đề về thiết kế trong hệ thống tệp tin phân tán

Các vấn đề về thiết kế

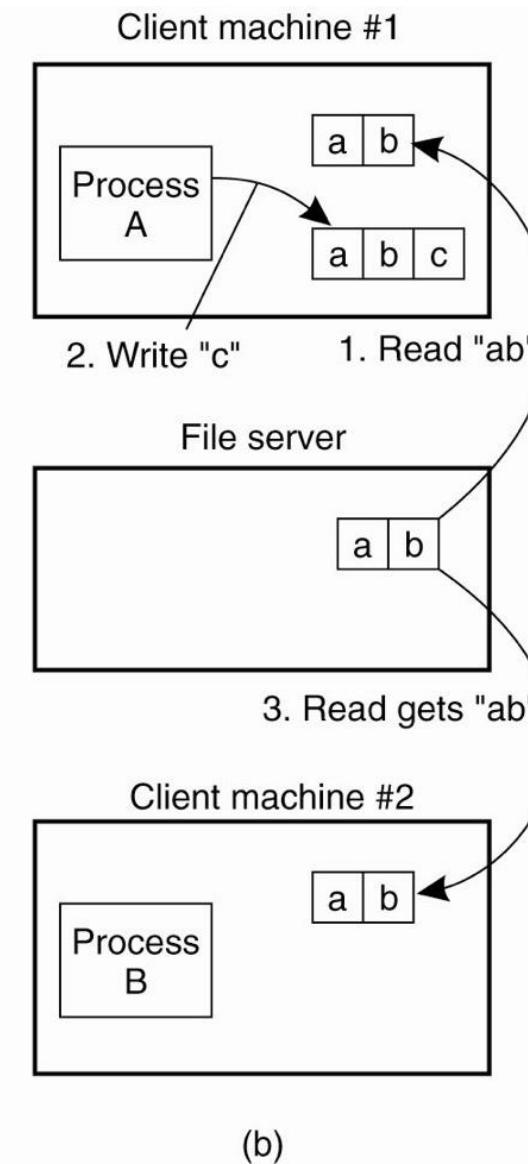
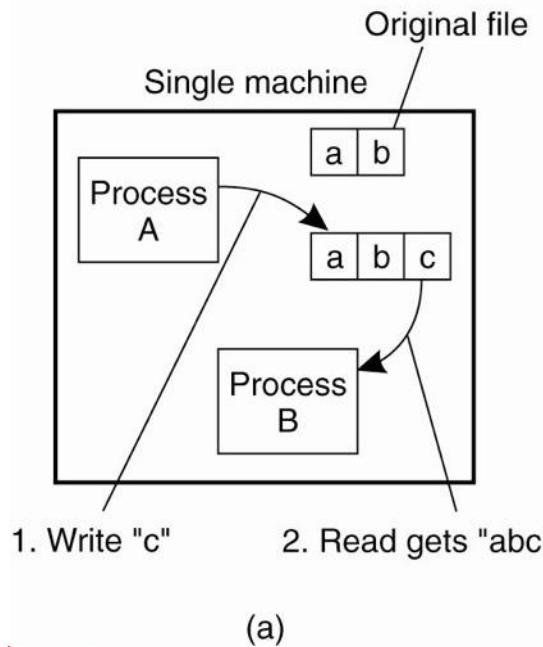
- Định danh và phân giải tên (định danh)
- Lựa chọn mô hình semantic khi có sự tương tranh
- Kỹ thuật vùng đệm
- Kỹ thuật nhân bản

Định danh và phân giải tên

- Không gian tên / name space: tập hợp các tên khả dụng
- Phân giải tên: ánh xạ từ tên tệp tới một đối tượng (thường là vị trí của dữ liệu tương ứng)
- Có 3 hướng tiếp cận
 - Nối tên máy chủ với tên tệp được lưu trữ trong máy chủ đó
 - Mount một thư mục từ xa thành thư mục trong hệ thống tệp tin cục bộ
 - Cung cấp 1 cấu trúc cây thư mục toàn cục duy nhất

Lựa chọn mô hình semantic khi có sự tương tranh [1]

- Trong hệ phân tán, cần xem xét lựa chọn mô hình semantic phù hợp (tính nhất quán được đảm bảo) khi có các thao tác đọc / ghi tương tranh



Lựa chọn mô hình semantic khi có sự tương tranh [2]

- Giả định các bước thao tác với tệp tin:
 - Mở; các thao tác đọc / ghi; đóng tệp tin
- **UNIX semantics:**
 - Giá trị đọc được là giá trị thuộc về thao tác ghi cuối cùng. Thao tác ghi tới 1 tệp đang mở được thấy ngay lập tức tới các thao tác khác mà đã mở tệp đó ở cùng thời điểm. Dễ thực hiện nếu có 1 máy chủ và không có cơ chế bộ đệm
- **Session semantics:**
 - Thao tác ghi tới một tệp đang mở không cần phải được thấy ngay lập tức bởi các thao tác của người khác đã mở tệp đó. Khi tệp được đóng lại thì các thay đổi mới cần phải được thấy bởi các phiên mở tệp tiếp theo.
- **Immutable-Shared-Files semantics:**
 - Tệp không sửa và cập nhật được. Dễ thực hiện theo cơ chế này vì đơn giản.
- **Transactions:**
 - Mọi thay đổi tới tệp có tính chất giao dịch
 - W1,R1,R2,W2 không được phép nếu $P1 = W1;W2$ and $P2 = R1;R2$

Sử dụng bộ nhớ đệm

- Lưu đệm tại máy chủ, sử dụng bộ nhớ trong của máy chủ
 - Các vấn đề liên quan đến quản lý vùng đệm nói chung: lưu đệm bao nhiêu, chiến lược thay thế
 - Còn chậm do độ trễ giao tiếp qua mạng tới máy chủ
 - Ví dụ: Sử dụng trong các máy chủ tìm kiếm hiệu năng cao
- Lưu đệm tại phía máy khách, trong bộ nhớ trong
 - Có thể được sử dụng trên các máy trạm không có thiết bị lưu trữ ngoài
 - Cải thiện tốc độ
- Lưu đệm tại phía máy khách, sử dụng bộ nhớ ngoài
 - Có thể lưu đệm được các tệp lớn
 - Có thể tiếp tục truy xuất dữ liệu ngay cả khi đứt kết nối tới máy chủ

Sự đánh đổi khi sử dụng bộ nhớ đệm

- Giảm thiểu truy xuất từ xa -> giảm thiểu truyền thông trên mạng và giảm tải cho máy chủ
- Chi phí mạng (network overhead) thấp hơn khi chuyển gói dữ liệu kích thước lớn so với chuyển nhiều gói dữ liệu kích thước nhỏ
- Truy cập ổ đĩa cũng tối ưu hơn với kích thước lớn thay vì nhiều thao tác nhỏ, ở vị trí khác nhau
- Vấn đề nhất quán bộ nhớ đệm là yếu điểm quan trọng. Khi có nhiều thao tác ghi dữ liệu, giá cần phải trả để đảm bảo sự nhất quán là không nhỏ

Sự nhân bản

- Tệp dữ liệu có thể được nhân bản tới nhiều máy chủ lưu trữ khác nhau
- Mục tiêu
 - Tăng độ tin cậy
 - Tăng tính sẵn sàng
 - Cân bằng tải cho các máy chủ lưu trữ
- Các thách thức
 - Làm thế nào để sự nhân bản là trong suốt với người dùng?
 - Làm thế nào để đảm bảo sự nhất quán giữa các nhân bản?
 - Khi có hỏng hóc ở máy chủ
 - Khi có sự phân tách mạng



25
YEARS ANNIVERSARY
SOICT

VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG
SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

Chân thành
cảm ơn!!!

