

Hệ điều hành phân tán

- Giới thiệu về hệ thống phân tán
- Liên lạc trong hệ thống phân tán
- Vấn đề đồng bộ trong hệ thống phân tán
- Quản lý tiến trình và bộ xử lý
- Hệ thống tệp phân tán

Xu thế phân tán

- Từ 1945-1985: máy tính mainframe, mini. Các hệ thống này thường chạy đơn lẻ.
- Từ 1980: xuất hiện Microprocessors, Local Area Network (Mbs,Gbs)

Luật Grosch

- Dùng với mô hình mainframe cổ điển tính toán tập trung
 - khả năng tính toán tỷ lệ thuận với bình phương chi phí => xu thế tập trung hoá
- Không còn đúng cho mô hình micro processor
 - chi phí lớn hơn chỉ đem lại cùng loại CPU với tần số hoạt động lớn hơn; hiệu suất tính toán không tăng đáng kể
- Áp dụng với mô hình tính toán phân tán
 - tập hợp khả năng tính toán của nhiều LAN PC
 - tập hợp khả năng tính toán của nhiều CPU trên 1 máy

Ưu điểm hệ thống phân tán so với hệ thống tập trung

- Kinh tế: chi phí/hiệu suất cao hơn
- Tốc độ: Khả năng tính toán có thể đạt được hoặc hơn hệ thống mainframe
- Bản chất nhiều ứng dụng phân tán
- Độ tin cậy cao hơn
- Có thể phát triển mở rộng

Ưu điểm hệ thống phân tán so với PC độc lập

- Chia sẻ tài nguyên (dữ liệu, thiết bị)
- Liên lạc giữa người dùng
- Chia tải công việc

Nhược điểm hệ thống phân tán

- Phần mềm chưa sẵn sàng
- Mạng có thể tắc nghẽn
- An toàn hệ thống khó khăn hơn

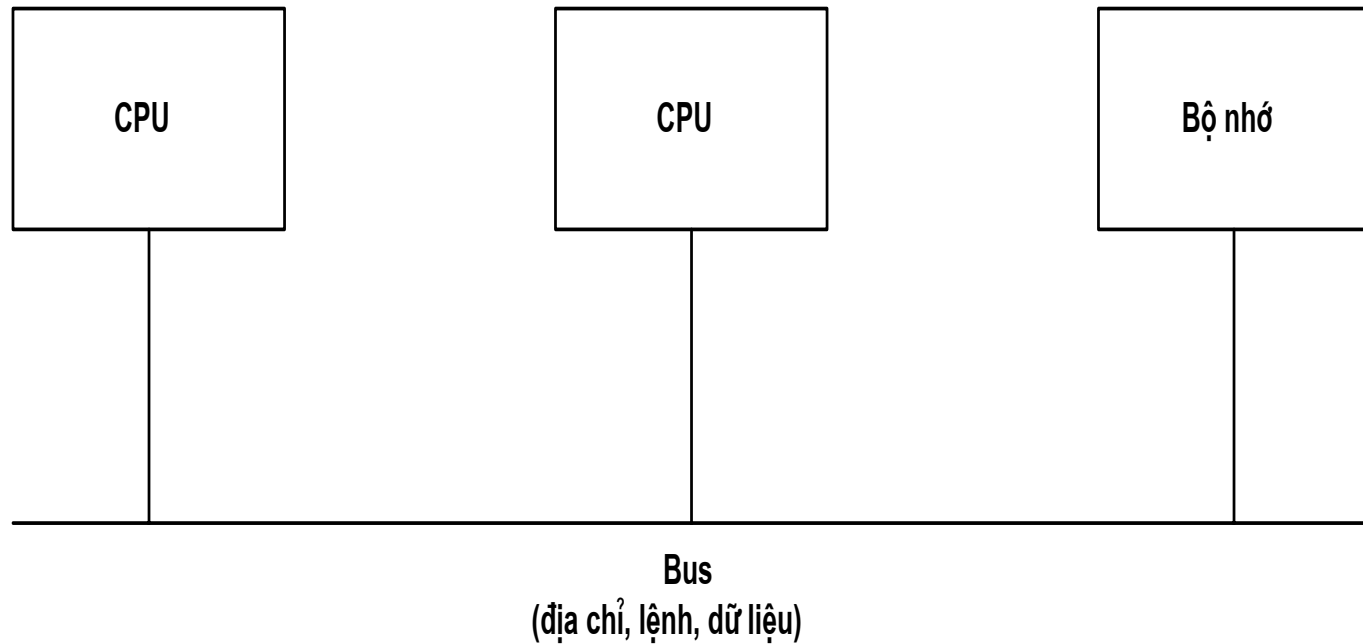
Khía cạnh phần cứng

- Phân loại:
 - SISD: Single Instruction Single Data stream. Ví dụ: PC đơn lẻ.
 - SIMD: Single Instruction Multiple Data stream. Ví dụ: tính toán với dữ liệu lặp.
 - MISD: Multiple Instruction Single Data stream. Không có loại máy nào theo kiểu này.
 - MIMD: Multiple Instruction Multiple Data stream. Ví dụ: tính toán song song và phân tán trên các máy.
- Khái niệm phân tán theo nghĩa rộng: cả multiprocessor và multicomputer

...khía cạnh phần cứng

- Phân loại theo kiến trúc kết nối:
 - kết nối dạng bus
 - kết nối dạng switch
 - Phân loại theo mức độ liên lạc:
 - kết nối chặt (tightly coupled)
 - kết nối lỏng (loosely coupled)
- Khó có thể định ranh giới rõ ràng chặt/lỏng

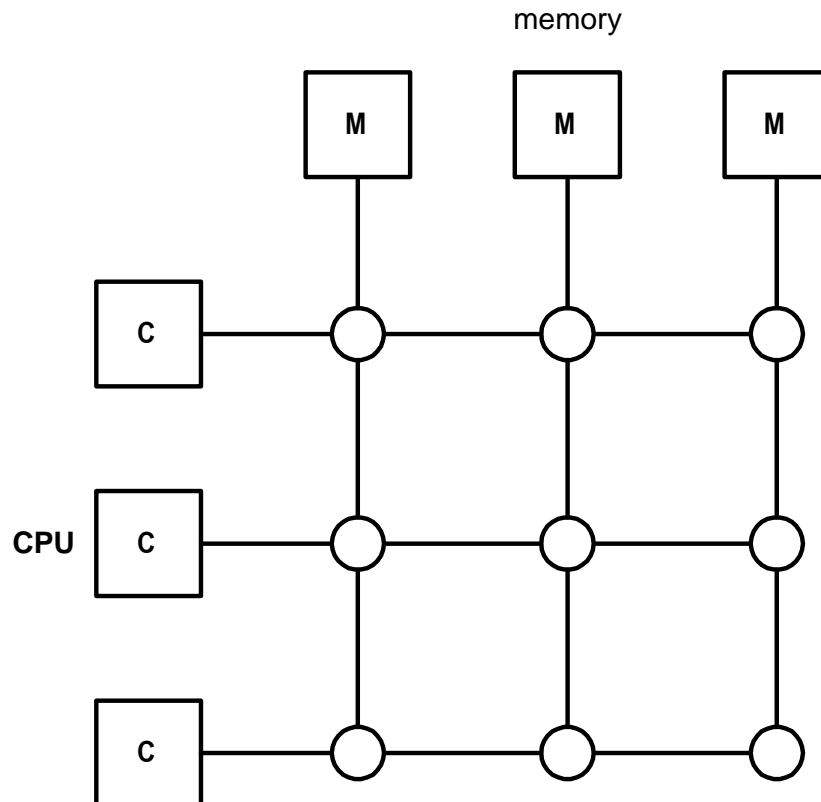
Đa xử lý kết nối dạng bus



...đa xử lý kết nối dạng bus

- Nhiều CPU, chia sẻ chung bộ nhớ
- Bus địa chỉ, dữ liệu, lệnh:
 - địa chỉ được đặt trên bus địa chỉ, lệnh được gửi qua bus lệnh và CPU nhận được dữ liệu qua bus dữ liệu
- Tính nhất quán dữ liệu trong đọc/ghi
- Tạo cache trong mỗi CPU tránh quá tải bus
- Write-through snoopy cache

Đa xử lý kết nối rẽ nhánh

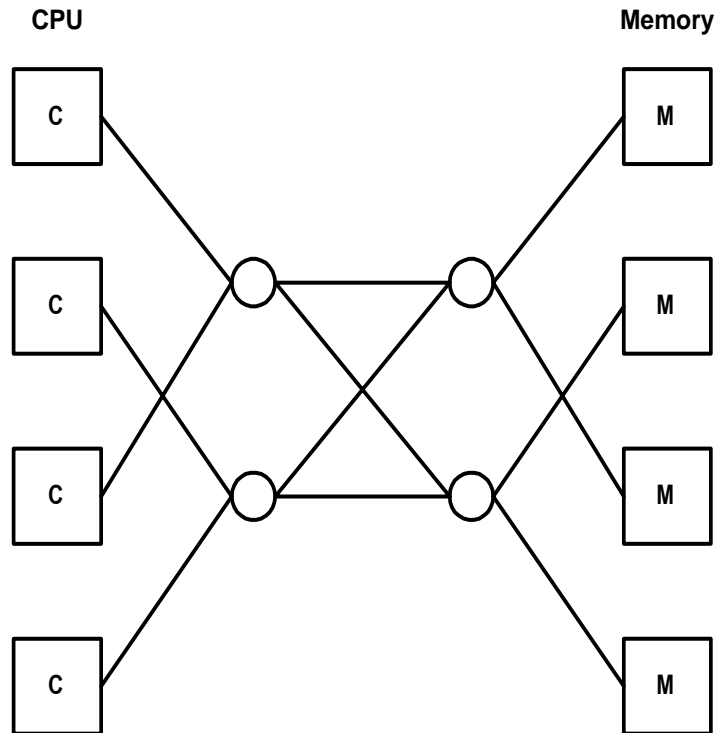


- Crossbar switch

- Với n CPU, n memory cần tới n^2 nút

- Có thể truy nhập song song nếu không trùng

...đa xử lý kết nối rẽ nhánh



- Omega switch
 - $O(1/2n\log_2n)$ nút
 - Thời gian dịch chuyển tương đối lâu hơn: qua \log_2n nút chuyển mạch

...đa xử lý kết nối rẽ nhánh

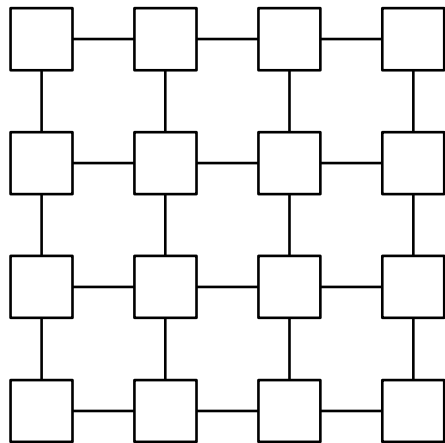
- NUMA: Non Uniform Memory Access
 - xây dựng theo cấu trúc phân cấp, chia sẻ một vùng nhớ chung giữa một số CPU
 - Mỗi CPU có bộ nhớ riêng thâm nhập nhanh, bộ nhớ chia sẻ thâm nhập chậm hơn
 - Ưu điểm: thời gian truy nhập tốt hơn, chi phí ít hơn
 - khó xây dựng phần mềm để cục bộ hoá mã chương trình và dữ liệu đối với mỗi CPU

Đa trạm dựa trên bus

- Mô hình LAN
 - mỗi máy có bộ nhớ riêng và tất nhiên CPU riêng

Đa trạm kết nối rẽ nhánh

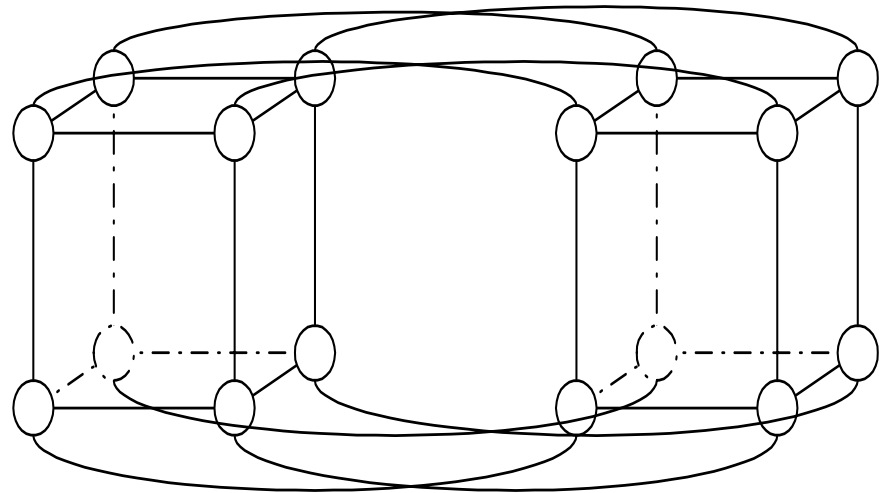
- Grid



- Thích hợp với 1 số bài toán: eye

- Thường gắn trên bảng mạch

- Hypercube



- Một nút nối với n nút khác

- Đường đi tín hiệu: $O(\log n)$

- Đã có với 1024 CPUs

Khía cạnh phần mềm

- Khó phân biệt tightly và loosely coupled
- Kết hợp với khía cạnh phần cứng => có các loại hệ thống:
 - phần cứng lỏng + phần mềm lỏng
 - phần cứng lỏng + phần mềm chặt
 - phần cứng chặt + phần mềm chặt

Hệ thống quản lý mạng

- Khái niệm hệ điều hành mạng:
 - cho phép quản lý tài nguyên thống nhất trong hệ thống mạng
 - Ví dụ:
 - Netware file server: quản lý tài nguyên tệp, máy in cho nhiều người dùng tham nhập từ máy trạm
 - CIFS, SAMBA: tương tự như Netware filer nhưng dựa trên giao thức SMB tương thích Win
 - Hệ thống quản lý tệp phân tán NFS trên Unix

Hệ thống quản lý tệp phân tán

- NFS (Sun micro system)
 - dựa trên giao thức gọi thủ tục từ xa (RPC) và giao thức biểu diễn dữ liệu (XDR)
 - hoàn toàn trong suốt đối với người dùng cuối

XDR	
RPC	
UDP	TCP
IP	
Vật lý và LKDL	

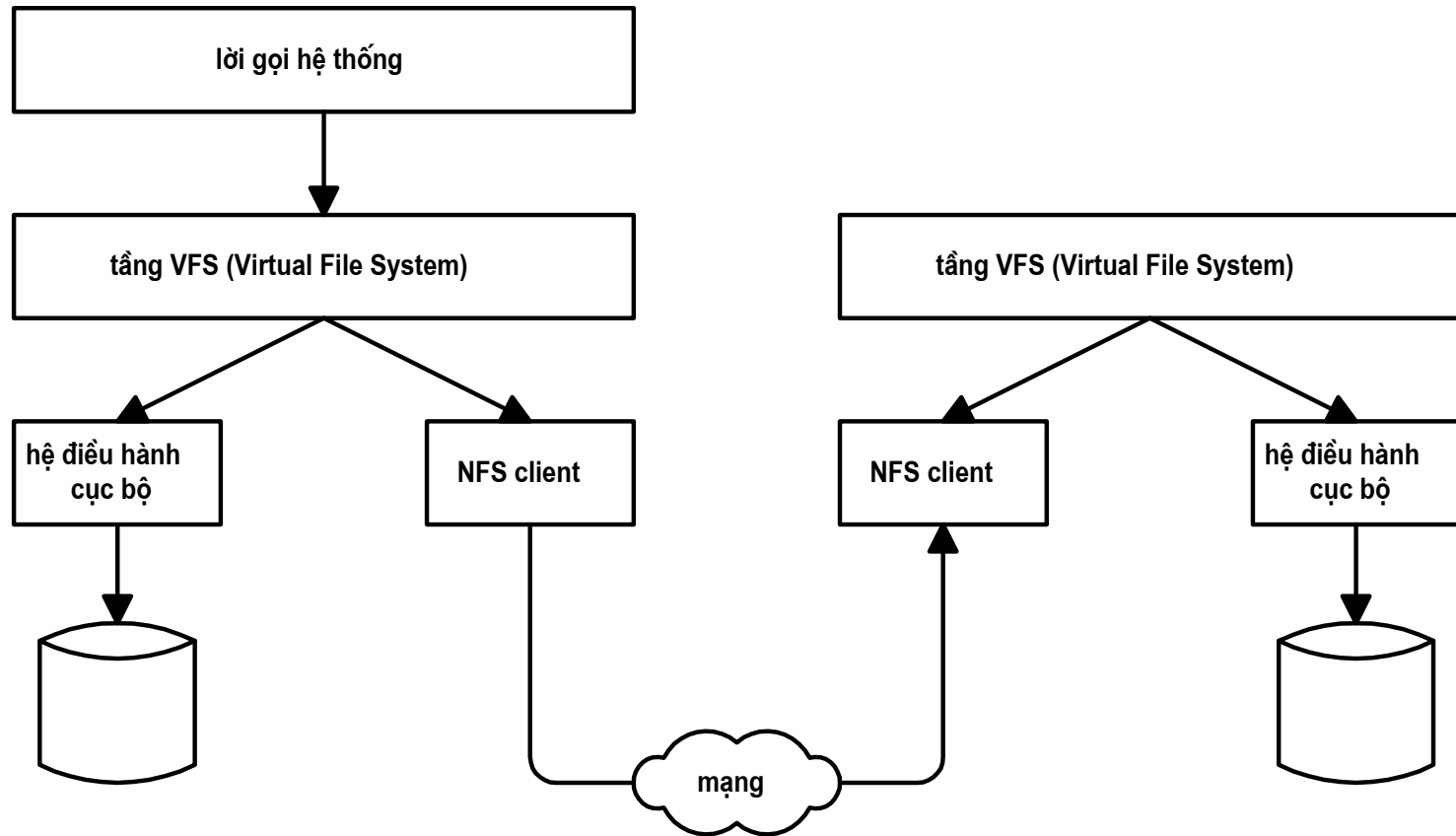
NFS – phương thức chia sẻ

- Server:
 - tạo danh sách các thư mục có thể chia sẻ trong */etc/exports*
- Client:
 - kết nối thư mục ở xa vào cây thư mục qua lệnh *mount*
 - gắn kết tự động phía khách được khai báo trong tệp */etc/rc*

NFS – cài đặt

- Tầng VFS:
 - tạo ra v-node (virtual node)
 - v-node trỏ tới i-node hoặc r-node
 - các thao tác tệp chỉ làm việc với v-node, do vậy không phân biệt được tệp là cục bộ hay ở xa

...NFS – cài đặt



...NFS-cài đặt

- Stateless server:
 - chỉ cung cấp file-handle cho client
 - file-handle gồm i-node và các thuộc tính
 - không lưu trữ bảng các files đang mở, do vậy là stateless
 - server có thể hỏng và khởi động nhanh lại mà không gây ảnh hưởng
- Ngược lại: RFS Unix system V có lưu trạng thái các tệp mở, chỉ số tệp

NFS-ngữ nghĩa thâm nhập

- NFS client thường sử dụng kỹ thuật read-ahead và caching để tăng tốc
- Mọi thay đổi tệp thường được cập nhật sau 30 giây
- Mọi thay đổi thư mục thường được cập nhật sau 3 giây

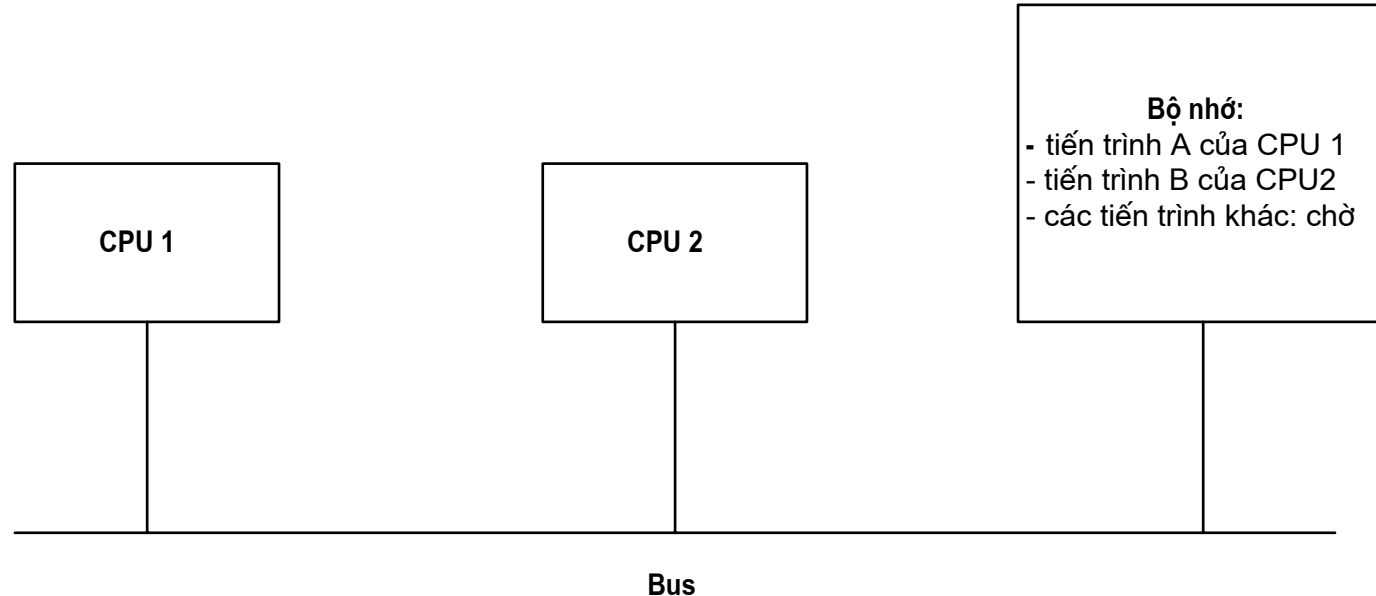
NFS-bảo mật

- Ban đầu:
 - client gửi UID, GID tới server và server gửi kết quả
 - server tin tưởng ở client
- Lựa chọn bảo mật:
 - public-key được áp dụng giữa client và server
 - NIS là bên thứ 3 cung cấp public-key

Hệ phân tán thực sự

- Phần cứng lỏng – phần mềm chặt
- Định nghĩa: chạy trên 1 nhóm máy không có bộ nhớ chung, tạo cho người dùng hình ảnh của 1 máy
- Điều kiện: cùng kernel trên khắp mọi máy, 1 tiến trình có thể lựa chọn nơi chạy tốt nhất

Hệ thống đa xử lý



- phần cứng chặt – phần mềm chặt
- các CPU chia sẻ chung bộ nhớ thông qua 1 bus tốc độ cao
- vấn đề phân bổ tiến trình và caching thích hợp

So sánh các mô hình

Tiêu chí	N.O.S	Dist.O.S	Multiproces sor O.S
Tạo hình ảnh của 1 bộ xử lý ảo?	không	có	có
Phải chạy cùng một hệ ĐH?	không	có	có
Cách liên lạc	qua tệp chia sẻ	thông báo	qua bộ nhớ chia sẻ
Cần giao thức chung?	có	có	không
Cần có hàng đợi tiến trình?	không	không	có
Ngữ nghĩa tài nguyên rõ ràng?	thường không	có	có

Các tiêu chí cho 1 hệ điều hành phân tán

- Tính trong suốt (transparency)
- Tính mềm dẻo (flexibility)
- Độ tin cậy (Reliability)
- Hiệu suất (performance)
- Tính mở rộng (scalability)

Tính trong suốt

- Vị trí:
 - người dùng không biết vị trí tài nguyên
- Dịch chuyển:
 - có thể chuyển dịch tài nguyên mà không phải thay đổi tên
- Sao lặ:
 - người dùng không biết số lượng bản copy
- Đồng thời:
 - nhiều người dùng có thể đồng thời chia sẻ
- Song song:
 - có thể hoạt động song song không cần tác động của người dùng

...tính trong suốt

- Vị trí:
 - tên của tài nguyên không thể gắn với vị trí của tài nguyên, ví dụ: `\\math\public\a.txt`
 - tính trong suốt về vị trí có thể thấy được trong hệ thống NFS
- Dịch chuyển:
 - ví dụ: người dùng nhóm kế toán xài file-server *ketoan*; nhóm văn phòng xài file server *vanphong*. Đối với SAMBA, 2 file servers trên có thể được tích hợp vào trên 1 máy nếu tạm thời có 1 máy hỏng.

...tính trong suốt

- Sao lặp:
 - ví dụ: hệ thống file servers với kết nối hình dạng vòng, tạo thành 1 hệ thống fault-tolerant trong suốt với client
- Đồng thời:
 - hệ thống phải thực hiện khoá, mở khoá các tài nguyên chung và làm cho người dùng cảm giác sở hữu tài nguyên riêng (serialisation)

...tính trong suốt

- Song song:
 - hệ thống phải tự phân rã công việc để thực hiện song song
 - đây là điểm mà phần mềm còn lâu mới đáp ứng được
 - ví dụ: chương trình chơi cờ sẽ tự phân các bàn cờ con cho các hệ thống phân tán thực hiện song song; người lập trình không hề phải bận tâm chuyện đó
- Trong suốt: với người dùng khác xa với lập trình viên

Tính mềm dẻo

- Trường phái monolithic kernel:
 - kernel cung cấp tất cả các dịch vụ cho chương trình ứng dụng người dùng (Unix, NT...)
- Trường phái micro kernel:
 - kernel chỉ cung cấp các dịch vụ cơ sở, càng ít càng tốt
 - các dịch vụ khác được cung cấp ở mức người dùng, liên lạc theo kiểu client-server
 - có thể có nhiều OS khác nhau

...tính mềm dẻo

- Micro kernel là mềm dẻo. Chỉ cung cấp:
 - cơ chế liên lạc giữa các tiến trình
 - một số dịch vụ quản lý bộ nhớ cục bộ
 - quản lý tiến trình và lập lịch, hạn chế ở mức cục bộ
 - dịch vụ vào ra mức cơ sở
- Monolithic: hiệu suất cao hơn (tạm thời)
- Micro kernel: mở rộng, nâng cấp dễ dàng

Độ tin cậy

- Tính sẵn sàng
 - Lamport: Hệ thống phân tán là hệ thống mà đôi khi tôi không thể làm việc chỉ vì 1 máy tôi chưa bao giờ nghe tên
- Khả năng hứng chịu lỗi
- Vấn đề: tính bảo mật

Hiệu suất

- Yếu tố đường truyền:
 - tốc độ đường truyền không phải là vấn đề chính (vd: Gb ethernet)
 - cản trở tốc độ: việc đóng/bóc các giao thức
- Yếu tố độ lớn công việc
 - nhiều công việc nhỏ, đường truyền chậm: rõ ràng nên làm cục bộ
 - một vài công việc lớn: có thể thực hiện ở xa

Tính mở rộng

- Không thể duy trì 1 hệ thống tập trung duy nhất với thuật toán xử lý tập trung
- Các đặc tính của hệ thống với thuật toán phân tán:
 - không máy nào có toàn bộ thông tin hệ thống
 - quyết định tại mỗi máy dựa trên thông tin cục bộ
 - sự cố của 1 máy không ảnh hưởng tới thuật toán
 - không cần tới giả định của 1 đồng hồ chung cho toàn hệ thống
- Việc bổ sung mở rộng trong hệ thống phân tán là dễ dàng

Một vài so sánh từ góc nhìn người dùng

- Unix
- NT
- Netware

Khía cạnh đa người dùng

	Làm việc từ máy trạm	Làm việc cục bộ
NT	có, nếu chạy terminal server	không
Netware	không	không
Unix	có	có

Khía cạnh quyền thâm nhập

	Phương thức quản lý	Các loại quyền
Netware	Chức năng quyền	Read, write, create, delete, append, modify, soạn
NT	2 mức: ACL và qua kết nối mạng	Read, write, delete, tạo folder/append, duyệt folder/chạy, đọc/ghi thuộc tính tệp Cho phép tạo danh sách deny các quyền
Unix	RWX như thuộc tính tệp	3 vùng xác lập quyền thâm nhập: chủ, nhóm, khác

Khía cạnh chia sẻ tệp

	cách thức	chia sẻ tài nguyên
NT	map ổ đĩa, có thể kết hợp giải pháp DFS	không dễ dàng chia sẻ tài nguyên riêng
Netware	map ổ đĩa	có thể chia sẻ vùng đã phân cấp
Unix	NFS	toàn quyền chia sẻ tài nguyên riêng

Ngữ nghĩa thâm nhập

	Đồng nhất tại trạm và máy chủ
Netware	Không làm việc tại máy chủ. Định dạng đặc biệt NLM cho ứng dụng trên máy chủ.
NT	Môi trường phát triển ứng dụng thống nhất. Môi trường làm việc không đồng nhất
Unix	Đồng nhất khắp mọi nơi

Tích hợp chia sẻ tệp

Với	Netware	NT	Unix
Netware	mount vào cây NDS thành các volume	?	mount theo NFS vào cây NDS thành volume
NT	NW gateway, map thành ổ	map thành ổ qua CIFS	map thành ổ qua CIFS/SMB hoặc mount qua PCNFS
Unix	mount qua IPX/SPX thành thư mục	mount qua SMB thành thư mục	mount qua NFS thành thư mục