

Kiến trúc hệ CSDL đa phương tiện

Nguyễn Đình Hóa

dinhhoa@gmail.com 0942807711

Tóm tắt nội dung bài 4

- ▶ Nén dữ liệu đa phương tiện
 - ▶ Nén dữ liệu văn bản
 - ▶ Mã hóa Huffman
 - ▶ Mã hóa run-length
 - ▶ Mã hóa LZW
 - ▶ Nén dữ liệu âm thanh
 - ▶ Lượng tử hóa phi tuyến
 - ▶ Mã hóa dự đoán
 - ▶ MPEG-audio
 - ▶ Nén dữ liệu hình ảnh
 - ▶ Nén dữ liệu video

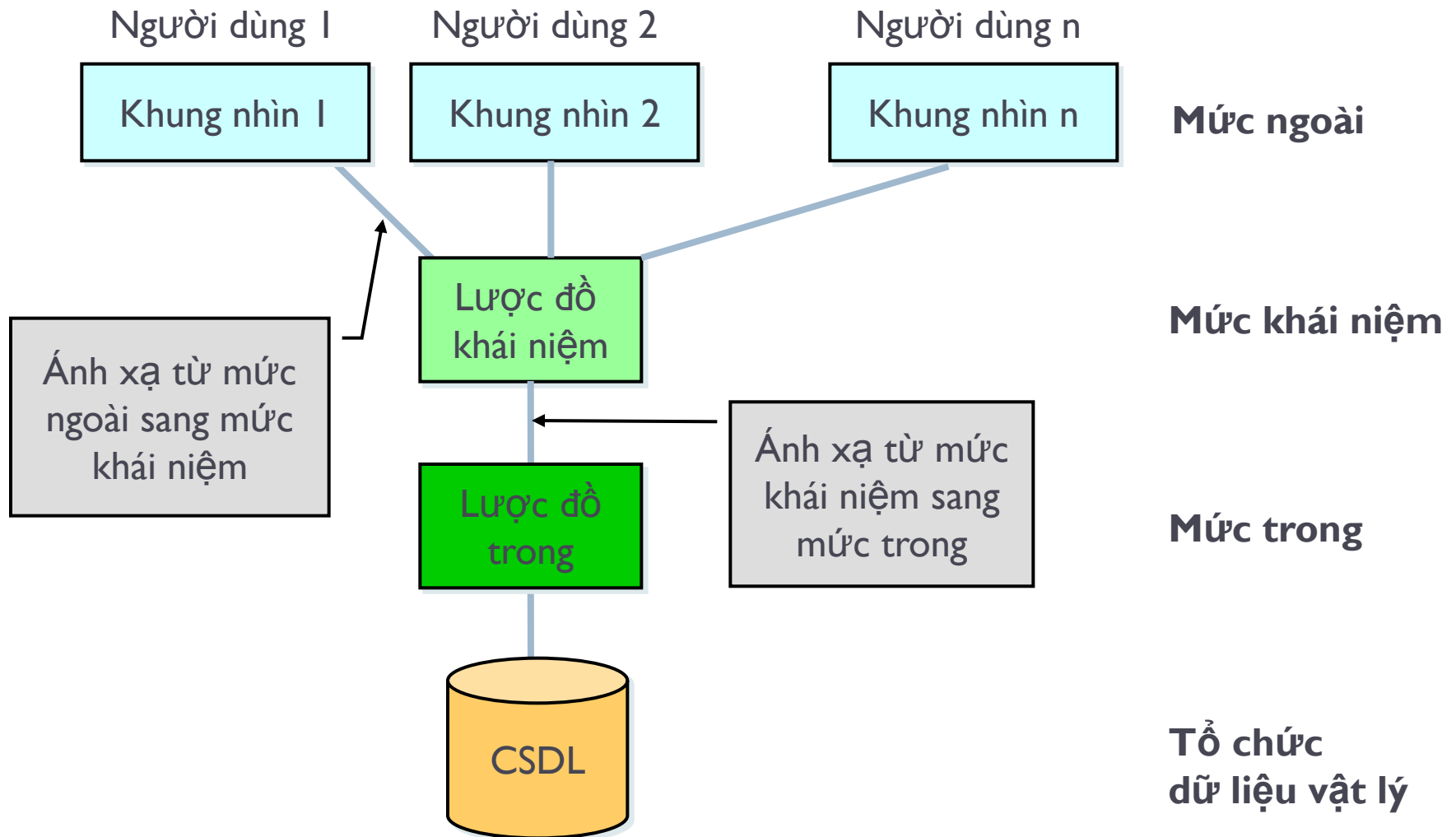
Các yêu cầu đối với kiến trúc CSDL ĐPT

- ▶ Kiến trúc hệ CSDL ĐPT phải đảm bảo hỗ trợ tìm kiếm dữ liệu cho các giao dịch cụ thể.
- ▶ 4 tính chất của một giao dịch CSDL (thuộc tính ACID)
 - **Tính nguyên tử (Atomicity):** mỗi giao dịch là một thể thống nhất, không chia cắt được. Giao dịch chỉ có thể được thực hiện hoặc không.
 - **Tính nhất quán (Consistency):** Mỗi giao dịch phải chuyển đổi CSDL một cách nhất quán từ trạng thái này sang trạng thái khác.
 - **Tính biệt lập (Isolation):** các giao dịch khác nhau phải được thực hiện độc lập với nhau.
 - **Tính duy trì (Durability):** những thay đổi tới CSDL bởi một giao dịch thành công sẽ được lưu trữ vĩnh viễn trong CSDL.

Kiến trúc hệ CSDL đa phương tiện

- ▶ Hệ CSDL đa phương tiện có thể có kiến trúc phức tạp
- ▶ Tùy thuộc vào các yêu cầu cụ thể mà kiến trúc tối ưu cho một hệ CSDL đa phương tiện có thể khác nhau.

Mô hình trừu tượng 3 lớp



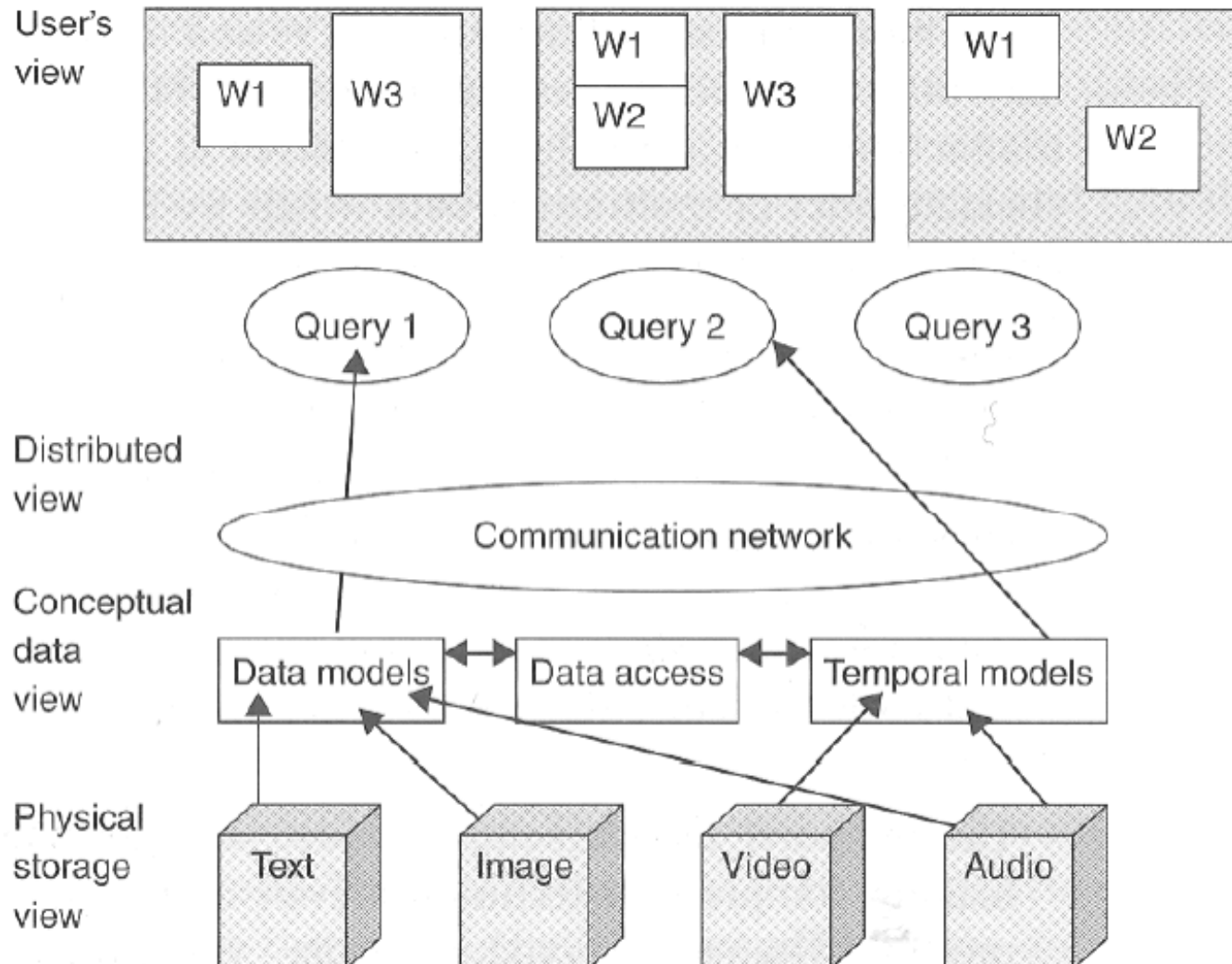
Mô hình trừu tượng 3 lớp

- ▶ **Mức ngoài** là khung nhìn của người sử dụng CSDL, mô tả phần CSDL tương ứng với người dùng đó.
 - ▶ *Mức ngoài bao gồm một số khung nhìn khác nhau về CSDL.*
- ▶ **Mức khái niệm** là khung nhìn của người thiết kế CSDL, mô tả dữ liệu nào được lưu trong CSDL và mối quan hệ giữa chúng.
- ▶ **Mức trong** thể hiện biểu diễn về mặt vật lý của CSDL, mô tả cách thức lưu trữ dữ liệu trong CSDL.

Kiến trúc hệ CSDL đa phương tiện

- ▶ Chất lượng của hệ CSDL ĐPT phụ thuộc vào tốc độ truyền tin từ kho dữ liệu đến bộ xử lý dữ liệu.
- ▶ Kiến trúc hệ CSDL phụ thuộc vào chất lượng máy tính cũng như hệ thống mạng máy tính
 - ▶ Hệ CSDL tập trung
 - ▶ Hệ CSDL cá nhân
 - ▶ Hệ CSDL trung tâm
 - ▶ Hệ CSDL khách - chủ (client - server)
 - ▶ Hệ CSDL phân tán
 - ▶ Thuần nhất
 - ▶ Không thuần nhất

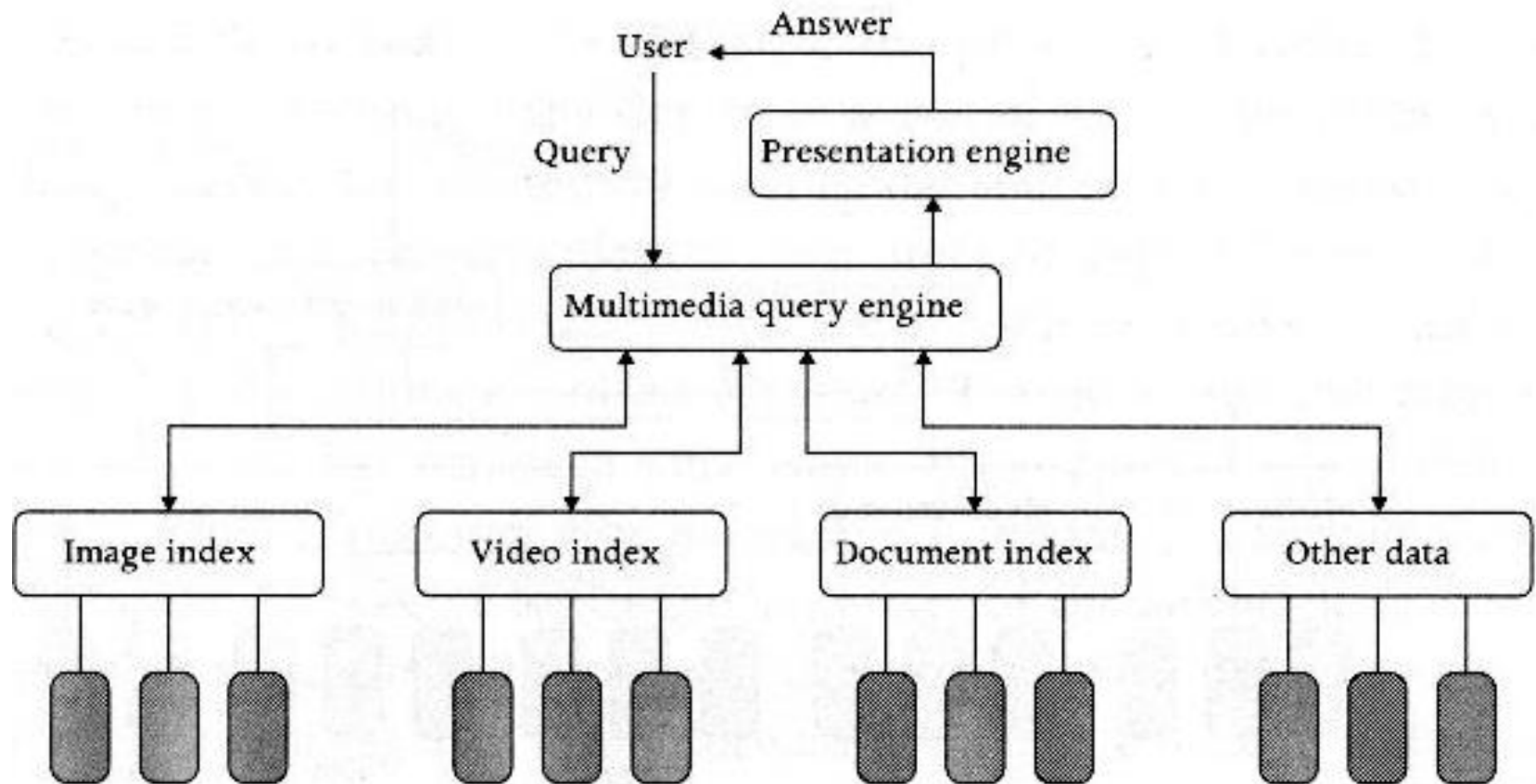
Kiến trúc hệ CSDL đa phương tiện



Kiến trúc hệ CSDL đa phương tiện

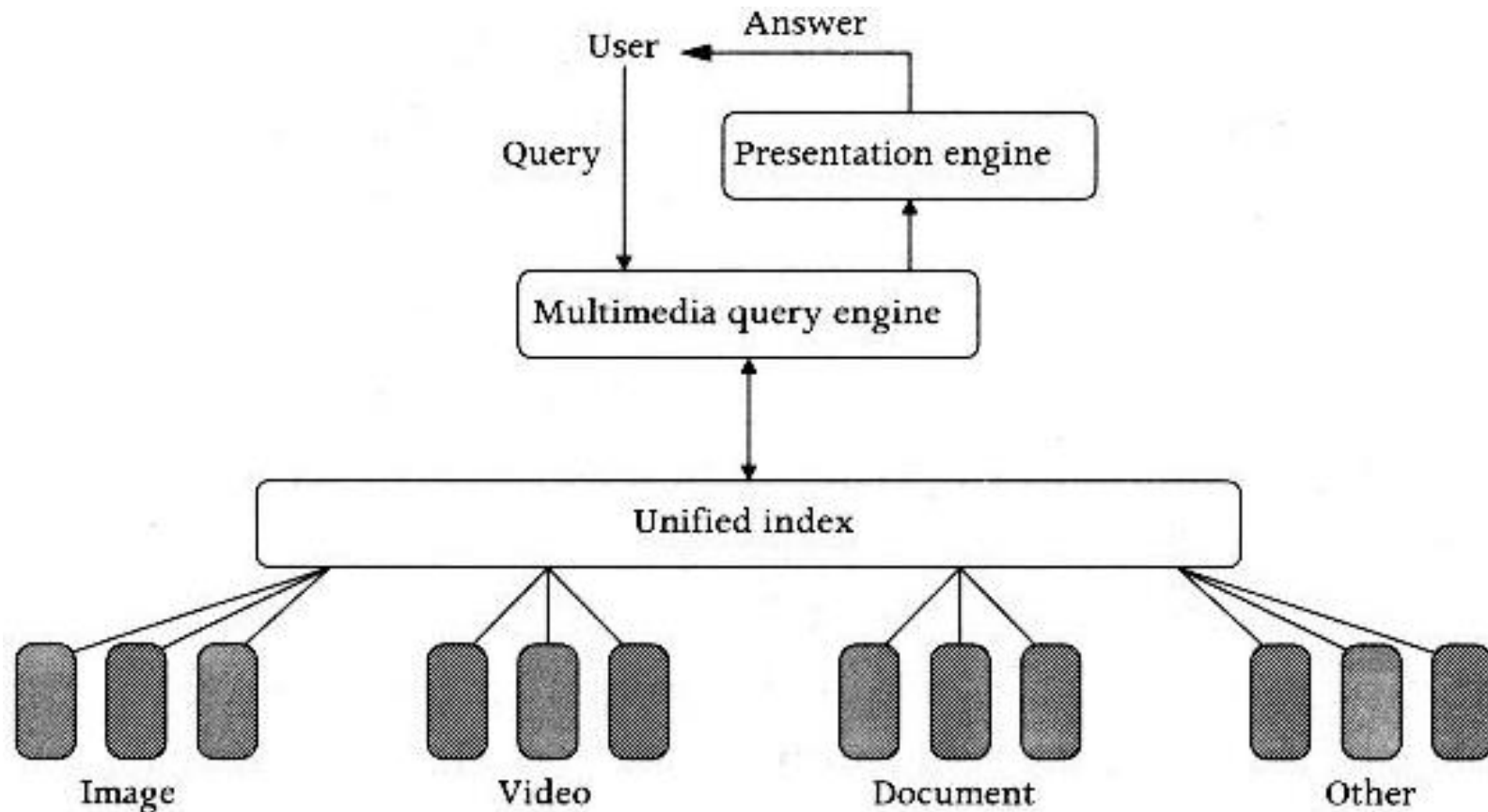
- ▶ Nguyên lý riêng biệt: mỗi loại dữ liệu đơn phương tiện được lưu trữ tại các máy chủ riêng biệt theo các cấu trúc phù hợp với đặc tính của từng loại
- ▶ Nguyên lý đồng nhất: tất cả các loại dữ liệu đa phương tiện được lưu trữ trong cùng một cấu trúc dữ liệu
- ▶ Nguyên lý kết hợp: kế hợp của cả hai loại trên

Kiến trúc hệ CSDL đa phương tiện



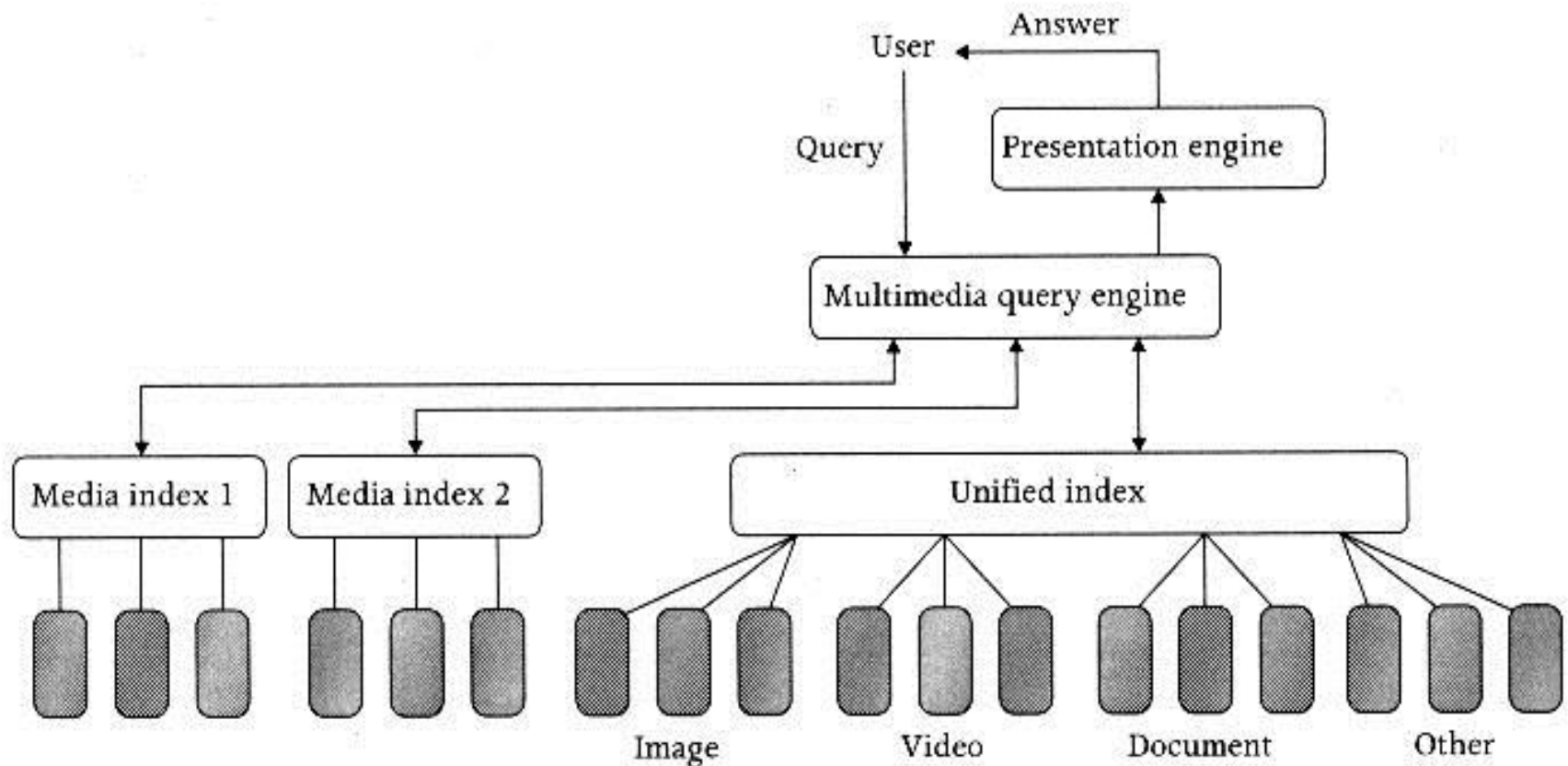
Architecture of a multimedia database under the principle of autonomy

Kiến trúc hệ CSDL đa phương tiện



Architecture of a multimedia database under the principle of uniformity

Kiến trúc hệ CSDL đa phương tiện

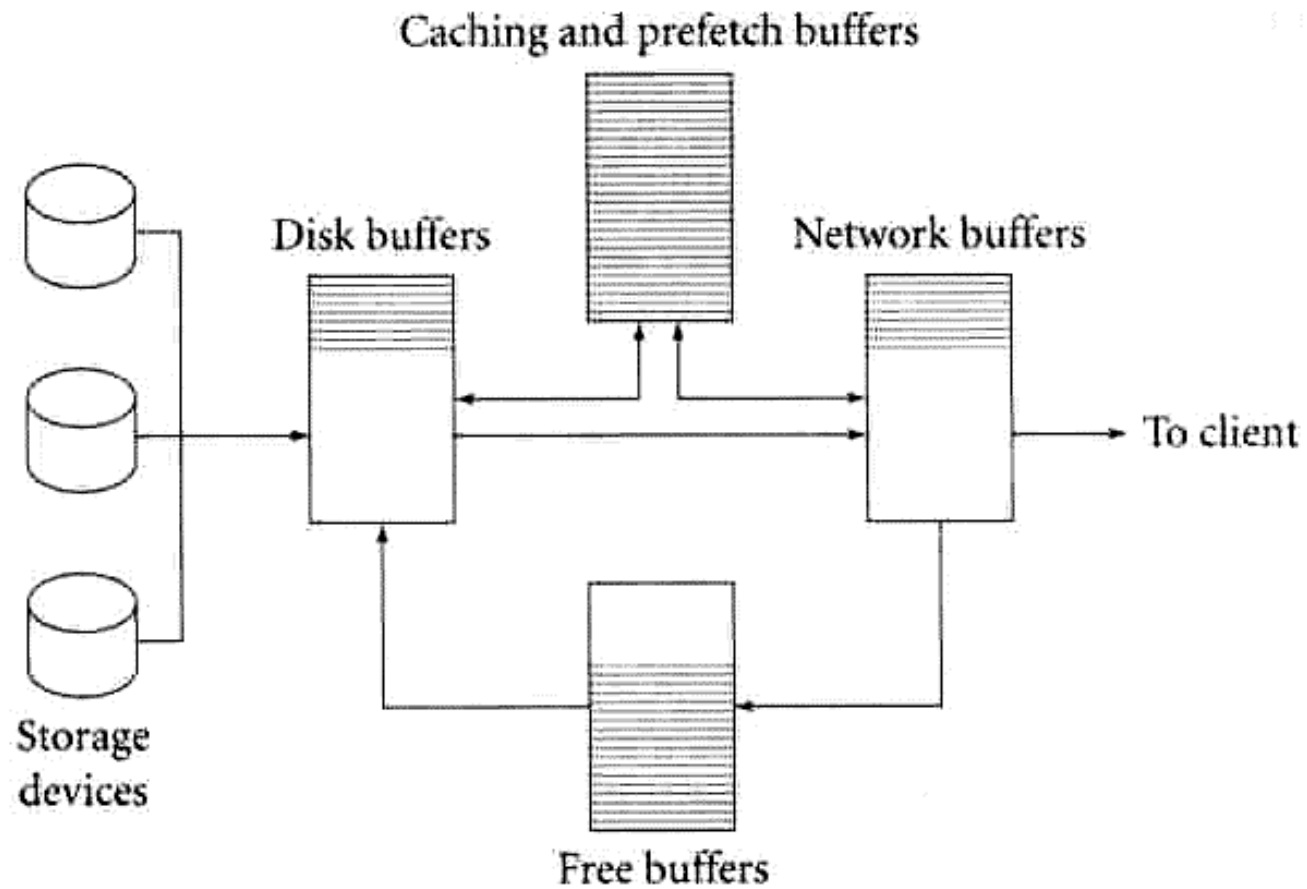


Architecture of a multimedia database under the principle of hybrid organization

Các yêu cầu đối với máy chủ ĐPT

- ▶ Đáp ứng yêu cầu của nhiều loại ứng dụng
- ▶ Thỏa mãn
 - ▶ Nhu cầu của nhiều người sử dụng khác nhau
 - ▶ Băng thông truy cập tốt
 - ▶ Dung lượng lưu trữ lớn (ứng với nhiều loại định dạng dữ liệu đa phương tiện)

Máy chủ đa phương tiện



Máy chủ đa phương tiện

- ▶ Tối thiểu hóa thời gian chờ đợi của yêu cầu dịch vụ
- ▶ Luôn sẵn sàng
- ▶ Khả năng cân bằng tải
- ▶ Có khả năng cung cấp ít nhất một lượng đường truyền dịch vụ nhất định
- ▶ Đáp ứng các dịch vụ thời gian thực
- ▶ Có khả năng tổng hợp nhu cầu người dùng
- ▶ Có khả năng tương tác với người dùng: tua nhanh, quay lại, tạm dừng,...

Hệ CSDL đa phương tiện phân tán

- ▶ Các bảng CSDL có thể được chia nhỏ với các mối quan hệ phụ
 - ▶ Chia nhỏ theo chiều ngang
 - ▶ Chia nhỏ theo chiều dọc
- ▶ CSDL được phân mảnh
- ▶ Các phân đoạn dữ liệu được nhân thành nhiều bản sao, có thể được lưu trữ ở các điểm khác nhau
- ▶ LOB: Các dữ liệu (video, audio) được lưu trữ tại các điểm gần nơi yêu cầu dịch vụ.

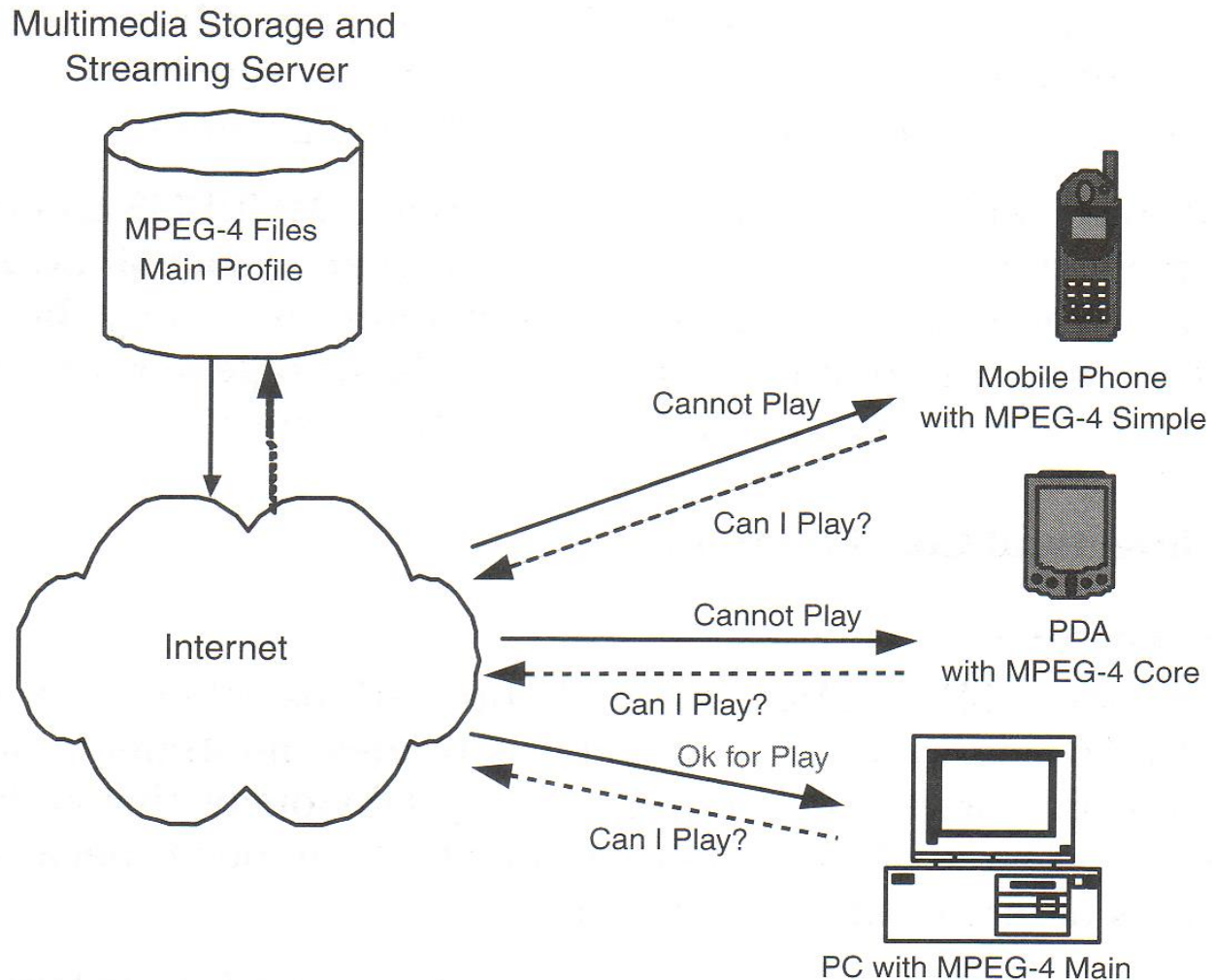
Siêu máy chủ

- ▶ Lấy dữ liệu từ nhiều máy chủ khác nhau
 - ▶ Gặp vấn đề khi một (hoặc một vài) máy chủ được lựa chọn theo mặc định của hệ thống, hoặc theo sự lựa chọn của khách hàng
 - ▶ Lựa chọn máy chủ theo chế độ tĩnh gây mất cân bằng tải
- ▶ Lựa chọn máy chủ theo chế độ động: luân phiên lấy dữ liệu từ các máy chủ ở các vị trí khác nhau.
- ▶ Các yêu cầu về dữ liệu ĐPT sẽ được chuyển đến máy chủ phù hợp (tùy thuộc vào vị trí điểm yêu cầu dịch vụ, trạng thái tải của các máy chủ, vị trí địa lý của các máy chủ, băng thông của mạng)

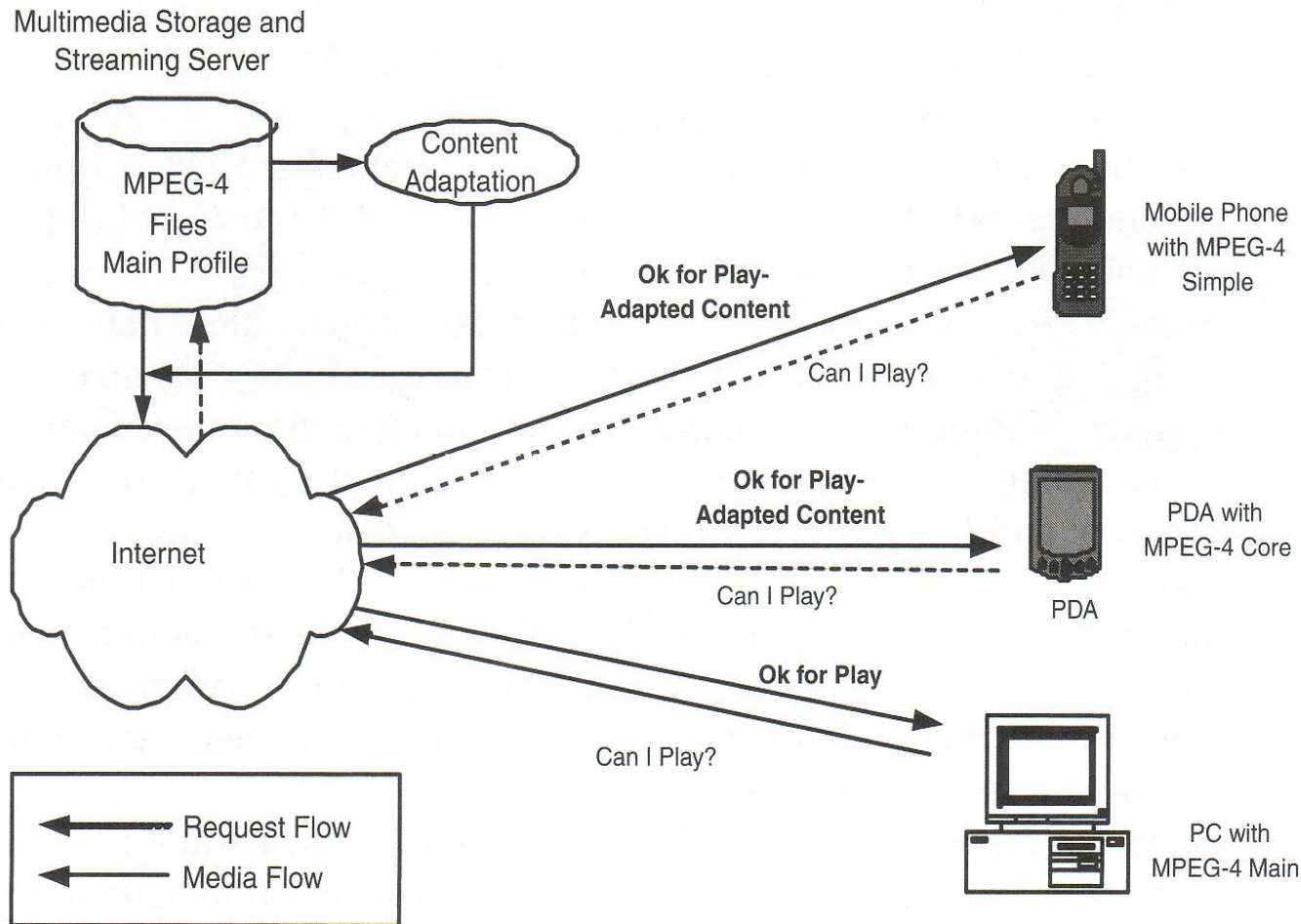
Hệ thống khách-chủ (client-server)

- ▶ Khách hàng gửi yêu cầu dịch vụ
- ▶ Yêu cầu được xử lý nội bộ tại một máy chủ
- ▶ Nhược điểm: có khả năng mất cân bằng hệ thống
 - ▶ Do số lượng khách hàng
 - ▶ Sự không đồng nhất về cấu hình máy khách
 - ▶ Kích thước dữ liệu được yêu cầu

Hệ thống khách-chủ (client-server)



Hệ thống khách-chủ (client-server)

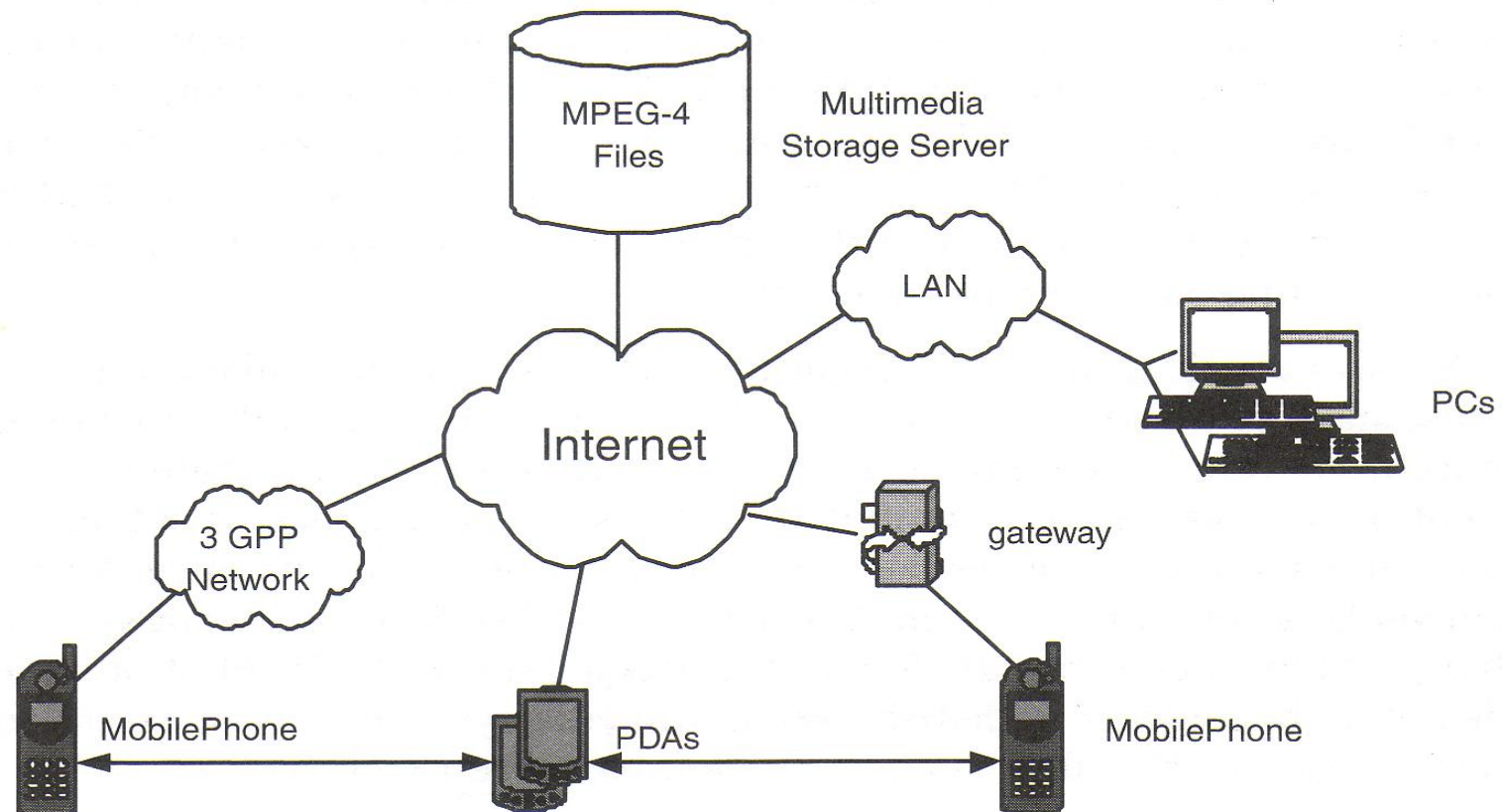


Hệ thống khách-chủ (client-server)

- ▶ Content adaptation (bộ thích ứng nội dung)
 - ▶ Đảm bảo chất lượng dịch vụ phù hợp với các máy đầu cuối khác nhau
 - ▶ Đảm bảo chất lượng nội dung phù hợp với yêu cầu của mạng dịch vụ
 - ▶ Đảm bảo chất lượng thông tin (độ phân giải, tốc độ gói tin,...)

Hệ thống điểm-điểm (peer-to-peer)

- ▶ Các thiết bị cầm tay trao đổi dữ liệu trực tiếp với nhau



Hệ thống điểm-điểm (peer-to-peer)

▶ Các tính năng chính

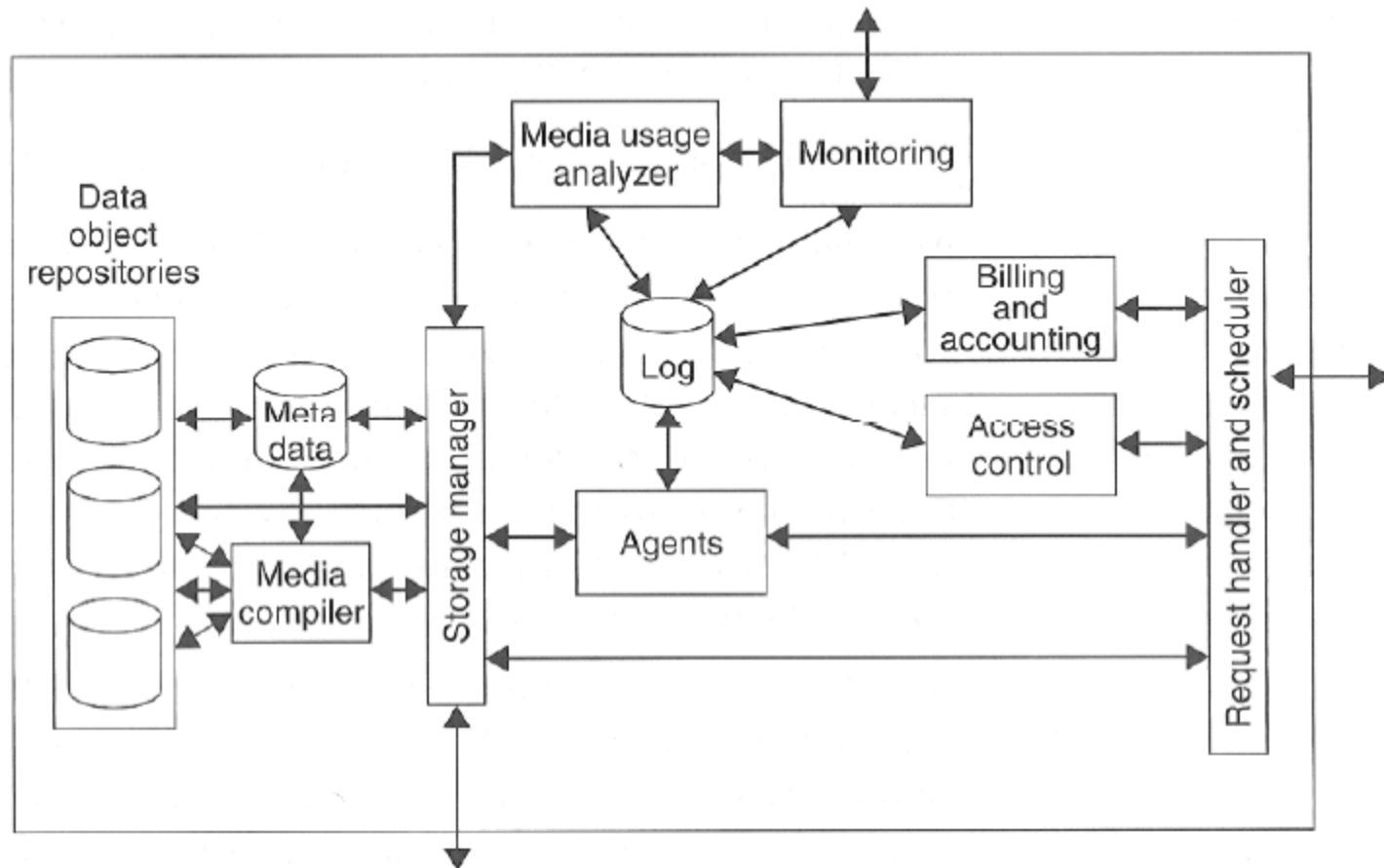
- ▶ Lưu trữ thông tin trước khi truyền
- ▶ Kiểm tra tính tương thích của máy bên kia trước khi truyền dữ liệu
- ▶ Tạo siêu dữ liệu mô tả dữ liệu nhận được

▶ Các vấn đề về kỹ thuật

- ▶ Cần có chuẩn chung về kết nối thông tin giữa các máy (P2P communication protocols, VD: CC-PP, JXTA v1.0)
- ▶ Giao thức liên kết giữa các máy (mobile agent)

Các hệ thống phân tán hỗn hợp

- ▶ Tồn tại nhiều hệ thống quản trị CSDL khác nhau, nhiều mô hình dữ liệu khác nhau



Quản lý nội dung

- ▶ Kết hợp nhiều công nghệ quản lý khác nhau
- ▶ Quản lý theo ngữ nghĩa
 - ▶ Thông tin ngoài lề (ngày tháng, địa điểm,...)
 - ▶ Thông tin bên trong (nội dung câu từ, hình ảnh chủ đạo, phân bố màu sắc,...)
 - ▶ ...
- ▶ Phân chia dữ liệu thành các phần nhỏ theo các siêu dữ liệu, xây dựng cấu trúc chỉ số hóa.
 - ▶ Thực hiện theo cơ chế tự động

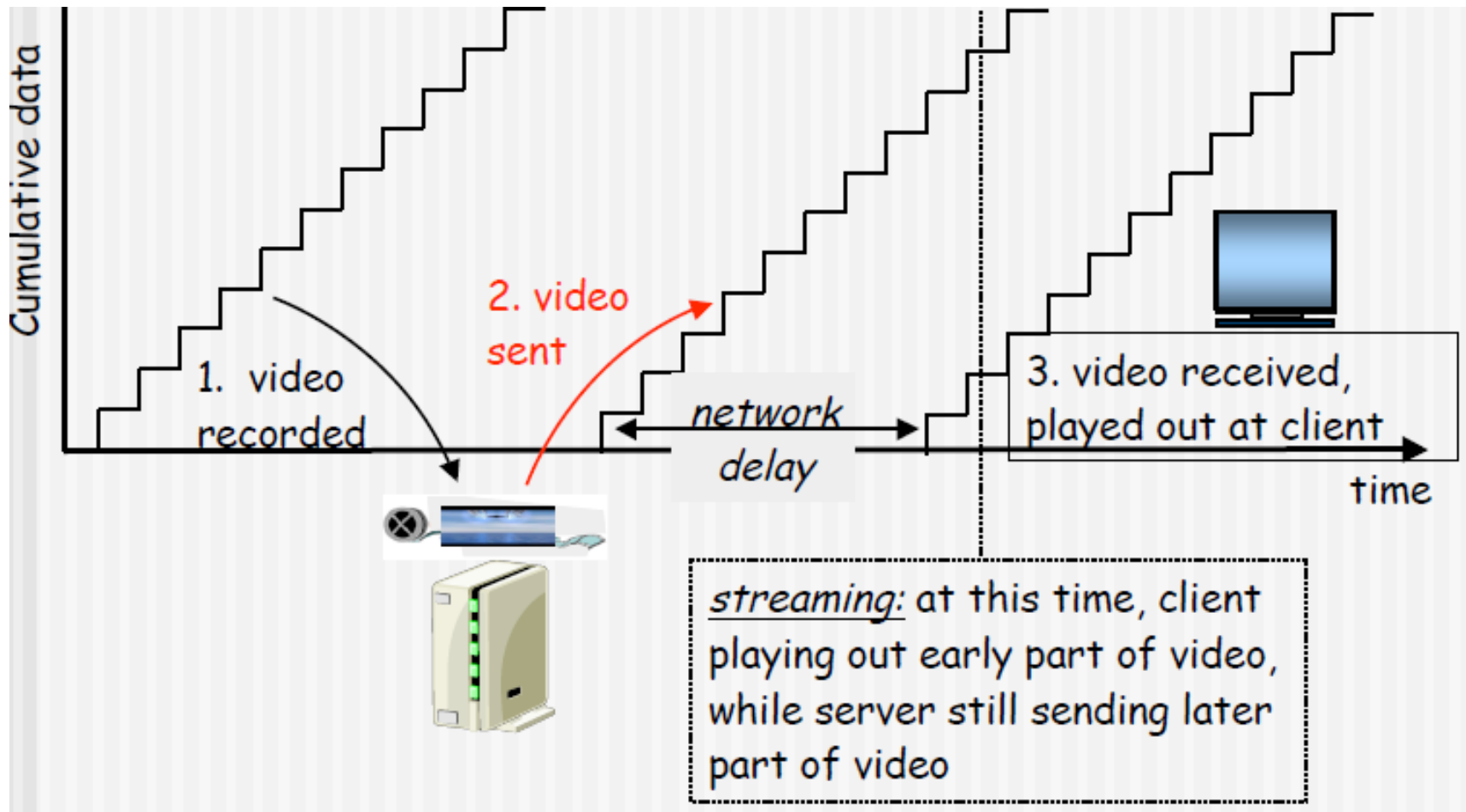
Các luồng dữ liệu đa phương tiện

- ▶ Các hệ thống đa phương tiện được thiết kế để truyền dữ liệu với tốc độ ổn định
- ▶ Streaming: kỹ thuật truyền dữ liệu theo các luồng liên tục và ổn định
- ▶ Streaming cho phép người dùng có thể xem nội dung trong quá trình truyền file dữ liệu
 - ▶ Nếu tốc độ truyền quá nhanh so với tốc độ xem nội dung thì dữ liệu sẽ được lưu tại bộ đệm
 - ▶ Nếu tốc độ truyền quá chậm so với tốc độ xem nội dung thì xảy ra hiện tượng giật, không liên tục

Ứng dụng Đa phương tiện trên mạng

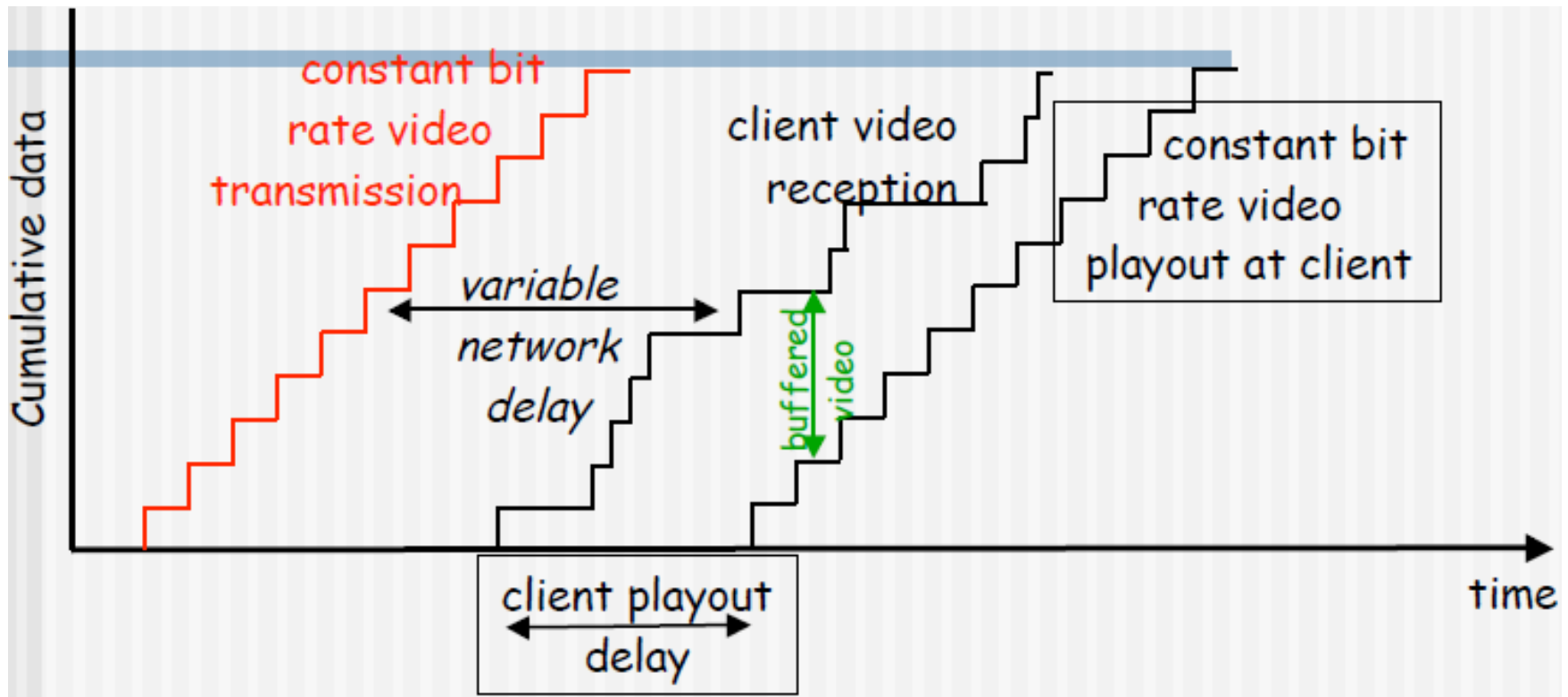
- ▶ Truyền dữ liệu đã được lưu tại máy chủ
- ▶ Truyền dữ liệu (video, audio) trực tiếp
- ▶ Tương tác video, audio thời gian thực
- ▶ Các đặc tính chính
 - ▶ Dễ bị trễ
 - ▶ Trễ giữa điểm đầu và điểm cuối
 - ▶ Trễ các gói tin trên cùng đường truyền
 - ▶ Khả năng bù thông tin: những thông tin bị mất (ngẫu nhiên) trong quá trình truyền tin không ảnh hưởng nhiều đến chất lượng tin.

Truyền dữ liệu đã lưu trong máy chủ



Truyền dữ liệu cần bộ đệm

- ▶ Bộ đệm: cho phép bù trễ đường truyền và trễ thông tin từ điểm phát đến điểm thu.



Một số yêu cầu đối với dịch vụ ĐPT

Application	Reliability	Delay	Jitter	Bandwidth
E-mail	High	Low	Low	Low
File transfer	High	Low	Low	Medium
Web access	High	Medium	Low	Medium
Remote login	High	Medium	Medium	Low
Audio on demand	Low	Low	High	Medium
Video on demand	Low	Low	High	High
Telephony	Low	High	High	Low
Videoconferencing	Low	High	High	High

Yêu cầu về băng thông

<u>Application</u>	<u>Bandwidth</u>
Telephony	64 kbps
Telephony	16 - 32 kbps
Teleconferencing	48 - 64 kbps
2-channel audio	128 - 384 kbps
5-channel audio	320 kbps
Videoconferencing	64 - 153 kbps
High-Definition TV	17 Mbps

Các đặc tính của lưu lượng truyền ĐPT

- ▶ Thường bao gồm các dòng truyền liên tục của dữ liệu video và audio số cùng với các ảnh số
- ▶ Các dòng truyền video và audio thường tách biệt nhau
- ▶ Dòng truyền bao gồm nhiều gói dữ liệu
- ▶ Các gói dữ liệu được truyền liên tục với nguyên lý đồng nhất.