Chương 2: Kiến trúc trong hệ phân tán

Họ tên SV: Đặng Mạnh Trường. MSSV: 20134209

**Câu 1**: Chức năng các tầng trong mô hình OSI

* Physical Layer (tầng vật lý): Truyền tải bit trên đường truyền vật lý
* Data-link layer (tầng liên kết dữ liệu): Truyền tải các khung dữ liệu (frame) giữa 2 máy tính
* Network layer (tầng mạng): Tìm đường đi cho dữ liệu đến các đích khác nhau trong mạng
* Transport layer (tầng giao vận): Đảm bảo truyền tải dữ liệu
* Session Layer (tầng phiên): Cho phép các ứng dụng thiết lập, sử dụng và xóa các kênh giao tiếp giữa chúng
* Presentation Layer (tầng biểu diễn): Đảm bảo các máy tính có kiểu định dạng dữ liệu khác nhau vẫn có thể trao đổi thông tin cho nhau
* Application Layer (tầng ứng dung): Cung cấp các ứng dụng truy xuất đến các dịch vụ mạng

Ví dụ: Khi chúng ta đổi địa chị IP máy từ IPv4 sang IPv6 thì việc truy cập vẫn không bị ảnh hưởng gì

**Câu hỏi 2**: Cho ví dụ và phân tích một mô hình kiến trúc thuê bao/xuất bản

(publish/subscribe)

Ví dụ: Trang thông tin tuyển dụng lập trình viên (LTV)

1 số trang tuyển dụng LTV (itviec,….) phân loại thông tin tuyển dụng theo ngôn ngữ lập trình (ví dụ: Java, C,Python,….). Một LTV đang tìm việc sẽ đăng ký tài khoản và chọn những ngôn ngữ lập trình mà mình biết. Thông tin tuyển dụng, sau khi được nhà tuyển dụng gửi tới admin, sẽ được admin đăng lên kèm theo ngôn ngữ lập trình được yêu cầu. Tất cả người dùng đã chọn ngôn ngữ đó sẽ được thông báo về tin tuyển dụng đó.

Ở đây

- Publisher là admin của trang tuyển dụng LTV

- Topic là ngôn ngữ lập trình

- Subscriber là những người dùng trên trang

**Câu hỏi 3**: Sự khác nhau giữa phân tán dọc và phân tán ngang là gì?

- Phân tán dọc (vertical distribution): refers to the distribution of the different layers in a multitiered architecture across multiple machines (chia thành nhiều tầng, phân chia theo chức năng). Nhược điểm: chi phí truyền thông lớn

- Phân tán ngang (horizontal distribution): deals with the distribution of a single layer across multiple machines, such as distributing a single database. (chỉ có 1 tầng, phân bố theo tải). Nhược điểm: Mất nhiều thời gian để truy cập do không có phân tầng riêng biệt

Tham khảo: Distributed Systems: Principles and Paradigms (Tanenbaum, van Steen)

**Câu hỏi 4**: Phân tích ưu nhược điểm của kiến trúc tập trung và kiến trúc không

tập trung.

Kiến trúc tập trung:

* Ưu điểm: Tăng năng suất thông qua việc sử dụng các user-interface, có hiệu quả về chi phí, tăng cường lưu trữ dữ liệu, kết nối rộng rãi và các dịch vụ ứng dụng đáng tin cậy. Cải thiện việc chia sẻ dữ liệu, tích hợp được các dịch vụ, chia sẻ tài nguyên giữa các nền tảng khác nhau, đảm bảo tính toàn vẹn của cơ sở dữ liệu, khả năng xử lý dữ liệu ở mọi nơi, dễ dàng bảo trì bảo mật.
* Nhược điểm: Có thể gây ra tình trạng máy chủ quá tải khi nhiều client truy cập; khi 1 máy chủ server quan trọng bị lỗi, các yêu cầu khác từ phía client sẽ không được đáp ứng

Kiến trúc không tập trung:

* Ưu điểm : Dễ cài đặt và cấu hình; tất cả tài nguyên và nội dung được chia sẻ giữa tất cả các peer, không phụ thuộc peer khác. Người dung có thể kiểm soát việc chia sẻ tài nguyên của họ. Chi phí bảo mật thấp so với kiến trúc tập trung
* Nhược điểm: Khó quản lý, phục hồi sao lưu dữ liệu khó khăn

**Câu hỏi 5**: Trong một mạng overlay có cấu trúc, các thông điệp được định

tuyến dựa theo hình trạng mạng (topology). Nhược điểm quan trọng của hướng

tiếp cận này là gì?

- Độ trễ lớn, cũng vậy, vì truyền qua 1 nút mất 1 khoảng thời gian để nó xác định định tuyến (tăng độ phức tạp)

- Do liên kết giữa các node là liên kết logic nên không thấy được liên kết vật lý bên dưới

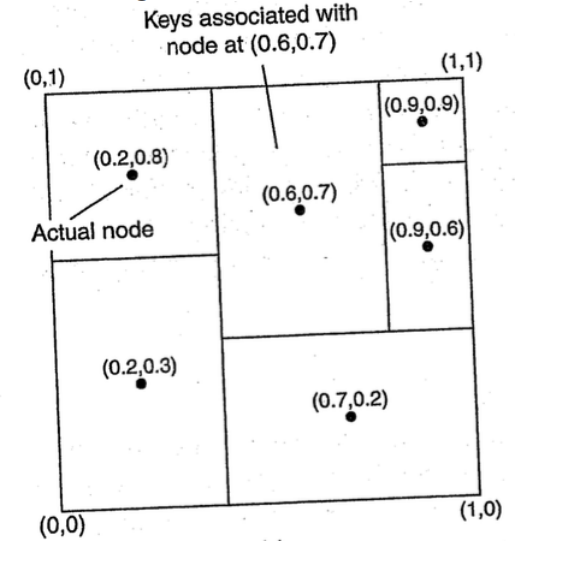
Câu hỏi 6: Vấn đề nảy sinh:

Nếu trong tiến trình P\_i+1 nào đó xuât hiện lỗi không trả lời được tiến trình Pi thì sẽ dẫn tới treo toàn bộ hệ thống.

Ngoài ra với kiến trúc đa tầng, thời gian đáp ứng request (phản hồi) sẽ lâu hơn do phải mất thời gian chuyển qua lại nhiều tầng.

1 vấn đề khác nảy sinh là hiệu năng bị ảnh hưởng từ các tầng khác nhau

**Câu hỏi 7:**



Giải thuật này nhìn chung là tốt, vì: Khi so sánh với các thuật toán định tuyến như Distance-Vector (DV) và Link-State (LS), ta thấy là DV và LS sử dụng thông tin về hình trạng mạng, nên với những mạng có hình trạng mạng thay đổi liên tục, với DV và LS sẽ cần cập nhật thông tin liên tục, trong khi giải thuật được nêu ra không phụ thuộc vào hình trạng mạng mà chỉ phụ thuộc vào hash function (hàm băm) từ cặp (key,value) tới không gian tọa độ.

Mặt khác, việc phân chia không gian tọa độ thành các vùng hoàn toàn phụ thuộc vào hash function (hàm băm) mà không liên quan gì tới khoảng cách vật lý/ IP hop giữa các bộ định tuyến, nên với giải thuật được nêu, 1 bộ định tuyến được coi là “hàng xóm gần mình nhất” có thể lại khá xa về mặt vật lý và như vậy IP hop sẽ lớn. Trong [1] các tác giả đề xuất 1 số metric định tuyến đặc biệt để đối phó với vấn đề này.

Tham khảo: [1] Ratnasamy et al (2001). “A scalable Content-Addressable Network”. In Proceedings of ACM SIGCOMM 2001