

Đề Thi Kết Thúc Học Phần
CSE702063 - Ứng Dụng Phân Tán (Đề Số 1)
Time: **90 phút**

Đề thi bao gồm:

- **20** câu hỏi trắc nghiệm (5 điểm)
- **Ba** câu hỏi tự luận (5 điểm)

Sinh viên không được sử dụng tài liệu. Giám thị trông thi không giải thích gì thêm.

Chúc các bạn hoàn thành bài thi tốt!

Trắc Nghiệm

Câu 1. Hệ Thống Ngân Hàng Truyền Thống: Trong những hệ thống này, tất cả giao dịch và quản lý dữ liệu của khách hàng được xử lý và lưu trữ trong một máy chủ trung tâm hoặc trung tâm dữ liệu được kiểm soát bởi ngân hàng. Hệ thống trên là:

- A) Hệ thống tập trung
- B) Hệ thống phân tán
- C) Hệ thống phi tập trung
- D) Hệ thống phân tán hoặc phi tập trung

Câu 2. Cấu Trúc Hệ Thống Phi Tập Trung nổi bật ở điểm nào sau đây?

- A) Có nút điều phối trung tâm quan trọng
- B) Tính tự chủ của các nút không được nhấn mạnh
- C) Thường không có nút điều phối trung tâm
- D) Tập trung vào phân bổ trên một nút trung tâm

Câu 3. Để đảm bảo tính linh hoạt và đáng tin cậy trong truyền thông phân tán, công nghệ nào thường được sử dụng nhiều nhất? (Chọn tất cả các đáp án đúng)

- A) Blockchain
- B) Cloud computing
- C) Hệ thống phân tán P2P (Peer-to-Peer)
- D) Mạng cảm biến không dây (Wireless Sensor Network)

Câu 4. Trong hệ thống phân tán, vai trò của máy chủ và máy khách là gì?

- A) Máy chủ lưu trữ dữ liệu, máy khách truy cập dữ liệu.
- B) Máy khách lưu trữ dữ liệu, máy chủ truy cập dữ liệu.
- C) Cả máy chủ và máy khách đều có thể lưu trữ và truy cập dữ liệu.
- D) Máy chủ quản lý mạng, máy khách quản lý dữ liệu.

Câu 5. Giao thức truyền thông trong hệ thống mạng máy tính đóng vai trò như thế nào?

- A) Giao thức truyền thông quản lý việc kết nối vật lý giữa các máy tính.
- B) Giao thức truyền thông là quy tắc được định nghĩa để cho phép truyền thông giữa các máy tính qua mạng.
- C) Giao thức truyền thông là một loại phần mềm để kiểm soát tốc độ truyền dữ liệu.
- D) Giao thức truyền thông là công nghệ chỉ sử dụng trong các mạng cục bộ.

Câu 6. Giao thức khóa phân tán thường thuộc về lớp nào trong hệ thống mạng?

- A) Physical layer
- B) Transport layer
- C) Application layer
- D) Middleware

Câu 7. Khi một quy trình trên máy tính A gọi một thủ tục trên máy tính B, quy trình gọi trên A bị tạm ngưng, và thực thi của thủ tục được gọi xảy ra trên B. Thông tin có thể được truyền từ người gọi đến người được gọi trong các tham số và có thể trở lại trong kết quả của thủ tục. Không có thông điệp truyền qua lại nào được nhìn thấy bởi lập trình viên. Phương pháp này được biết đến là cuộc gọi thủ tục từ xa, hoặc thường chỉ là RPC. Trong RPC, quy trình nào bị tạm ngưng khi thực hiện cuộc gọi thủ tục từ xa?

- A) Quy trình gọi trên máy tính A
- B) Quy trình gọi trên máy tính B
- C) Cả hai quy trình trên máy tính A và B
- D) Không có quy trình nào bị tạm ngưng

Câu 8. Một cơ chế truyền tham số khác cũng tồn tại, tuy nhiên nó không được sử dụng trong hầu hết các ngôn ngữ lập trình. Nó được gọi là "gọi theo sao chép/phục hồi". Nó bao gồm việc sao chép biến vào ngăn xếp bởi người gọi, như trong gọi theo giá trị, và sau đó sao chép trở lại sau khi gọi, ghi đè giá trị ban đầu của người gọi. Dưới hầu hết các điều kiện, điều này đạt được cùng hiệu quả như gọi theo tham chiếu, nhưng trong một số tình huống, chẳng hạn như cùng một tham số xuất hiện nhiều lần trong danh sách tham số, ngữ nghĩa sẽ khác nhau. Cơ chế truyền tham số nào không phổ biến nhưng vẫn tồn tại trong lập trình?

- A) Gọi theo giá trị
- B) Gọi theo tham chiếu
- C) Gọi theo sao chép/phục hồi
- D) Gọi theo con trỏ

Câu 9. Trong hệ thống phân tán, khái niệm "tiến trình" có thể là (chọn các đáp án đúng)

- A) Quá trình chuyển giao dữ liệu giữa các nút trong mạng.
- B) Các bước cụ thể để cài đặt phần mềm trên các thiết bị khác nhau.
- C) Sự tiến triển của các công nghệ mạng mới.
- D) Quá trình thực thi các tác vụ hoặc dịch vụ trên các nút trong hệ thống.

Câu 10. Trong ứng dụng phân tán, thread thường được hiểu là gì?

- A) Một luồng dữ liệu được truyền giữa các nút trong mạng.
- B) Một tiến trình riêng biệt chạy đồng thời với các tiến trình khác trên cùng một máy tính.
- C) Một phần tử trực tuyến trong cơ sở dữ liệu phân tán.
- D) Một phần của một tiến trình chính, giúp thực hiện các tác vụ đồng thời trong một ứng dụng phân tán.

Câu 11. Trong quá trình tạo một quá trình mới, hệ điều hành cần thực hiện những hoạt động nào?

- A) Khởi tạo các đăng ký CPU.
- B) Sao chép dữ liệu từ một quá trình khác.
- C) Xóa dữ liệu từ một phân đoạn nhớ.
- D) Tất cả các đáp án trên

Câu 12. Định danh (Authentication) chủ yếu liên quan đến:

- A) Xác định cá nhân
- B) Xác minh danh tính
- C) Mã hóa dữ liệu
- D) Truy cập thông tin không được phép

Câu 13. Xác thực dựa trên hành động của một người dùng là gì?

- A) Xác định dựa trên kiến thức
- B) Xác thực dựa trên hành động
- C) Xác thực hai yếu tố
- D) Xác thực dựa trên thiết bị

Câu 14. Định nghĩa nào sau đây phù hợp nhất với đồng bộ hóa trong ngữ cảnh của quy trình?

- A) Đảm bảo hai tập dữ liệu giống nhau.
- B) Quản lý tương tác giữa các hoạt động trong hệ thống phân tán.
- C) Đảm bảo rằng một quy trình chờ đợi quá trình khác hoàn tất hoạt động của nó.
- D) Đảm bảo các tương tác và phụ thuộc giữa các hoạt động được quản lý.

Câu 15. Trong một hệ thống phân tán, việc đạt được sự đồng thuận về thời gian là:

- A) Đơn giản và dễ dàng.
- B) Không có ý nghĩa.

C) Khá phức tạp và không dễ dàng.

D) Chỉ cần một cuộc họp giữa các quy trình để thống nhất.

Câu 16. Liệu chúng ta có thể xem một mạch số như một hệ thống phân tán không? (Chọn các đáp án đúng)

A) Có, bởi vì mạch số bao gồm nhiều thành phần hoạt động đồng thời và liên kết với nhau.

B) Không, bởi vì mạch số không sử dụng giao thức truyền thông để trao đổi thông tin giữa các thành phần.

C) Có, bởi vì mạch số có thể phân phối công việc giữa các thành phần khác nhau.

D) Không, bởi vì mạch số không có khả năng chịu lỗi như hệ thống phân tán.

Câu 17. Để hiểu vai trò của tính chịu lỗi trong các hệ thống phân tán, trước tiên chúng ta cần xem xét kỹ hơn về ý nghĩa thực sự của việc một hệ thống phân tán chịu được lỗi như thế nào. Việc chịu lỗi chặt chẽ liên quan đến những gì được gọi là các hệ thống đáng tin cậy. Đáng tin cậy là một thuật ngữ bao gồm một số yêu cầu hữu ích cho các hệ thống phân tán, bao gồm các yêu cầu: (chọn tất cả các đáp án đúng)

A) Tính sẵn sàng (Availability)

B) Tính tin cậy (Reliability)

C) Tính an toàn (Safety)

D) Tính có thể bảo trì (Maintainability)

Câu 18. Độ tin cậy $R(t)$ của một thành phần trong khoảng thời gian $[0, t)$ được định nghĩa một cách chính xác như xác suất có điều kiện rằng nó đã hoạt động đúng trong khoảng thời gian đó, biết rằng nó đã hoạt động đúng vào thời điểm $T = 0$. Theo Pradhan [1996], để thiết lập $R(t)$, chúng ta xem xét một hệ thống gồm N thành phần giống nhau. Đặt $N_0(t)$ là số thành phần hoạt động đúng vào thời điểm t và $N_1(t)$ là số thành phần gặp sự cố. Ta có:

$$R(t) = \frac{N_0(t)}{N} = 1 - \frac{N_1(t)}{N} = \frac{N_0(t)}{N_0(t) + N_1(t)}.$$

Cho các mệnh đề sau:

(I) Độ tin cậy chỉ nhận giá trị rời rạc.

(II) Độ tin cậy luôn không âm và bé hơn 1.

(III) Cận trên của độ tin cậy là 1.

(IV) Từ chỉ số độ tin cậy và số máy hoạt động không chính xác có thể tính được số máy hoạt động chính xác.

Các mệnh đề đúng là:

A) I, II, III, IV

B) II, III, IV

C) II, III

D) II, IV

Câu 19. Tính chịu lỗi trong hệ thống phân tán thường được đo bằng khả năng chịu được bao nhiêu lỗi mà hệ thống vẫn có thể tiếp tục hoạt động đúng cách. Nếu một hệ thống có ba máy chủ và mỗi máy chủ có xác suất gặp sự cố là 0.1 trong một ngày, tính xác suất để hệ thống vẫn hoạt động đúng sau một ngày?

A) 0.97

B) 0.999

C) 0.9999

D) 0.9

Câu 20. 24 giờ sau khi được đồng bộ, chênh lệch tối đa giữa hai đồng hồ là bao nhiêu, với tốc độ lệch tối đa của đồng hồ là 1 trong 10^6 ?

A) 86.4 [ms]

B) 86.4 [s]

C) 864 [ms]

D) 864 [s]

Tự Luận

Câu 1 (2 điểm)

Nêu định nghĩa và giải thích các mô hình thương mại hoá và triển khai của một ứng dụng phân tán:

0.5 điểm Pay-as-you-go

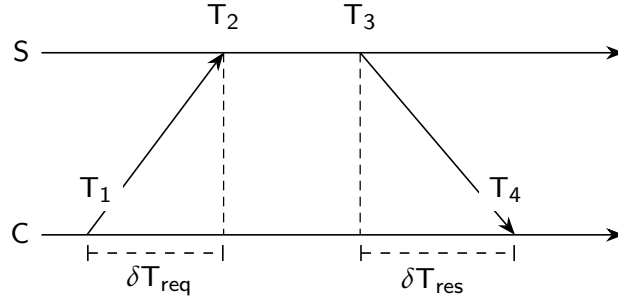
0.5 điểm PaaS

0.5 điểm SaaS

0.5 điểm IaaS

Câu 2 (2 điểm)

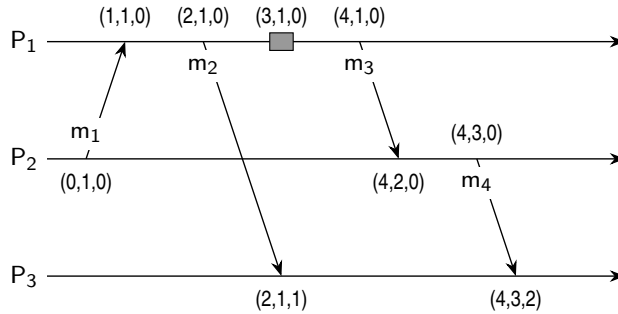
A. (1 điểm)



Hình 1: Sử dụng cho Câu A.

Ở thời gian $T_1 = 2$ [ms], một yêu cầu (request, req) được gửi từ máy khách C tới máy chủ S và máy khách này nhận được phản hồi ở mốc thời gian $T_4 = 12$ [ms]. Thời gian xử lý yêu cầu này ở trên máy chủ (được coi như một tiến trình nội bộ: được xử lý độc lập và duy nhất bởi máy chủ) là 7 [ms]. Biết rằng thời gian truyền tin từ $C \rightarrow S$ và $S \rightarrow C$ là bằng nhau, ký hiệu là δT_{req} . **Tính** δT_{req} .

B. (1 điểm)



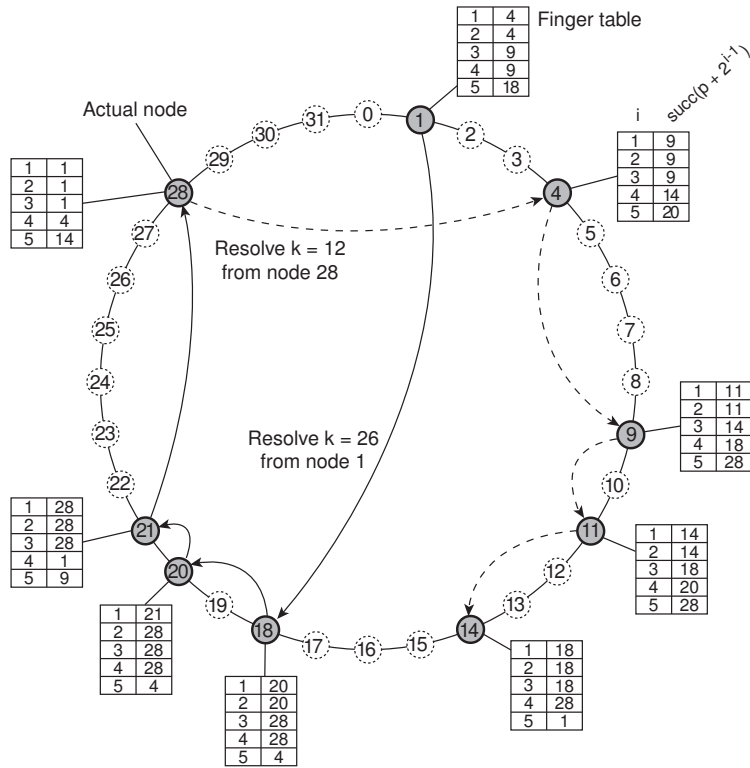
Hình 2: Sử dụng cho Câu B.

0.25 điểm Mô hình đồng hồ trong biểu diễn tiến trình ở Hình 2 là gì?

0.75 điểm Nếu mỗi đơn vị đồng hồ trên cả ba tiến trình biểu diễn 0.01 [ms] trong thời gian thực. Thời điểm tiến trình thứ 2 nhận được tin nhắn m_3 từ tiến trình thứ nhất là bao nhiêu tính trong thời gian thực? Biết rằng $P_2(0, 1, 0) = 0$ [ms].

Thời Gian	Số Xe Yêu Cầu Sạc
1 (6:00-10:00)	N_1
2 (10:00-13:00)	N_2
3 (13:00-17:00)	N_3
4 (17:00-24:00)	N_4
5 (24:00-6:00)	N_5

Bảng 1: Finger table



Hình 3: Hệ Thống Phân Tán Quản Lý Sạc Thông Minh cho Xe Điện

Câu 3 (1 điểm)

Một start-up về công nghệ nhận được một dự án: "Xây dựng hệ thống sạc thông minh trong nội thành Hà Nội". Công ty này dự kiến xây dựng 32 cụm máy phân tán (trạm), đánh số từ 0 tới 31, biểu diễn ở Hình 3. Dữ liệu trên mỗi máy được biểu diễn bằng một Finger table có cấu trúc như Bảng 1. (N_i là số xe yêu cầu sạc tại mốc thời gian i). Sử dụng thuật toán Map-Reduce để tính số lượng xe yêu cầu sạc từ 6:00 tới 17:00 tại trạm 1, 4 và 11.

Ghi chú: Cán bộ coi thi không giải thích gì thêm.
Các thông số cung cấp cho SV (nếu có)

Kết Thúc Đề Thi

Lưu ý: Giáo viên ra đề và trưởng khoa/bộ môn ký duyệt vào mặt sau của đề. Phần ngày thi và giờ thi bỏ trống và sẽ được cán bộ in sao đề thi ghi (xóa phần này khi làm đề thi)

Trưởng Khoa/Bộ môn

(Ký và ghi rõ họ tên, học vị và chức danh khoa học, nếu có)

Giảng viên ra đề thi