



## Ung dung Big Data: Truyen du lieu trong thoi gian thuc\_ Nhóm 02

[Nhà của tôi](#) / [Các khoá học của tôi](#) / [BDAS436177\\_23\\_1\\_02](#) / [Day 14: Practice for final exam #1](#) / [Bài mock exam #1](#)

**Bắt đầu vào lúc** Friday, 1 December 2023, 2:12 PM

**Trạng thái** Đã xong

**Kết thúc lúc** Friday, 1 December 2023, 2:25 PM

**Thời gian thực hiện** 12 phút 26 giây

**Điểm** 10,00 trên 10,00 (100%)

### Câu hỏi 1

Đúng

Đạt điểm 0,25 trên 0,25

Trong kiến trúc publish/subscribe pattern, gói tin gửi đến client là đầu ra của ...

- ☐ a. Topic
- ☐ b. Producer
- ☒ c. Subscription
- ☐ d. Slave



Your answer is correct.

### Câu hỏi 2

Đúng

Đạt điểm 0,25 trên 0,25

Khác biệt giữa request/response và request/acknowledge là?

- ☐ a. Định dạng dữ liệu trong gói tin yêu cầu
- ☒ b. Mục đích gói tin hồi đáp
- ☐ c. Cách thức xác thực
- ☐ d. Giao thức mạng



Your answer is correct.

### Câu hỏi 3

Đúng

Đạt điểm 0,25 trên 0,25

Xe thông minh gửi thông tin vị trí cho hệ thống để nhận về tình trạng giao thông tuyến đường gần đó và thể hiện lên màn hình dữ liệu vừa nhận. Hệ thống trên thuộc pattern nào?

- ☐ a. Request/acknowledge pattern
- ☐ b. One-way pattern
- ☒ c. Request/response pattern
- ☐ d. Publish/subscribe pattern



Your answer is correct.

### Câu hỏi 4

Đúng

Đạt điểm 0,25 trên 0,25

Ví dụ nào dưới đây là streaming pattern?

- ☐ a. Đọc báo từ trang chủ tuoitre.vn
- ☐ b. Tất cả đều sai
- ☒ c. Nghe nhạc trên Spotify
- ☐ d. Ứng dụng cập nhật mail cứ mỗi 5 phút



Your answer is correct.

### Câu hỏi 5

Đúng

Đạt điểm 0,25 trên 0,25

Nếu một gói tin không thể bị mất và được đọc một lần duy nhất bởi consumer thì đó là kiểu đảm bảo gì?

- ☐ a. At most once
- ☐ b. Once
- ☒ c. Exactly once
- ☐ d. At least once



Your answer is correct.

## Câu hỏi 6

Đúng

Đạt điểm 0,25 trên 0,25

Nếu một gói tin không thể bị mất và có thể được đọc nhiều lần bởi consumer thì đó là kiểu đảm bảo gì?

- ☐ a. Once
- ☐ b. Exactly once
- ☐ c. At most once
- ☒ d. At least once



Your answer is correct.

## Câu hỏi 7

Đúng

Đạt điểm 0,25 trên 0,25

Mô tả nào sau đây khớp với sender-based message logging?

- ☒ a. Gói tin được lưu ngay trước khi gửi.
- ☐ b. Gói tin được lưu ngay sau khi nhận.
- ☐ c. Tất cả đều sai.
- ☐ d. Gói tin được lưu ngay trước khi xử lý.



Your answer is correct.

## Câu hỏi 8

Đúng

Đạt điểm 0,25 trên 0,25

Đâu là phương pháp kháng lỗi cho tầng thu thập dữ liệu?

- ☐ a. Imaging
- ☒ b. Logging
- ☐ c. Recovery
- ☐ d. Clone



Your answer is correct.

## Câu hỏi 9

Đúng

Đạt điểm 0,25 trên 0,25

Điều nào sau đây là đúng với checkpointing?

- ☐ a. Mỗi nút ít nhất phải có một checkpoint.
- ☐ b. Tất cả đều sai.
- ☒ c. Mỗi nút có thể có số checkpoint khác nhau trong cùng khoảng thời gian.
- ☐ d. Trong một nút, các checkpoint sau phụ thuộc vào checkpoint trước.



Your answer is correct.

## Câu hỏi 10

Đúng

Đạt điểm 0,25 trên 0,25

Trong logging, thuật ngữ tiếng Anh nào ám chỉ việc gửi lại gói tin?

- ☐ a. Recreate
- ☐ b. Resend
- ☐ c. Repeat
- ☒ d. Replay



Your answer is correct.

## Câu hỏi 11

Đúng

Đạt điểm 0,25 trên 0,25

Kiểu đảm bảo nào dưới đây có hiệu năng trung bình?

- ☐ a. At most once
- ☐ b. Exactly once
- ☐ c. Once
- ☒ d. At least once



Your answer is correct.

## Câu hỏi 12

Đúng

Đạt điểm 0,25 trên 0,25

Hiện tượng nào sau đây thuộc nhóm lỗi response failure?

- ☐ a. Tất cả đều sai.
- ☒ b. Gói tin bị mất một phần khi đến nơi.
- ☐ c. Gói tin bị mất hoàn toàn trong khi gửi.
- ☐ d. Hệ thống không phản hồi ACK sau khi nhận gói tin.



Your answer is correct.

## Câu hỏi 13

Đúng

Đạt điểm 0,25 trên 0,25

Lợi thế của phương pháp concurrent checkpointing?

- ☐ a. Sử dụng file system đặc dụng.
- ☒ b. Loại bỏ sự phụ thuộc vào checkpoint bất kỳ.
- ☐ c. Thời gian khôi phục nhanh.
- ☐ d. Không cần dung lượng lưu trữ lớn.



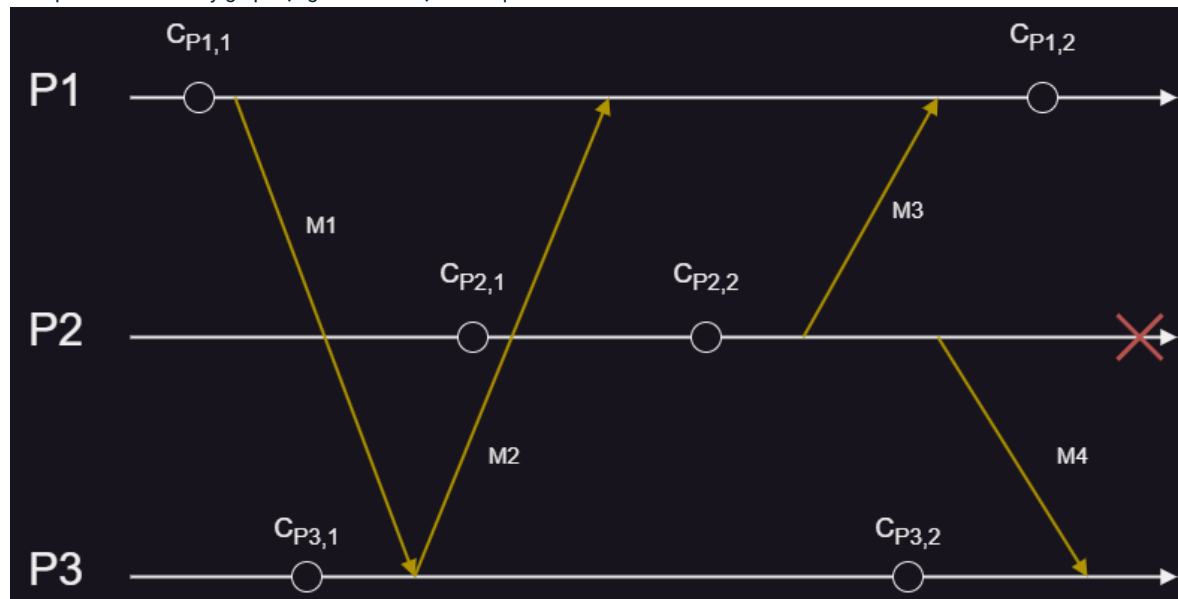
Your answer is correct.

## Câu hỏi 14

Đúng

Đạt điểm 0,25 trên 0,25

Cho biểu đồ thể hiện quá trình trao đổi gói tin giữa các tiến trình như sau. Trong trường hợp một trong các hệ thống bị lỗi tại X, tập checkpoint nào sau đây giúp trạng thái toàn cục nhất quán?



- ☐ a.  $\{CP_{1,2}, CP_{2,2}, CP_{3,1}\}$
- ☐ b.  $\{CP_{1,1}, CP_{2,1}, CP_{3,2}\}$
- ☐ c.  $\{CP_{1,2}, CP_{2,2}, CP_{3,2}\}$
- ☒ d.  $\{CP_{1,1}, CP_{2,1}, CP_{3,1}\}$



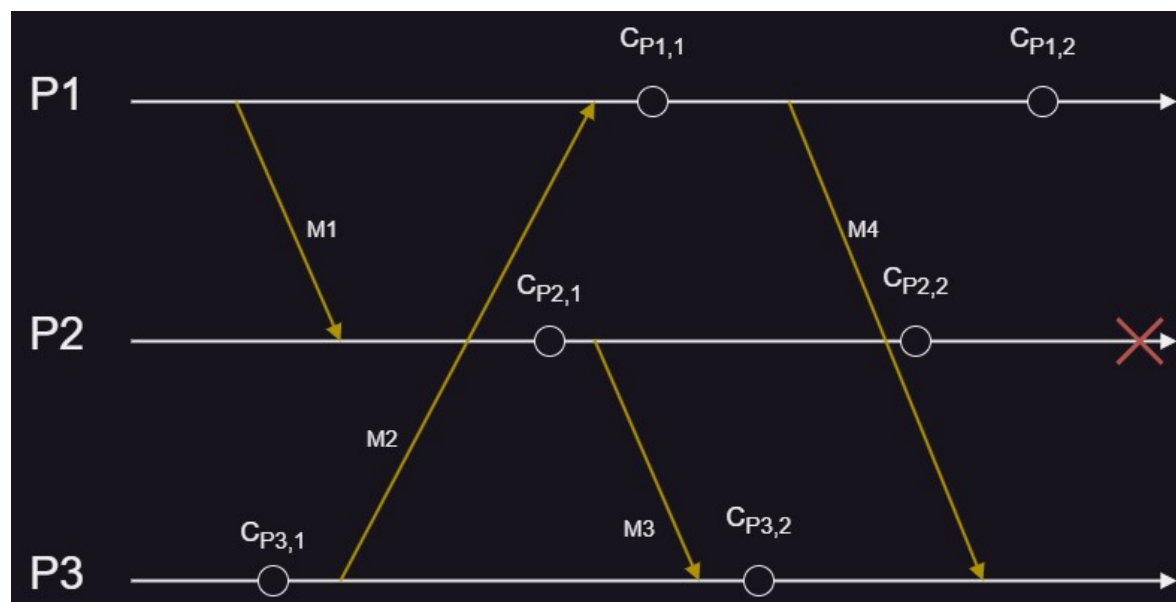
Your answer is correct.

## Câu hỏi 15

Đúng

Đạt điểm 0,25 trên 0,25

Cho biểu đồ thể hiện quá trình trao đổi gói tin giữa các tiến trình như sau. Trong trường hợp một trong các hệ thống bị lỗi tại X đỏ, tập checkpoint nào sau đây khiến trạng thái toàn cục KHÔNG nhất quán?



- ☐ a. Tất cả đều sai
- ☐ b.  $\{CP_{1,2}, CP_{2,2}, CP_{3,2}\}$
- ☒ c.  $\{CP_{1,2}, CP_{2,1}, CP_{3,2}\}$
- ☐ d.  $\{CP_{1,1}, CP_{2,2}, CP_{3,2}\}$

✓

Your answer is correct.

## Câu hỏi 16

Đúng

Đạt điểm 0,25 trên 0,25

Hai cổng nào của RabbitMQ server được khởi chạy mặc định?

- ☐ a. 5674, 25674
- ☐ b. 5673, 25673
- ☒ c. 5672, 25672
- ☐ d. 5671, 25671

✓

Your answer is correct.

## Câu hỏi 17

Đúng

Đạt điểm 0,25 trên 0,25

Trong RabbitMQ, Persistent Delivery Mode và Durable lần lượt miêu tả tính chất của?

- ☒ a. Gói tin và hàng đợi
- ☐ b. Hàng đợi và gói tin
- ☐ c. Hàng đợi và broker
- ☐ d. Broker và hàng đợi



Your answer is correct.

## Câu hỏi 18

Đúng

Đạt điểm 0,25 trên 0,25

Trong DNS-based discovery, giả sử máy A cần tìm host của máy B, IP trung gian có vai trò?

- ☐ a. Hỗ trợ truy vấn host của máy A.
- ☒ b. Hỗ trợ truy vấn host của máy B.
- ☐ c. Là IP của máy B.
- ☐ d. Là IP của máy A.



Your answer is correct.

## Câu hỏi 19

Đúng

Đạt điểm 0,25 trên 0,25

Một topic trong Kafka bao gồm các hàng dữ liệu được gọi là?

- ☐ a. Stack
- ☐ b. Bin
- ☐ c. Dataframe
- ☒ d. Partition



Your answer is correct.



**Câu hỏi 20**

Đúng

Đạt điểm 0,25 trên 0,25

Trong thuật toán HyperLogLog, số lượng bit lấy ban đầu trong dãy bit dữ liệu được xác định bằng công thức nào dưới đây?

- ☒ a.  $\log_2(m)$
- ☐ b.  $\log_{10}(m^2)$
- ☐ c.  $\log_2(m^2)$
- ☐ d.  $\log_{10}(m)$



Your answer is correct.

**Câu hỏi 21**

Đúng

Đạt điểm 0,25 trên 0,25

Trong phương pháp tumbling của tăng phân tích, chính sách loại bỏ là?

- ☐ a. Khi không có dữ liệu
- ☒ b. Khi đã đầy dữ liệu
- ☐ c. Cả 3 câu đều sai
- ☐ d. Khi đạt được lượng dữ liệu cho trước



Your answer is correct.

**Câu hỏi 22**

Đúng

Đạt điểm 0,25 trên 0,25

Hai kiểu phương pháp tumbling của tăng phân tích là?

- ☒ a. Count-based, temporal
- ☐ b. Count-based, time
- ☐ c. Quantity-base, temporal
- ☐ d. Quantity-base, time



Your answer is correct.

## Câu hỏi 23

Đúng

Đạt điểm 0,25 trên 0,25

Đâu là bất đẳng thức chính xác nhất khi so sánh stream time và event time?

- ☒ a. Stream time  $\geq$  event time
- ☐ b. Stream time  $>$  event time
- ☐ c. Stream time  $<$  event time
- ☐ d. Stream time  $\leq$  event time



Your answer is correct.

## Câu hỏi 24

Đúng

Đạt điểm 0,25 trên 0,25

Ví dụ nào dưới đây liên quan đến ràng buộc concept drift?

- ☒ a. Hệ thống gợi ý không còn chính xác do mô hình bị lỗi thời.
- ☐ b. Số lượng người dùng thường xuyên truy cập trang web ở mức 500-1000 người/giờ.
- ☐ c. Vì một sự kiện giảm giá, lượng dữ liệu vào hệ thống streaming tăng đột biến.
- ☐ d. Dữ liệu đã xử lý không thể được xử lý lại.



Your answer is correct.

## Câu hỏi 25

Đúng

Đạt điểm 0,25 trên 0,25

Trong tầng phân tích, công cụ thực hiện chính sách kích hoạt của phương pháp cửa sổ trượt là?

- ☒ a. Cửa sổ trượt
- ☐ b. Cửa sổ thời gian
- ☐ c. Cửa sổ dữ liệu
- ☐ d. Dòng thời gian



Your answer is correct.

**Câu hỏi 26**

Đúng

Đạt điểm 0,25 trên 0,25

Trong Spark Streaming, lớp trừu tượng thể hiện stream dữ liệu là?

- ☒ a. DStream
- ☐ b. DtStream
- ☐ c. DataStream
- ☐ d. Stream



Your answer is correct.

**Câu hỏi 27**

Đúng

Đạt điểm 0,25 trên 0,25

Trong hệ thống streaming data, dữ liệu có n bảng ghi vào thì số lượng bảng ghi được phân tích thường là?

- ☐ a. Bằng n
- ☐ b. Lớn hơn n
- ☐ c. Bằng 0
- ☒ d. Nhỏ hơn n



Your answer is correct.

**Câu hỏi 28**

Đúng

Đạt điểm 0,25 trên 0,25

Cấu hình cửa sổ trượt cho tăng phân tích như sau: Kích thước cửa sổ là 20 giây, thanh trượt là 5 giây. Khi chạy thử hệ thống, người ta nhận thấy thời gian dữ liệu vào hệ thống sớm nhất là t giây thì dữ liệu ra trễ nhất là t + 60 giây. Tính số cửa sổ nằm gọn trong sự kiện.

- ☐ a. 15
- ☐ b. 6
- ☒ c. 9
- ☐ d. 14



Your answer is correct.

**Câu hỏi 29**

Đúng

Đạt điểm 0,25 trên 0,25

Cấu hình cửa sổ trượt cho tầng phân tích như sau: Kích thước cửa sổ là 20 giây, thanh trượt là 5 giây. Khi chạy thử hệ thống, người ta nhận thấy thời gian xử lý vào hệ thống sớm nhất là  $t$  giây thì dữ liệu ra trễ nhất là  $t + 60$  giây. Tính số cửa sổ cần quan tâm.

- ☒ a. 15
- ☐ b. 14
- ☐ c. 9
- ☐ d. 6



Your answer is correct.

**Câu hỏi 30**

Đúng

Đạt điểm 0,25 trên 0,25

Đâu là tên một công nghệ in-memory database (IMDB)?

- ☒ a. Apache Geode
- ☐ b. Redis
- ☐ c. SQLite
- ☐ d. IndexedDB



Your answer is correct.

**Câu hỏi 31**

Đúng

Đạt điểm 0,25 trên 0,25

Trong thuật toán Count-min sketch (CMS), ma trận có số cột là?

- ☐ a. Giới hạn giá trị băm
- ☒ b. Giới hạn giá trị băm của hàm băm lớn nhất
- ☐ c. Tần suất dữ liệu
- ☐ d. Số lượng hàm băm



Your answer is correct.

**Câu hỏi 32**

Đúng

Đạt điểm 0,25 trên 0,25

Write-back là phương pháp caching nâng cấp từ phương pháp?

- ☐ a. Refresh-ahead
- ☐ b. Read-through
- ☐ c. Write-around
- ☒ d. Tất cả đều sai



Your answer is correct.

**Câu hỏi 33**

Đúng

Đạt điểm 0,25 trên 0,25

Thiết bị Universal Serial Bus thuộc kiểu nào dưới đây?

- ☐ a. NVMe
- ☐ b. Lưu trữ ngắn hạn
- ☐ c. Bộ nhớ
- ☒ d. Lưu trữ dài hạn



Your answer is correct.

**Câu hỏi 34**

Đúng

Đạt điểm 0,25 trên 0,25

Cho ma trận Count-min sketch như hình sau.

18	44	47	77
15	24	5	57
67	13	31	2
8	4	5	43

Giả sử kết quả các hàm băm fi (với i là chỉ mục hàng trong ma trận, bắt đầu từ 0) như sau:

f0("bigdata") = 1

f1("bigdata") = 2

f2("bigdata") = 2

f3("bigdata") = 0

Tần suất ước lượng của dữ liệu chuỗi "bigdata" là?

- ☐ a. 8
- ☐ b. 4
- ☐ c. 2
- ☒ d. 5



Your answer is correct.

**Câu hỏi 35**

Đúng

Đạt điểm 0,25 trên 0,25

Cho ma trận Count-min sketch như hình sau.

41	44	28	23	43
40	11	17	36	45
49	43	26	33	20
30	42	40	40	19
28	10	39	17	44

Giả sử kết quả các hàm băm  $f_i$  (với  $i$  là chỉ mục hàng trong ma trận, bắt đầu từ 0) như sau: $f_0(\text{"cntt"}) = 3$  $f_1(\text{"cntt"}) = 1$  $f_2(\text{"cntt"}) = 0$  $f_3(\text{"cntt"}) = 1$  $f_4(\text{"cntt"}) = 4$ 

Tần suất ước lượng của dữ liệu chuỗi "cntt" là?

- ☐ a. 20
- ☒ b. 11
- ☐ c. 10
- ☐ d. 17



Your answer is correct.

**Câu hỏi 36**

Đúng

Đạt điểm 0,25 trên 0,25

Cho ma trận Count-min sketch như hình sau.

18	44	47	77
15	24	5	57
67	13	31	2
8	4	5	43

Giả sử kết quả các hàm băm fi (với i là chỉ mục hàng trong ma trận, bắt đầu từ 0) như sau:

 $f_0(\text{"ktdl"}) = 2$  $f_1(\text{"ktdl"}) = 1$  $f_2(\text{"ktdl"}) = 1$  $f_3(\text{"ktdl"}) = 0$ 

Tần suất ước lượng của dữ liệu chuỗi "ktdl" là?

- ☒ a. 8
- ☐ b. 23
- ☐ c. 2
- ☐ d. 5



Your answer is correct.

**Câu hỏi 37**

Đúng

Đạt điểm 0,25 trên 0,25

Một hàm băm cho ra kết quả băm với độ dài không đổi là 20 bytes. Khi thể hiện kết quả đó dưới dạng số nhị phân thì ta sẽ có bao nhiêu chữ số?

- ☐ a. 140
- ☐ b. 200
- ☐ c. 120
- ☒ d. 160



Your answer is correct.



**Câu hỏi 38**

Đúng

Đạt điểm 0,25 trên 0,25

Một hàm băm cho ra kết quả băm với độ dài không đổi là 8 bits. Giả sử ta giới hạn giá trị nhỏ nhất của kết quả là 0, giá trị số nguyên tối đa mà hàm có thể thể hiện là?

- ☐ a. 63
- ☒ b. 255
- ☐ c. 64
- ☐ d. 256



Your answer is correct.

**Câu hỏi 39**

Đúng

Đạt điểm 0,25 trên 0,25

Sử dụng thuật toán Count-min sketch, cần hàm băm tối đa bao nhiêu bits để dùng với ma trận 10 cột?

- ☒ a. 3
- ☐ b. 2
- ☐ c. 5
- ☐ d. 4



Your answer is correct.

**Câu hỏi 40**

Đúng

Đạt điểm 0,25 trên 0,25

Sử dụng thuật toán Count-min sketch, cần hàm băm tối đa bao nhiêu bits để dùng với ma trận 15 cột?

- ☒ a. 3
- ☐ b. 5
- ☐ c. 4
- ☐ d. 2



Your answer is correct.

◀ **Bài tập cộng điểm (nhóm)**

Chuyển tới...

