



Ung dung Big Data: Truyen du lieu trong thoi gian thuc_ Nhóm 01

[Nhà của tôi](#) / [Các khoá học của tôi](#) / [BDAS436177_23_1_01](#) / [Day 13: Practice for final exam #1](#) / [Bài mock exam #1](#)

Bắt đầu vào lúc Thursday, 16 November 2023, 3:49 PM

Trạng thái Đã xong

Kết thúc lúc Thursday, 16 November 2023, 3:53 PM

Thời gian thực hiện 3 phút 27 giây

Điểm 10,00 trên 10,00 (100%)

Câu hỏi 1

Hoàn thành

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 8 cột?

- ☒ a. 3
- ☐ b. 2
- ☐ c. 4
- ☐ d. 5

Your answer is correct.

Câu hỏi 2

Hoàn thành

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 < event time
- ☐ b. Stream time > event time
- ☒ c. Stream time >= event time
- ☐ d. Stream time <= event time

Your answer is correct.

Câu hỏi 3

Hoàn thành

Đạt điểm 0,25 trên 0,25

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

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

Your answer is correct.

Câu hỏi 4

Hoàn thành

Đạt điểm 0,25 trên 0,25

Một hàm băm cho ra kết quả băm với độ dài không đổi là 1 byte. 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. 255
- ☐ b. 1023
- ☐ c. 1024
- ☐ d. 256

Your answer is correct.

Câu hỏi 5

Hoàn thành

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

Your answer is correct.

Câu hỏi 6

Hoàn thành

Đạt điểm 0,25 trên 0,25

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

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

Your answer is correct.

Câu hỏi 7

Hoàn thành

Đạt điểm 0,25 trên 0,25

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

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

Your answer is correct.

Câu hỏi 8

Hoàn thành

Đạt điểm 0,25 trên 0,25

Ví dụ nào dưới đây là streaming pattern?

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

Your answer is correct.

Câu hỏi 9

Hoàn thành

Đạt điểm 0,25 trên 0,25

Cho ma trận Count-min sketch như hình sau.

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

$f_0(\text{"xinchao"}) = 2$

$f_1(\text{"xinchao"}) = 1$

$f_2(\text{"xinchao"}) = 1$

$f_3(\text{"xinchao"}) = 0$

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

- ☐ a. 75
- ☐ b. 31
- ☒ c. 42
- ☐ d. 35

Your answer is correct.

Câu hỏi 10

Hoàn thành

Đạt điểm 0,25 trên 0,25

Đâu là tên của một công nghệ caching?

- ☐ a. InfluxDB
- ☐ b. Apache Ignite
- ☒ c. Hazelcast
- ☐ d. Perset

Your answer is correct.

Câu hỏi 11

Hoàn thành

Đạt điểm 0,25 trên 0,25

Cấu hình cửa sổ trượt cho tầng phân tích như sau: Kích thước cửa sổ là 15 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 $t + 30$ giây. Tính số cửa sổ nằm gọn trong khoảng sự kiện.

- ☐ a. 3
- ☐ b. 1
- ☒ c. 4
- ☐ d. 8

Your answer is correct.

Câu hỏi 12

Hoàn thành

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. $\{C_{P1,2}, C_{P2,1}, C_{P3,2}\}$
- ☐ b. Tất cả đều sai
- ☐ c. $\{C_{P1,2}, C_{P2,2}, C_{P3,2}\}$
- ☐ d. $\{C_{P1,1}, C_{P2,2}, C_{P3,2}\}$

Your answer is correct.

Câu hỏi 13

Hoàn thành

Đạt điểm 0,25 trên 0,25

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

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

Your answer is correct.

Câu hỏi 14

Hoàn thành

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

Your answer is correct.

Câu hỏi 15

Hoàn thành

Đạt điểm 0,25 trên 0,25

Cho ma trận Count-min sketch như hình sau.

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. 5
- ☒ b. 8
- ☐ c. 2
- ☐ d. 23

Your answer is correct.

Câu hỏi 16

Hoàn thành

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

Your answer is correct.

Câu hỏi 17

Hoàn thành

Đạt điểm 0,25 trên 0,25

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

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

Your answer is correct.

Câu hỏi 18

Hoàn thành

Đạt điểm 0,25 trên 0,25

Cho biểu đồ thể hiện quá trình trao đổi gói tin giữa các tiến trình như sau. Trạng thái toàn cục tại thời điểm checkpoint $C_{P3,2}$ được khởi tạo là?

- ☐ a. Gửi M1, nhận M2, gửi M3.
- ☐ b. Nhận M1, gửi M2, nhận M3.
- ☒ c. Nhận M1, nhận M2, gửi M3.
- ☐ d. Nhận M1, gửi M2, gửi M3.

Your answer is correct.

Câu hỏi 19

Hoàn thành

Đạt điểm 0,25 trên 0,25

Hai kiểu phương pháp tumbling của tăng phân tích là?

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

Your answer is correct.

Câu hỏi 20

Hoàn thành

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

Your answer is correct.

Câu hỏi 21

Hoàn thành

Đạt điểm 0,25 trên 0,25

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

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

Your answer is correct.

Câu hỏi 22

Hoàn thành

Đạt điểm 0,25 trên 0,25

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

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

Your answer is correct.

Câu hỏi 23

Hoàn thành

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

Your answer is correct.

Câu hỏi 24

Hoàn thành

Đạt điểm 0,25 trên 0,25

Cấu hình cửa sổ trượt cho tăng phân tích như sau: Kích thước cửa sổ là 35 giây, thanh trượt là 7 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 $t + 70$ giây. Tính số cửa sổ nằm gọn trong khoảng sự kiện.

- ☐ a. 4
- ☒ b. 6
- ☐ c. 8
- ☐ d. 14

Your answer is correct.

Câu hỏi 25

Hoàn thành

Đạt điểm 0,25 trên 0,25

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

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

Your answer is correct.

Câu hỏi 26

Hoàn thành

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

Your answer is correct.

Câu hỏi 27

Hoàn thành

Đạt điểm 0,25 trên 0,25

Cho ma trận Count-min sketch như hình sau.

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

$$f_0(\text{"fit"}) = 0$$

$$f_1(\text{"fit"}) = 3$$

$$f_2(\text{"fit"}) = 2$$

$$f_3(\text{"fit"}) = 1$$

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

- ☐ a. 9
- ☐ b. 26
- ☐ c. Không thể xác định
- ☒ d. 0

Your answer is correct.

Câu hỏi 28

Hoàn thành

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

Your answer is correct.

Câu hỏi 29

Hoàn thành

Đạt điểm 0,25 trên 0,25

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

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

Your answer is correct.

Câu hỏi 30

Hoàn thành

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 32 cột?

- ☒ a. 5
- ☐ b. 8
- ☐ c. 7
- ☐ d. 6

Your answer is correct.

Câu hỏi 31

Hoàn thành

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 32

Hoàn thành

Đạt điểm 0,25 trên 0,25

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

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

Your answer is correct.

Câu hỏi 33

Hoàn thành

Đạt điểm 0,25 trên 0,25

Hiện tượng nào sau đây thuộc nhóm lỗi response failure?

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

Your answer is correct.

Câu hỏi 34

Hoàn thành

Đạt điểm 0,25 trên 0,25

Lợi thế của phương pháp concurrent checkpointing?

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

Your answer is correct.

Câu hỏi 35

Hoàn thành

Đạt điểm 0,25 trên 0,25

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

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

Your answer is correct.

Câu hỏi 36

Hoàn thành

Đạt điểm 0,25 trên 0,25

Thiết bị Universal Serial Bus thuộc kiểu nào dưới đây?

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

Your answer is correct.

Câu hỏi 37

Hoàn thành

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

Your answer is correct.

Câu hỏi 38

Hoàn thành

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. Slave
- ☒ d. Subscription

Your answer is correct.

Câu hỏi 39

Hoàn thành

Đạt điểm 0,25 trên 0,25

Một hàm băm cho ra kết quả băm với độ dài không đổi là 8 bytes. Khi thể hiện kết quả đó dưới dạng số nhị phân thì ta sẽ có bao nhiêu chữ số?

- ☒ a. 64
- ☐ b. 8
- ☐ c. 16
- ☐ d. 32

Your answer is correct.

Câu hỏi 40

Hoàn thành

Đạt điểm 0,25 trên 0,25

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

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

Your answer is correct.

◀ **Bài tập cộng điểm (nhóm)**

Chuyển tới...

Short essay exercise (group) ▶