

SISTEMI DISTRIBUITI E CLOUD COMPUTING A.A. 2023/24
Seconda prova intermedia - 22/1/2024

Cognome _____ **Nome** _____

Matricola _____

Domanda 1 (punti 9)

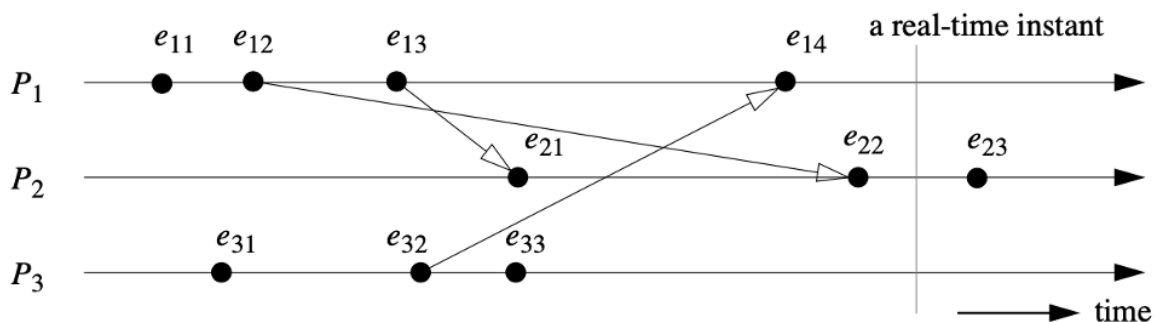
- a) Con riferimento alla migrazione live delle macchine virtuali, spiegare l'approccio pre-copy per la migrazione della memoria. In presenza di un workload di tipo write-intensive rispetto alla memoria, quale problema presenta questo approccio e quale soluzione può essere adottata per risolvere il problema?
- b) Considerando il seguente Dockerfile, spiegare la strutturazione in layer dell'immagine che verrà costruita.

```
FROM python:3.7-alpine
WORKDIR /code
ENV FLASK_APP=app.py
ENV FLASK_RUN_HOST=0.0.0.0
RUN apk add --no-cache gcc musl-dev linux-headers
COPY requirements.txt requirements.txt
RUN pip install -r requirements.txt
EXPOSE 5000
COPY . .
CMD ["flask", "run"]
```

- c) Spiegare cosa è un volume in Docker e quali vantaggi introduce.
- d) Perché l'uso di un volume agevola la scalabilità orizzontale e la migrazione dei container?
- e) Con riferimento alle applicazioni a microservizi, si presenti un pattern a scelta tra circuit breaker, CQRS e database per service, indicando anche quale obiettivo persegue.
- f) Si spieghi come Kubernetes fornisce auto-scaling.

Domanda 2 (punti 8)

- a) Si presenti l'algoritmo di sincronizzazione fisica di Cristian, spiegando anche se l'algoritmo può essere usato in un sistema distribuito asincrono, come viene determinata l'accuratezza e come è possibile migliorarla; infine, si discutano i possibili vantaggi e svantaggi dell'algoritmo.
- b) Descrivere i protocolli di aggiornamento dei clock scalare e vettoriale e spiegare perché il clock vettoriale aumenta la conoscenza sul sistema rispetto al clock scalare.
- c) Si determinino i valori del clock vettoriale di tutti gli eventi nel sottostante diagramma temporale.



- d) Dato il valore del clock vettoriale determinato al punto c), si discuta quali delle seguenti affermazioni è vera, motivando opportunamente la risposta.
1. $e_{12} \rightarrow e_{33}$
 2. $e_{21} \rightarrow e_{14}$
 3. $e_{11} \rightarrow e_{23}$

Cosa si può invece affermare conoscendo solo il valore del clock scalare degli eventi?

- e) Discutere se nel diagramma sopra riportato si verifica una violazione della causalità, motivando la risposta.

Domanda 3 (punti 8)

- a) Si spieghino le differenze di garanzie tra consistenza causale, linearizzabile e sequenziale.
- b) Si presenti un esempio di archivio di dati distribuito che soddisfa la consistenza sequenziale ma non quella linearizzabile e un archivio che soddisfa la consistenza causale ma non quella linearizzabile e sequenziale. *NB: gli esempi devono essere diversi da quelli in c).*
- c) Determinare qual è il massimo grado di consistenza data centrica soddisfatto dagli archivi di dati distribuiti sottostanti, motivando opportunamente la risposta.

P1: W(y)1

P2: W(y)2 R(y)1 R(y)2

P1: W(x)a W(x)b R(x)c W(x)d

P2: R(x)a W(x)c R(x)d

P3: R(x)a R(x)c R(x)b

- d) Si spieghi l'algoritmo del multicast causalmente ordinato, spiegando inoltre se una variante di questo algoritmo può essere eventualmente utilizzata per implementare un protocollo di consistenza causale.

Domanda 4 (punti 7)

- a) Si presenti un algoritmo di mutua esclusione distribuita per un sistema distribuito a scelta tra Lamport distribuito, Ricart-Agrawala e Maekawa.
- b) Nell'algoritmo considerato, come viene garantita la proprietà di safety?
- c) Cosa accade rispetto alla proprietà di liveness se nell'algoritmo considerato un processo subisce un crash? Quale meccanismo occorre adottare per far fronte ad eventi di crash di un processo?
- d) Nell'algoritmo di elezione ad anello, spiegare se e perché vengono garantite le proprietà di safety e liveness se due processi avviano contemporaneamente l'elezione, assumendo che nel frattempo nessun processo subisca un crash.