

SISTEMI DISTRIBUITI E CLOUD COMPUTING A.A. 2017/18
Seconda prova intermedia - 12/01/2018

Cognome _____ **Nome** _____

Matricola _____

Domanda 1 (punti 6)

- a) Si spieghi come funziona il full content delivery attuato dalla CDN Akamai.
- b) Quali problematiche presenta in generale il meccanismo di redirectione basato su DNS e quali sono le possibili soluzioni?
- c) Analizzando il seguente output del comando `dig www.repubblica.it` individuare la soluzione di distribuzione del contenuto adottata dal sito del quotidiano La Repubblica, motivando opportunamente la risposta.

```
>>> dig 9.8.3-P1 <<>> www.repubblica.it
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 63674
;; flags: qr rd ra; QUERY: 1, ANSWER: 3, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;www.repubblica.it.      IN      A

;; ANSWER SECTION:
www.repubblica.it. 60      IN      CNAME      www.repubblica.it.edgekey.net.
www.repubblica.it.edgekey.net. 13695 IN CNAME      e7047.e12.akamaiedge.net.
e7047.e12.akamaiedge.net. 20      IN      A          95.100.93.182

;; Query time: 29 msec
;; SERVER: 192.168.1.1#53(192.168.1.1)
;; WHEN: Fri Jan 12 07:40:36 2018
;; MSG SIZE rcvd: 129
```

Domanda 2 (punti 6)

- a) Si descrivano i protocolli quorum-based per la consistenza.
- b) Quale modello di consistenza supportano?
- c) Il database NoSQL Cassandra offre diversi livelli di consistenza, tra cui ALL ($N_R=N_W=N$), ONE ($N_R=N_W=1$) e QUORUM ($N_R=N_W=N/2+1$). Quali sono le differenze tra le tre configurazioni in termini di latenza, consistenza e disponibilità, considerando anche che Cassandra può operare su una scala geografica? Quale delle tre configurazioni è più scalabile all'aumentare del numero di repliche N?

Domanda 3 (punti 7)

- a) Si definisca la consistenza sequenziale.
- b) Si discutano le differenze tra consistenza sequenziale e linearizzabile e tra sequenziale e causale.
- c) L'archivio di dati sottostante soddisfa la consistenza sequenziale? Motivare la risposta.

```
P0:      R(x)a  W(x)b                W(x)a      R(x)b
P1:                        R(x)a  R(x)b
P2:      W(x)b  R(x)a                R(x)b
```

- d) L'archivio di dati sottostante soddisfa la consistenza causale? Motivare la risposta.

```
P0:      R(x)a  R(x)b                W(x)a      W(x)d
P1:      R(x)a  R(x)b                R(x)a  W(x)c
P2:                        R(x)b  R(x)a                R(x)a  R(x)d  R(x)c
```

Domanda 4 (punti 7)

- a) Si descriva l'obiettivo, le assunzioni rispetto ai modelli di sistema e di failure considerati dell'algoritmo dei generali bizantini oppure dell'algoritmo di Paxos, descrivendone brevemente il funzionamento.
- b) Sotto le assunzioni dell'algoritmo dei generali bizantini, si può raggiungere l'accordo bizantino in presenza di 5 generali leali e 2 traditori? Quante fasi sono necessarie e qual è l'ordine del numero di messaggi scambiati durante l'esecuzione dell'algoritmo? Motivare le risposte.

- c) Sotto le assunzioni dell'algoritmo di Paxos, può verificarsi la non terminazione dell'algoritmo e perché? Come può essere affrontato questo problema?

Domanda 5 (punti 6)

Si supponga di dover realizzare un'applicazione distribuita per lo storage di file (tipo Dropbox) in grado di servire un vasto insieme di utenti distribuiti su scala geografica. L'applicazione consente agli utenti registrati le operazioni CRUD (Create, Read, Update, Delete) su file e la condivisione di file tra utenti. Si presenti una possibile architettura ad alto livello di un sistema a bassa latenza ed elevata disponibilità in grado di supportare tale applicazione, discutendone vantaggi e possibili svantaggi. Si discuta la relazione tra l'architettura proposta ed il teorema CAP ed il suo impatto sulle scelte progettuali. Discutere infine quali soluzioni possono essere adottate per minimizzare i conflitti di scrittura su un file condiviso tra un gruppo di utenti.