

SISTEMI DISTRIBUITI E CLOUD COMPUTING A.A. 2013/14
Seconda prova intermedia - 4/2/2014

| | |
|------------------------|-------------------|
| Cognome _____ | Nome _____ |
| Matricola _____ | |

Domanda 1 (punti 6)

Si definisca la consistenza causale e si consideri il seguente scenario in un sito di social network tipo Facebook:

1. Anna non trova suo figlio Marco. Aggiunge il post S sul suo account: “Temo che Marco sia scomparso”.
2. Subito dopo che Anna ha aggiunto il post, Marco la chiama al cellulare, avvisandola che sta giocando a casa di un amico. Anna modifica S, che diviene S*, in “Falso allarme! Marco era uscito a giocare”.
3. Laura, un’amica di Anna, legge S* e aggiunge in risposta il post J: “Che sollievo!”

Se Andrea, un amico di Anna, legge i post S e J ma non S*, cosa capisce? Se Facebook garantisce consistenza causale, può verificarsi che Andrea veda J senza aver visto S*?

Si consideri inoltre il seguente archivio di dati:

| | | | | |
|-----|-------|-------|-------|-------|
| P1: | R(y)a | R(y)c | W(x)b | R(x)b |
| P2: | W(y)b | R(x)a | R(y)c | R(x)b |
| P3: | R(x)b | W(x)c | W(y)c | R(x)c |
| P4: | R(x)a | R(y)b | W(x)d | W(y)d |

Si discuta se tale archivio di dati soddisfa la consistenza causale, motivando la risposta. Quale grado massimo di consistenza data-centrica è soddisfatto se l’operazione R(y)c viene rimossa da P1?

Domanda 2 (punti 6)

Si spieghi come funziona la redirectione DNS in un sistema replicato geograficamente. Perché un DNS server autoritativo che attua la redirectione DNS non riesce ad instradare la totalità delle richieste di risoluzione di indirizzo generate dai client e quali soluzioni possono essere adottate per affrontare tale problema? Infine, si discutano vantaggi e svantaggi dell’uso dell’anycast per il DNS, presentando un esempio reale di utilizzo di tale meccanismo.

Domanda 3 (punti 6)

Con riferimento alla tolleranza ai guasti, si considerino le seguenti domande.

- a) Qual è la cardinalità di un gruppo di processi tollerante a k guasti?
- b) Si consideri il protocollo primary-backup per la consistenza con un processo primario e due processi di backup. Per quale valore di k tale schema è tollerante a k guasti e sotto quali assunzioni? Si motivi opportunamente la risposta.
- c) Si consideri un gruppo tollerante a k guasti, con $k > 1$. Se un processo del gruppo fallisce, il gruppo è ancora tollerante a k guasti? Si motivi opportunamente la risposta.
- d) Si consideri un gruppo formato da N processi che eseguono l’algoritmo di Paxos. Assumendo un modello di failure di tipo crash, quanti processi falliti può tollerare il gruppo per continuare a soddisfare le specifiche?
- e) L’algoritmo di Paxos si compone di due fasi. Si spieghi la prima fase, fornendo il flusso di messaggi scambiati tra i proposer e gli acceptor e descrivendone i rispettivi stati e transizioni di stato.

Domanda 4 (punti 5)

Si descrivano gli approcci esaminati a lezione per realizzare un sistema di naming semplice, discutendo vantaggi e svantaggi. Si possono verificare guasti bizantini in un sistema di naming gerarchico come il DNS e come possono essere eventualmente affrontati?

Domanda 5 (punti 7)

Si supponga di dover realizzare un’applicazione distribuita per lo storage personale di file (tipo Dropbox) in grado di servire un’ampia popolazione di utenti distribuita su scala geografica. Si presenti una possibile architettura ad alto livello di un sistema altamente scalabile ed affidabile in grado di supportare tale applicazione, discutendo vantaggi e possibili svantaggi della soluzione presentata anche derivanti dal teorema CAP. Oltre alla registrazione di un utente, l’applicazione deve consentire le tradizionali operazioni CRUD su file e la condivisione di file tra gruppi di utenti. Discutere infine quali soluzioni possono essere adottate per ridurre l’overhead relativo all’aggiornamento di file, in particolare di grandi dimensioni.