

Sistemi Distribuiti - A.A. 2009/10

Progetto 1 per modalità B di esame:

File system distribuito transazionale con replicazione

Docente: Valeria Cardellini
Università di Roma "Tor Vergata"
cardellini@ing.uniroma2.it

Requisiti del progetto

Lo scopo del progetto è realizzare in linguaggio C usando l'API del socket di Berkeley un **file system distribuito di tipo transazionale con replicazione**. Una transazione è una sequenza di operazioni su file che soddisfano le proprietà ACID (Atomicità, Consistenza, Isolamento e Durabilità). Ciò significa che ogni sequenza di operazioni su file deve apparire come atomica, deve lasciare il file system in uno stato consistente, deve essere eseguita in modo isolato ed indipendente dalle altre transazioni e deve resistere a failure dei server o dei client del file system distribuito.

Il file system distribuito transazionale con replicazione deve supportare sia letture, sia scritture in base alla semantica specificata di seguito.

La semantica delle operazioni di lettura richiede che i file siano sempre letti interamente (ovvero non sono ammesse letture di parti di un file). In caso di lettura, il client del file system invia al server il nome del file da leggere ed il server restituisce il file completo.

In assenza di failure, la semantica delle operazioni di scrittura richiede che all'inizio di una transazione il client richieda al server un nuovo ID della transazione, specificando il file che sarà scritto durante la transazione. Alla ricezione della richiesta, il server genera un ID univoco della transazione e lo invia al client. Se il file specificato dal client non esiste, il server lo crea. Tutti i successivi messaggi client/server appartenenti alla transazione conterranno l'ID della transazione. Il client, ricevuto l'ID della transazione, invia al server una sequenza di operazioni di scrittura sul file oggetto della transazione e chiede al server di effettuare il commit della transazione. Alla ricezione della richiesta di commit, il server risponde con un ack. Il file scritto non deve essere visibile nel file system finché il server non effettua il commit della transazione.

Il file system distribuito transazionale con replicazione deve soddisfare le caratteristiche elencate di seguito.

- La comunicazione tra client e server deve avvenire usando TCP come protocollo di trasporto.
- La comunicazione tra client e server deve avvenire usando un protocollo applicativo da definire.
- Il client del file system distribuito può decidere di abortire una transazione (dopo aver inviato alcune richieste di operazioni di scrittura) inviando al server una richiesta di tipo abort. Il server deve garantire che una transazione abortita non abbia alcun effetto sul file system.

- Il file system distribuito deve essere tollerante ai seguenti tipi di failure.
Per il client: omissioni, failstop, bizantini.
Per il server: failstop.
- Nel caso di failure di tipo omissione sul client, può accadere che il client perda alcuni messaggi inviati nella transazione. Il server deve implementare un meccanismo per fronteggiare questo tipo di failure e garantire che non venga effettuato il commit della transazione finché questo tipo di failure non viene risolto.
- Nel caso di failure di tipo failstop sul client, può accadere che il client subisca un crash prima del commit della transazione. Il server deve garantire che non venga effettuato il commit della transazione. Il server può prevedere un meccanismo basato su timeout per attendere il ripristino del client o per cancellare traccia della transazione non completata.
- Nel caso di failure di tipo bizantino sul client, può accadere che il client presenti un comportamento anomalo, non corrispondente alle regole di funzionamento del file system. Esempi di comportamenti anomali includono l'invio di un messaggio con ID inesistente della transazione, oppure l'invio di messaggi non formattati secondo il protocollo applicativo definito. Il server deve garantire che non venga effettuato il commit della transazione.
- Il server deve essere stateful, ovvero deve sapere quali operazioni stava eseguendo prima del suo crash.
- Il server deve essere in grado di gestire transazioni concorrenti riguardanti lo stesso file. Nel caso in cui due client effettuino simultaneamente il commit della transazione sullo stesso file, è compito del server stabilire quale transazione viene completata per prima.
- La replicazione deve basarsi sullo schema primary copy, in modo tale che il client comunichi con un solo server per volta. In caso di crash del server primario, i file relativi a transazioni completate dovranno essere presenti sul nuovo server primario.

Il client ed i server del file system distribuito devono essere eseguibili nello spazio utente e senza richiedere privilegi di root. Si richiede inoltre la presenza di un file di configurazione, tramite il quale sia possibile specificare i valori dei parametri di configurazione (tra cui: l'indirizzo IP e la porta del server primario, il path della directory in cui il server primario memorizzerà i file).

Si richiede di effettuare un testing completo del file system realizzato e di presentarne i risultati nella relazione. Per il testing, usare il client sviluppato e considerare sia il comportamento esente da failure (anche in presenza di concorrenza), sia tutti i casi di failure sopra specificati.

Scelta e consegna del progetto

Il progetto può essere realizzato da un gruppo composto al massimo da **tre** studenti ed è riservato agli studenti che hanno optato per la modalità B dell'esame di Sistemi Distribuiti.

Per poter sostenere l'esame nell'A.A. 2009/10, **entro lunedì 25/1/2010** è necessario prenotarsi per il progetto, comunicando a Valeria Cardellini le seguenti informazioni:

- nominativi ed indirizzi di e-mail dei componenti del gruppo;
- progetto scelto.

Per ogni comunicazione via e-mail è necessario specificare [SD] nel subject della mail. Il progetto è valido **solo** per l'A.A. 2009/10: esso scadrà inderogabilmente con la sessione d'esame di settembre 2010.

La consegna del progetto deve avvenire **almeno dieci giorni** prima della data stabilita per l'appello di Sistemi Distribuiti in cui si intende sostenere la discussione del progetto e la prova orale (se non già superata). Il materiale relativo al progetto deve essere consegnato a Valeria Cardellini.

La consegna del progetto consiste in:

1. un CD-ROM contenente tutti i sorgenti (opportunamente commentati) necessari per il funzionamento e la copia elettronica della relazione (in formato pdf);
2. la copia cartacea della relazione.

La relazione contiene:

- la descrizione dettagliata dell'architettura del sistema e delle scelte progettuali effettuate;
- la descrizione dell'implementazione;
- la descrizione delle eventuali limitazioni riscontrate;
- l'indicazione della piattaforma software usata per lo sviluppo ed il testing del sistema;
- alcuni esempi di funzionamento;
- un breve manuale per l'installazione, la configurazione e l'esecuzione del sistema.

Valutazione del progetto

I principali criteri di valutazione del progetto saranno:

1. rispondenza ai requisiti;
2. originalità;
3. efficienza;
4. leggibilità del codice;
5. modularità del codice;
6. organizzazione, chiarezza e completezza della relazione;
7. semplicità di installazione e configurazione del software realizzato in ambiente Linux.