

Ingegneria del Web - A.A. 2007/08

Progetto C: Sistema di prefetching client-side per traffico Web

Ing. Valeria Cardellini

Requisiti del progetto

Lo scopo del progetto è realizzare in linguaggio C usando l'API del socket di Berkeley un **sistema di prefetching client-side per traffico Web**. Il prefetching (o caching proattivo) si pone l'obiettivo di superare le limitazioni del tradizionale caching di tipo passivo tramite un precaricamento delle risorse Web effettuato in anticipo rispetto alle successive richieste degli utenti.

I requisiti salienti dell'applicazione sono elencati di seguito.

- Il supporto del protocollo HTTP/1.1 (in particolare, connessioni persistenti e chunked encoding) nel gestire l'interazione con i server Web;
- il supporto del prefetching di risorse Web sia statiche che dinamiche;
- il supporto di almeno un meccanismo di predizione delle risorse da precaricare. Esistono molti approcci proposti in letteratura (vedi riferimenti successivi), a partire da una semplice strategia basata su una soglia prefissata, in cui sulla base degli accessi passati si seleziona un insieme di risorse candidate e si effettua il prefetching di quelle risorse nell'insieme la cui probabilità di accesso supera la soglia prefissata, fino a strategie basate su modelli Markoviani o sul data mining;
- il supporto del prefetching di risorse specificate dall'utente, in base alle modalità da questi indicate (quando attivare il prefetching, ogni quanto, ...);
- la gestione dello spazio di memorizzazione su disco disponibile per il prefetching;
- il logging configurabile dell'attività di prefetching.

Per approfondire lo studio dei meccanismi di predizione per il prefetching client-side in ambito Web si possono considerare le soluzioni proposte in:

1. A. Balamash, M. Krunz, P. Nain, "Performance analysis of a client-side caching/prefetching system for Web traffic", *Computer Networks*, Vol. 51, No. 13, Sep. 2007. <http://www-sop.inria.fr/maestro/personnel/Philippe.Nain/PAPERS/PREFETCHING/Comnet.pdf>
2. Y. Jiang, M.-Y. Wu, W. Shu, "Web prefetching: costs, benefits and performance", *Proc. of 7th Int'l Workshop on Web Content Caching and Distribution*, Aug. 2002. <http://2002.iwcw.org/papers/18500186.pdf>
3. W.-G. Teng, C.-Y. Chang, and M.-S. Chen, "Integrating Web caching and Web prefetching in client-side proxies", *IEEE Trans. on Parallel and Distributed Systems*, Vol. 16, No. 5, pp. 444-455, May 2005.

oppure altri lavori sull'argomento (ad esempio quelli elencati nei riferimenti bibliografici in [1]), reperibili tramite <http://scholar.google.com/>.

Il sistema di prefetching deve essere eseguito nello spazio utente e senza richiedere privilegi di root. I suoi parametri di configurazione devono essere specificati tramite un apposito file di configurazione.

Il funzionamento del sistema di prefetching e la sua rispondenza ai requisiti possono essere testati tramite un browser oppure un tool a riga di comando quale cURL (disponibile all'indirizzo <http://curl.haxx.se/>).

Opzionale

Il progetto può prevedere anche la realizzazione di una semplice interfaccia grafica per la configurazione e l'uso del sistema di prefetching.

Consegna del progetto

Il progetto può essere realizzato da un gruppo composto al massimo da **tre** studenti.

Per poter sostenere l'esame nell'A.A. 2007/08, **entro mercoledì 12/3/2008** devono essere comunicate via e-mail 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 [IW] nel subject della mail. Eventuali modifiche relative al gruppo devono essere tempestivamente comunicate e concordate con il docente. Il progetto è valido **solo** per l'A.A. 2007/08: i progetti scadranno inderogabilmente con l'ultima sessione d'esame dell'A.A. 2007/08.

La consegna del progetto deve avvenire **almeno dieci giorni** prima della data stabilita per l'appello di Ingegneria del Web 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 dell'applicazione 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 dell'applicazione;
- un esempio di funzionamento;
- un manuale per l'installazione, la configurazione e l'esecuzione dell'applicazione.

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.