

Cache del disco

Index

- [Introduction](#)
 - [Usato meno di recente \(LRU\)](#)
 - [Usato meno di frequente \(LFU\)](#)
 - [Sostituzione basata su frequenza](#)
 - [Sostituzione basata su frequenza: 3 segmenti](#)
 - [Prestazioni della cache del disco](#)
-

Introduction

Anche questo un altro buffer in memoria principale usato esclusivamente per i settori del disco e contiene una copia di alcuni di essi.

Quando si fa una richiesta di I/O per dati che si trovano su un certo settore, si vede prima se tale settore è presente nella cache, se non c'è, il settore letto viene anche copiato nella cache. Ci sono vari modi per gestire questa cache

Spesso questa viene chiamata *page cache*, da non confondere con la cache spesso presente direttamente sui dischi (quest'ultima è hardware e quindi trasparente per il sistema operativo)

Usato meno di recente (LRU)

Se occorre rimpiazzare qualche settore nella cache piena, si prende quello nella cache da più tempo senza referenze.

La cache viene puntata da uno "stack" di puntatori e quello riferito più di frequente è in cima allo stack. Quindi ogni volta che un settore viene referenziato o copiato nella cache, il suo puntatore viene portato in cima allo stack

Non è un vero stack, perché non è acceduto usando solo push, pop e top

Usato meno di frequente (LFU)

Viene rimpiazzato il settore con meno referenze (il settore usato meno di recente non necessariamente è quello usato meno di frequente)

Ovviamente serve un contatore (inizialmente ad 1) per ogni settore che viene incrementato ad ogni riferimento, e si rimuove il settore con il contatore minimo

Risulta quindi essere più sensato rispetto all'LRU: meno vieni usato, meno servi
Ma la località potrebbe avere un effetto dannoso in questo caso. Prendiamo come esempio un settore acceduto varie volte di fila perché contiene dati acceduti secondo il principio di località. Dopo che si è finito di accedere a questi dati non serve più, ma non verrà sostituito in quanto ha un valore abbastanza alto

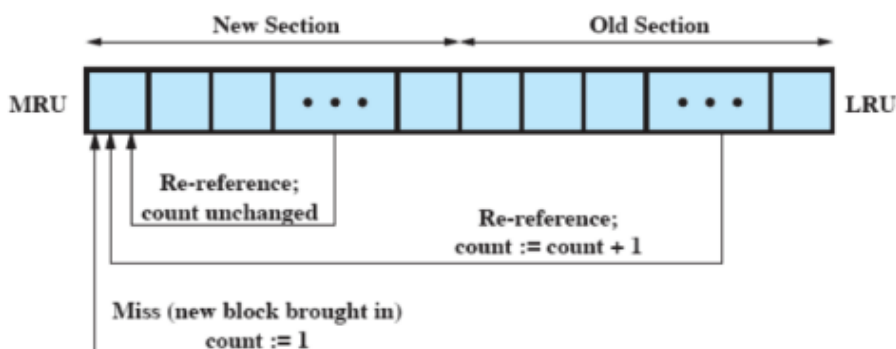
Sostituzione basata su frequenza

Si è quindi pensato ad una soluzione intermedia tra LRU e LFU.

Come nell'LRU c'è uno "stack" di puntatori (quando un blocco viene referenziato, lo si sposta all'inizio dello stack), ma è spezzato in due: una parte nuove e una vecchia.

Quindi ogni volta che ci sta un riferimento ad un settore nella cache, l'incremento avviene solo se si trova nella parte vecchia e per la sostituzione si sceglie il blocco con contatore minimo nella parte vecchia (in caso di parità il più recente)

Si passa dalla parte nuova alla parte vecchia per scorrimento: quando un blocco vecchio viene riferito e diventa nuovo, spinge l'ultimo dei nuovi a diventare il primo dei vecchi



Anche questa soluzione però ha dei problemi; infatti se non c'è presto un riferimento ad un blocco, questo potrebbe essere sostituito solo perché è appena arrivato nella parte vecchia anche se serviva e i suoi riferimenti sarebbero arrivati di lì a poco

Sostituzione basata su frequenza: 3 segmenti

Nuovo

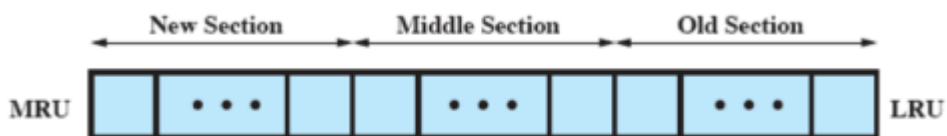
- unica parte dove i contatori non vengono incrementati
- non eleggibile per rimpiazzamento

Medio

- i contatori vengono incrementati
- non eleggibile per rimpiazzamento

Vecchio

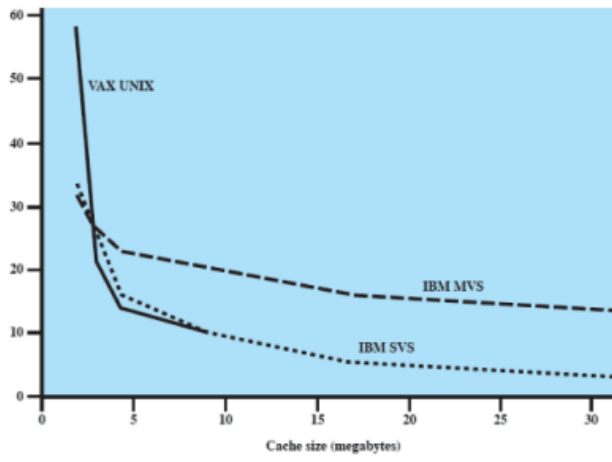
- i contatori vengono incrementati
- i blocchi sono eleggibili per il rimpiazzamento



In questa soluzione anche se non sono riferito da tempo ancora non posso essere sostituito in quanto non ancora arrivato nella sezione vecchia

Prestazioni della cache del disco

LRU



LFU

