

Sprint: Sistemi affidabili e ad elevata potenza di calcolo

La domanda di prestazioni e affidabilità elevate, assieme al crollo dei prezzi dell'hardware, hanno portato alla diffusione su larga scala di estesi gruppi di computer. Queste condizioni offrono un grosso potenziale per i sistemi di gestione di dati altamente efficienti ed affidabili. Per realizzare ciò, tuttavia, è necessario rivedere diverse ipotesi fondamentali alla base dei sistemi attuali di gestione dei dati. Questo è l'obiettivo di Sprint, un progetto della Facoltà di Scienze Informatiche dell'Università della Svizzera Italiana, finanziato dalla Fondazione Hasler e dal Fondo Nazionale Svizzero.

In questi ultimi anni, le applicazioni web sono diventate un elemento comune nella maggior parte dei servizi online. Questi sistemi richiedono dei tempi di risposta rapidi e la capacità di servire un gran numero di clienti allo stesso tempo e in modo continuo. Spesso le applicazioni web si basano su un sistema di banche dati, il cui compito è quello di immagazzinare il comportamento o lo stato dell'applicazione nell'arco del tempo. L'aumento del numero di clienti nel sistema crea immancabilmente una strozzatura nella banca dati. Questo succede poiché è più semplice duplicare gli altri componenti di un'applicazione web, ad esempio il web server, piuttosto che la banca dati stessa. Inoltre, quando vengono duplicati altri elementi, la banca dati diventa il punto critico, poiché l'affidabilità dell'applicazione dipende dalla sua robustezza. Le soluzioni attuali, che aumentano la performance e l'affidabilità dei sistemi di banche dati, spesso si basano su hardware dedicato e sono care. Il progetto Sprint ha come obiettivo di disegnare e implementare una piattaforma di gestione di dati per le applicazioni web. La piattaforma sarà sviluppata in base alle più moderne tendenze nel campo dei componenti hardware già disponibili. Molti sistemi in commercio possiedono una capacità di memoria principale vicina a quella del disco rigido di alcuni anni fa. Inoltre, l'hardware può essere assemblato in gruppi di computer molto potenti, costruiti a partire da componenti già disponibili e gestiti da software gratuito. Tuttavia, un'architettura di questo tipo richiede nuovi approcci al modo con cui sono progettati i sistemi software. Basta considerare ad esempio il divario tra la capacità dei dischi rigidi e la velocità di trasferimento dei dati da/per il disco. Infatti, la capacità del disco raddoppia ogni anno, mentre i tempi di accesso aumentano del 10% ogni anno. Ciò crea una strozzatura che limita la capacità del sistema di accedere ai dati e di elaborarli, anche se i processori sono più veloci. Si potrebbe risolvere questo problema concentrandosi sui sistemi di gestione di dati che utilizzano esclusivamente la memoria principale (ad esempio le specifiche banche dati). Oltre ai requisiti di



Un tipico ambiente Sprint (copyright Virginia Tech).

prestazione e di affidabilità (ossia la percentuale di tempo con cui la macchina è capace di rispondere ad una richiesta), i sistemi di gestione di dati dovranno far fronte alla crescente complessità delle architetture dei gruppi di computer. Questi sistemi dovrebbero essere in grado di adattarsi automaticamente ai cambiamenti della configurazione del sistema. Oltre a semplificare la gestione dei gruppi di computer, l'adattabilità può anche contribuire a migliorare la sua affidabilità e la sua performance: se un server importante va in panne, automaticamente ne subentra un altro che riesce a recuperare il suo lavoro, aumentando così l'affidabilità dell'intero sistema. Adattarsi alle caratteristiche delle richieste che arrivano dall'esterno presenta un ulteriore vantaggio. Se la maggior parte delle applicazioni si limitano a leggere la banca dati, le prestazioni possono essere migliorate aumentando il numero di copie della banca dati e leggendo in parallelo sulle diverse copie, incrementando così l'affidabilità dell'intero sistema. Se invece prevalgono le operazioni di aggiornamento (ossia delle richieste che modificano i dati), allora la presenza di più copie significa una maggiore perdita di prestazione per sincronizzarle e tenerle aggiornate. Il grado ottimale di duplicazione può essere raggiunto solamente da sistemi in grado di adattarsi alle caratteristiche delle richieste in tempo reale. L'obiettivo di Sprint è quello di risolvere questo tipo di problema.

L'approccio di Sprint

Il progetto Sprint sviluppa delle architetture per gruppi di server e si avvale di tecniche ad alta prestazione per superare i limiti dei tempi di risposta delle banche dati tradizionali basate sulla scrittura su disco rigido (che sono solitamente lunghi). Per ottenere affidabilità e prestazioni elevate, si combinano due strategie: dividere i dati tra computer diversi e allo stesso tempo duplicarli. Su ogni server è attiva una banca dati basata sulla memoria principale. Ciò significa che l'intera banca dati si trova nella memoria principale dei server, ma nessun server preso individualmente contiene l'intera immagine della banca dati. Di conseguenza, le operazioni di sola lettura della banca dati dipendono unicamente dalla memoria principale e dai tempi di risposta della rete, che sono inferiori di alcuni ordini di grandezza ai tempi di risposta del disco rigido. Per motivi dovuti al ripristino dei dati, le transazioni che modificano lo stato della banca dati devono essere scritte su disco. Sprint tratta questi casi accedendo al disco in modo sequenziale, procedura molto più rapida dell'accesso casuale. L'eliminazione dei ritardi dovuti alla trascrizione dei dati su disco consente di migliorare sia il numero di transazioni eseguite per unità di tempo che il tempo di risposta. La duplicazione viene utilizzata per incrementare sia l'affidabilità che la performance. Senza la duplicazione, i dati depositati in un server che va in panne andrebbero persi. Dal momento che ogni singolo dato viene duplicato, gli utenti non notano alcuna interruzione del servizio quando un server va in panne, visto che i dati possono essere prelevati da un altro server. La duplicazione aiuta pure a migliorare la performance: se alcuni dati vengono per lo più letti e raramente modificati, allora la loro duplicazione in diversi server permette di

Gestione di dati a grande distanza

Vi è un'altra tendenza che riguarda i sistemi futuri di gestione dei dati: le reti di gruppi di computer autonomi diffusi su ampie aree geografiche e interconnessi tra di loro attraverso collegamenti a lunga distanza. Queste architetture sono in grado di sopravvivere agli eventi catastrofici (come ad esempio un terremoto che potrebbe distruggere l'intero gruppo di computer). Una delle principali sfide è quella di ridurre gli elevati tempi di risposta di comunicazione causati dalla distanza tra i gruppi. I ricercatori dell'USI stanno analizzando questi aspetti nel contesto di GORDA, un progetto finanziato dall'Unione Europea. Il progetto GORDA collegherà diversi gruppi di banche dati sparsi ovunque in Europa. Il consorzio comprende partner accademici e industriali, come il produttore di banche dati MySQL.



I ricercatori dell'USI. Da sinistra a destra: Lasaro Camargos, Fernando Pedone, Vaide Zuikeviciute, Rim Moussa, Marcin Wieloch e Rodrigo Schmidt.

eseguire le transazioni in parallelo. La ripartizione dei dati viene utilizzata per aumentare la prestazione. L'idea è che se l'accesso alle singole parti della banca dati avviene frequentemente, allora queste ultime possono essere posizionate in diversi server, di modo che le operazioni possano essere svolte in parallelo. Il progetto Sprint intende contribuire alla ricerca sulla progettazione e l'implementazione di sistemi adattativi di gestione di dati altamente efficienti ed affidabili. Inoltre verranno rivisti gli algoritmi tradizionali di gestione dei dati dal punto di vista delle tecnologie emergenti dei gruppi di computer, cercando di capire in che modo queste tecnologie incidono sulla performance dei gruppi stessi. La richiesta di sistemi di banche dati potenti e robusti continuerà ad aumentare nei prossimi anni. Per raggiungere un livello di scalabilità e di affidabilità tali da soddisfare questa domanda, occorre disporre di tecniche che sfruttino le moderne tecnologie dei gruppi di computer e che si adattino automaticamente ai cambiamenti esterni. Sprint è un progetto open-source. Ciò significa che gli utenti potranno scaricare gratuitamente il software e utilizzarlo a piacere.

Informazioni:

Prof. Fernando Pedone
 Facoltà di Scienze informatiche
 Università della Svizzera Italiana
 Tel. +41 58 666 46 95; E-mail: fernando.pedone@unisi.ch
<http://www.inf.unisi.ch/pedone>

Indirizzi web:

<http://www.inf.unisi.ch>
<http://www.inf.unisi.ch/sprint>