



Tauler - M1.226 - Laboratori Sistemes Distribuïts Gran Escala a...

Rebut

Jordi Puigsegur Figueras

solució pràctica

A: Tauler Labora. Sistemes Distribuïts Gran Escala aula 1

Data: 17.05.2018 22:42

Arxius adjunts

 ["mapreduce-SOLUTION.js" \(downloadAttachment?pMailId=5602321_741320899&pAttachmentId=1&s=8107086a...](#)

Hola a tothom!

A continuació us comento per on havia d'anar la solució a la pràctica.

Les característiques importants de la eina per a complir els requeriments demanats serien:

- Tolerància a particions (partition tolerance)

En cas d'una partició de xarxa la web ha de seguir funcionant. És crític que no s'aturin ni les lectures ni les escriptures.

- Alta disponibilitat (high availability)

El sistema ha d'estar disponible 24x7, sense interrupcions. Bàsicament parlem d'alta disponibilitat de lectura.

- Alta capacitat lectura + baixa latència (read throughput + low latency)

L'objectiu principal ha de ser fer de motor a la web responent a un nombre molt elevat de peticions en un temps baix.

- Escalabilitat horitzontal (horizontal scalability)

La escalabilitat horitzontal és imprescindible ja inicialment per a poder aguantar el nombre de lectures i escriptures previstes i també per a poder assumir qualsevol creixement futur.

- Capacitat d'escriptura (write throughput)

L'eina ha de ser capaç de permetre les operacions d'escriptura per tal de poder enmagatzemar els logs corresponents a les visites.

- Capacitat d'indexació per diferents criteris

És important també poder disposar d'índexos secundaris per a poder realitzar cerques per altres camps. Per exemple un avió en concret.

També seria recomanable tenint en compte les característiques de l'aplicació:

- Esborrat d'informació per TTL

Es important que la eina suporti l'esborrat de les dades caducades, ja que sempre serà més eficient que si ho hem de fer nosaltres de forma programàtica.

- Distribució geogràfica (multi data-center awareness)

Pot ser interessant disposar de diferents centres de càlcul distribuïts geogràficament ja que tant la recollida de dades com la web són a nivell global. Ha de gestionar la topologia de la xarxa i ha de permetre seleccionar en quin centre de càlcul llegim i escrivim.

- Processament distribuït

L'eina escollida ha de suportar processament distribuït, tipus MapReduce, per a poder realitzar els càlculs que ens proposen, o d'altres que ens puguin aportar valor afegit.

D'entrada es descarten els SGBDR ja que són sistemes dissenyats per a ser altament consistens i com a mínim ens donaran problemes en la distribució geogràfica, la tolerància a particions de dades, i en la escalabilitat horitzontal. En definitiva, els SGBDR no están pensats per la tasca que ens han proposat, que realitzarà millor una eina NoSQL. En general les eines noSQL utilitzen menys recursos i estan pensades específicament per aquests tipus de problema.

La clau per escollir la eina adequada per al nostre problema es tracta en identificar correctament que estem davant d'un problema clarament AP, on la disponibilitat i la tolerància a particions és crítica i en canvi la consistència no ho és ja que no es modifiquen dades, només anem afegint nous registres i per tant no hi haurà conflictes.

Respecte a quina es la millor eina, no hi ha una única solució. El fet que necessitem disponibilitat d'escriptura dóna més punts a eines com Cassandra o CouchDB que son multi-master. La distribució geogràfica dóna més punts a Cassandra. En tot cas tenint en compte el volum de dades que hem proposat tampoc podem descartar les altres eines. En tots els casos hi ha tradeoffs i en la meua opinió l'èxit d'un projecte d'aquest tipus i envergadura també dependrà molt d'altres factors, com per exemple l'arquitectura de la web en si. Així mateix, no he tingut en compte altres paràmetres com suport i popularitat de la eina, que també es tindrien en compte al prendre una decisió per tal de garantir que continuarà en desenvolupament i suportada.

En l'avaluació no s'ha tingut en compte tant l'eina escollida sino l'argumentació que es feia, valorant especialment que fos correcta i breu.

S'adjunta l'arxiu javascript corresponent a la solució del map reduce.

Respecte l'exercici de Cassandra, es tracta igualment d'identificar que el problema és un AP. Al tractar-se sempre d'insercions en una colecció que ve a ser una sèrie temporal no es poden donar inconsistències. Com a molt hi haurà nodes amb les dades no del tot actualitzades. Per això, tant en el cas de les escriptures com les lectures es tracta d'afavorir la disponibilitat amb un nivell ONE. És important tenir clar que això no afecta al número de rèpliques d'una dada a cada centre de càlcul, que es regeix pel replication factor.

Jordi Puigsegur
jpuigsegur@uoc.edu