

BAZE DE DATE NOSQL

PANIOGLO Vadim

Universitatea Tehnică a Moldovei

Abstract: Bazele de date relaționale sunt tot mai des substituite din cauza lipsei de flexibilitate. Aceste tipuri de baze de date utilizează operații foarte complexe. În trecut bazele de date relaționale gen MySQL erau pe primul loc, însă în ultimul timp sunt mai preferate cele NoSQL. În sfera unde se utilizează MySQL s-au infiltrat și cei ce oferă soluții de baze de date cloud. Bazele de date cloud poartă numele de NoSQL- Not only SQL și sunt baze de date non relaționale. În articolul dat am realizat o prezentare a generației noi de baze de date NoSQL, care sunt considerate un pas important în evoluția bazelor de date, descriind avantajele și dezavantajele acestora față de modelul relațional tradițional. Deasemenea, articolul în cauză conține o analiză a funcționalităților bazei de date NoSQL MongoDB, pentru a arăta cât este de ușor de a trece de la SQL la NoSQL.

Cuvinte cheie: Baze De Date, NoSql, Sql, Model Relațional, Open-Source, MongoDB.

1. Introducere

Fiecare întreprindere adună cu timpul cantități mari de date care compuse din informații necesare pentru desfășurarea activității de business. În mare majoritate a cazurilor, organizațiile stochează aceste date în baze de date relaționale pentru a putea avea acces la ele și pentru a le utiliza în scopurile inițial predefinite. Cu toate acestea, în ultimul timp mulți dintre dezvoltatori au început să utilizeze baze de date non relaționale numite NoSQL-Not Only SQL. Diferitele baze de date NoSQL existente azi în domeniul dat exemplifică mai multe abordări. Ceea ce au în comun este că nu sunt relaționale. Principalul avantaj este că permit lucrul eficient cu date nestructurate precum e-mail, multimedia, procesoare de text. La momentul actual există multe companii care au creat propriile baze de date NoSQL. Cele mai răspândite sunt cele dezvoltate de către companiile mari Web 2.0, precum Amazon și Google, din necesitatea de a procesa cantități mari de date. Acestea au dezvoltat Dynamo și Big Table ce stau la fundamentul altor baze de date NoSQL disponibile acum.

2. Caracteristici generale

Bazele de date de tip NoSQL sunt caracterizate prin memorarea unor volume mari de date (organizațiile ca Google și Amazon au în gestiune de la 10 până la 100 mii de servere) fără a avea restricția de a stoca doar date cu structură fixă. În aceste baze de date pot fi stabilite legături (prin referințe la date memorate în alte baze de date) între date și pot fi memorate aceleași date pe mai multe servere (partajare și replicare). De asemenea bazele de date NoSQL nu utilizează operații de joncțiune la interogări (consumare foarte mare de timp). Aceste baze de date sunt soluții foarte bune dar pentru cazuri particulare, se necesită o analiză specializată pentru a face o alegere corectă.

Un exemplu bun de acest tip de bază de date ar fi MongoDB, care stochează datele folosind tabele nu ca o bază de date tradițională, ci sub o formă de documente JSON, cu scheme dinamice. Nu există schemă pentru câmpurile dintr-un document, acestea precum și tipurile lor pot varia. Astfel nu există operația de „alter table” pentru adăugare de coloane. În practică este obișnuit ca o colecție să aibă o structură omogenă, deși nu este o cerință, colecțiile putând avea structuri diferite. Această flexibilitate presupune ușurință în migrarea și modificarea imaginii de ansamblu asupra datelor.

Aceste baze de date de asemenea utilizează niște limbaje de interogare, cum ar fi UnSQL (Unstructured Query Language), care lucrează cu fișiere de tip JSON și Xquery – cu XML. Un exemplu real de manipulare a datelor ar fi:

```
INSERT INTO abc VALUE { type: "nested", content: { content: "nested object", x: 1, y: {str: "hi", str2: "there"}, z: true } };
```

```
SELECT { x:abc.type, y:abc.content.x, z:abc.content.x+50 } FROM abc;
```

3. Avantaje și dezavantaje

Care sunt situațiile în care se recomandă folosirea unei soluții NoSQL?

- baza de date tradițională nu mai poate fi scalată la un preț acceptabil;
- se generează foarte multe informații temporare cum ar fi: coșuri de cumpărături, chestionare incomplete, istorice de navigare, personalizări, etc;
- baza de date a fost deja denormalizată pentru a îmbunătăți performanțele;
- stocați cantități foarte mari de text și/sau imagini;
- aveți nevoie să rulați interogări de date care nu implică doar simple relații ierarhice. de exemplu: “toți oamenii dintr-o rețea socială care nu au cumpărat anul acesta o carte dar au legătură cu o persoană care a cumpărat”;

- tranzacțiile nu trebuie să fie perfect consistente; de exemplu un buton de “like”, dacă tranzacția eșuează, nu este o problemă, utilizatorul cel mai probabil va mai apăsa o dată butonul.

Bineînțeles, nu este totul perfect și există și dezavantaje:

- nu există un limbaj universal valabil. Bazele de date relaționale au SQL, care chiar dacă are multe extensii proprietare totuși utilizatorii știu la ce să se aștepte;

- maturitatea – majoritatea sistemelor NoSQL încă sunt la primele variante sau încă în plină dezvoltare;

- suport – fiind în general proiecte open source, iar firmele ce oferă suport sunt mici, de multe ori startup-uri și poate nu oferă suficientă credibilitate;

- disponibilitatea dezvoltatorilor – evident fiind o tehnologie nouă, comparativ cu bazele de date tradiționale sunt mult mai puțini dezvoltatori software NoSQL.

- nu se asigură consistența bazei de date (de către sistemul de gestiune)

- nu există metode performante pentru protecția datelor

- există posibilități limitate de interogare

4. Concluzii

Conform definiției formulate în articolul dat bazele de date NoSQL reprezintă o tranziție spre baze de date superioare ce vor îngloba flexibilitatea și performanțele lor actuale cu modelul relațional. Odată cu apariția bazelor de date NoSQL, dezvoltatorii au avantajul de a utiliza mai multă diversitate în modelul de date abordat. La fel aceste baze de date formează modelul optim pentru aplicațiile web. Din această cauză cunoașterea caracteristicilor lor este foarte importantă, mai ales înainte de a trece la o astfel de soluție. Baza de date MongoDB, nominalizată în acest articol, este ușor de înțeles și de manipulat. Este ideală atât pentru proiecte mici, de test, cât și pentru proiecte ce vor stoca un volum mare de date. Posibil, MongoDB este o bază de date ce va fi utilizată și mai mult pe viitor datorită tendinței prezente spre aplicații Web.

Bibliografie

1. Wikipedia: NoSQL, [Resursă electronică]. – Regim de acces: <http://en.wikipedia.org/wiki/NoSQL>
2. Wikipedia: MongoDB, [Resursă electronică]. – Regim de acces: <https://ro.wikipedia.org/wiki/MongoDB>
3. NoSQL: Avantaje și Dezavantaje, [Resursă electronică]. – Regim de acces: <http://www.books-express.ro/blog/nosql-avantaje-si-dezavantaje/>
3. Guide to the Non - Relational Universe, [Resursă electronică]. – Regim de acces: <http://nosql-database.org/>
4. Documentația oficială a MongoDB, [Resursă electronică]. – Regim de acces: <http://docs.mongodb.org/>
5. STONEBRAKER, M.: The "NoSQL" Discussion has Nothing to Do With SQL. November 2009. - Blog post of 2009-11-04. [Resursă electronică]. – Regim de acces: <http://cacm.acm.org/blogs/blog-cacm/50678-the-nosql-discussion-has-nothing-to-do-withsql/fulltext>
6. SHALOM, N.: The Common Principles Behind The NoSQL Alternatives, December 2009,- Blog post of 2009-12-15. [Resursă electronică]. – Regim de acces: http://natishalom.typepad.com/nati_shaloms_blog/2009/12/the-common-principlesbehind-the-nosql-alternatives.html