

EXPLORAREA ȘI BENEFICIILE BAZELOR DE DATE NOSQL

Nicolay-Alex AOUN

Departamentul Ingineria Software și Automatică, grupa TI-216, Calculatoare, Informatică și Microelectronică,
Universitatea Tehnică a Moldovei, Chișinău, Republica Moldova

Autorul corespondent: Aoun Nicolay-Alex, nicolay-alex.aoun@isa.utm.md

Îndrumător/coordonator științific: **Dorian SARANCIUC**, lector universitar, Departamentul Ingineria
Software și Automatică

Rezumat. Bazele de date NoSQL au devenit un aspect semnificativ în dezvoltarea administrării datelor cu timpul, constant evoluându-și extensibilitatea, flexibilitate și performanța. În acest articol este arătat beneficiul acestora, cum se utilizează și cum pot fi utilizate tehnologiile NoSQL pentru rezolvarea problemelor comune din zilele de astăzi a bazelor de date obișnuite. Ce accentuiază de ce este important să cunoaștem și să utilizăm bazele de date NoSQL și metodele prin care acestea pot fi utilizate în majoritatea industriilor pentru ușurarea vieții lucrătorilor.

Cuvinte cheie: stocare cloud, date, SQL, noSQL, optimizare, big data, arhitectura datelor

Introducere

În ultimii ani, tehnologia bazelor de date a trecut prin o multitudine de schimbări pentru a putea continua să răspundă la cerințele utilizatorilor, fie administrarea datelor sau dezvoltarea aplicațiilor. Bazele de date SQL au fost primele programe care utilizează modele structurate pentru a stoca și a recupera date. Aplicații precum MySQL, PostgreSQL și Microsoft SQL Server servesc ca fundația de baza a SQL, și sunt eficiente în menținerea datelor în siguranță cu proprietățile ACID (atomicity, consistency, isolation, durability).

Cu toate acestea, volumul imens de date și numărul semnificativ de informație nestructurată sau semi-structurată au făcut ca bazele de date tradiționale SQL să devină mult mai dificil de utilizat și menținut. Aceste schimbări au adus la crearea și apariția noilor baze de date NoSQL, care sunt mult mai flexibile și extensibile, și sunt capabile de a administra cu ușurință diferite tipuri de date.

MongoDB, Cassandra și Redis sunt unele cele mai folosite dintre aceste noi baze de date NoSQL care au apărut pentru a îmbunătăți viteza, flexibilitatea, și mai multe alte aspecte importante pentru aplicațiile moderne care lucrează cu cantitate enormă de date.

Structura și Tipurile de Baze de Date NoSQL

Există mai multe tipuri de baze de date NoSQL, fiecare dintre ele create pentru a se specializa pe tipuri specifice de date. În acest capitol vom vorbi despre aceste tipuri de date principale, acestea fiind Document Store, Key-Value, date graf și baze de date de tip Column-Family.

Bazele de date Document Store, de exemplu precum MongoDB, sunt folosite pentru administrarea datelor în formate de document, cum ar fi JSON sau XML. Aceste baze de date lucrează bine cu șiruri de informații complexe și bine structurate. Ele sunt folosite deseori ca baze de date pentru programe precum sistemele de administrare a informației și site-urile de comerț online, unde flexibilitatea și viteza sunt importante.

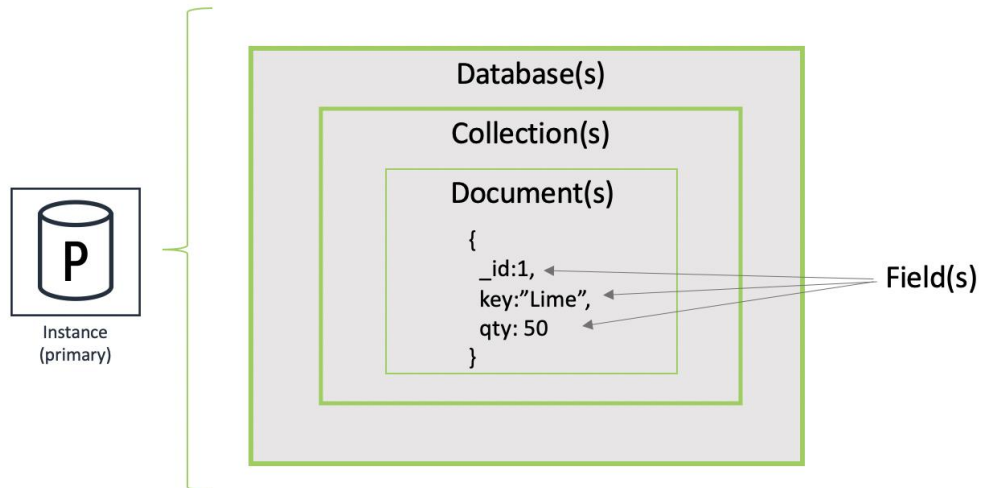


Figura 1. Structura unui document JSON într-un NoSQL Document Store

Redis este o Baza de Date de Tip Key-Value. NoSQLurile Key-Value sunt ușor de utilizat și salvează datele sub formă de perechi Key-Value. Acest design este creat cu scopul de a fi foarte rapid în găsirea informațiilor, ceea ce face acest tip de baze de date excelent pentru situațiile în care trebuie de găsit informații instantaneu, care sunt folosite de aplicațiile web ce stochează informații în cache.

Bazele de Date Graf, cum ar fi Neo4j, ne aranjează datele în forma de graf, cu noduri și legături, și se concentrează pe relațiile dintre entitățile de date. Acestea sunt utilizate mai ales în aplicații cum ar fi rețelele sociale sau search engines ca google și bing, unde este important de vizualizat cum sunt conectate datele.

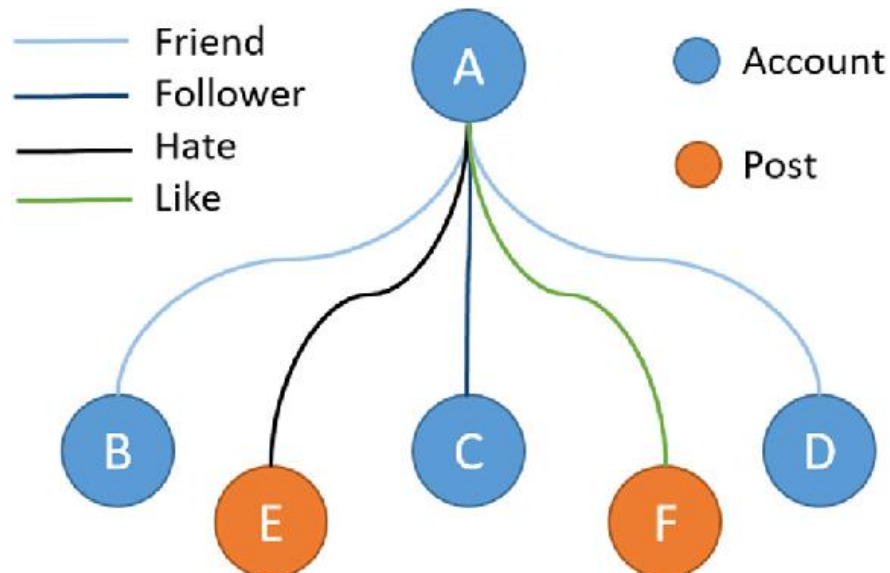


Figura 2. Ilustrarea unei baze de date graf

Bazele de Date de Tip Column-Family, ca de exemplu Apache Cassandra, sortează datele în coloane în loc de rânduri. Această structură lucrează bine pentru gestionarea seturilor mari de date și este foarte folositor în aplicații care necesită operațiuni frecvente de citire și scriere, cum ar fi algoritmi serii de timp și aplicațiile IoT.

Aceste diferite tipuri de baze de date NoSQL ne oferă o experiență îmbunătățită de management modern al datelor prin oferirea multor modalități de stocare și procesare a datelor.

Cu trecerea timpului, bazele de date NoSQL devin din ce în ce mai populare, deoarece sunt legate de big data, adică seturi de date enorme. Industriile, precum cele ale comerțului electronic, social media, și IoT (Internet of Things) au crescut semnificativ în volum și varietate, și va continua a crește cu trecerea timpului. Această cerere crescută a făcut trecerea la modalități mai puternice și flexibile de administrare a informațiilor mult mai necesară. Bazele de date NoSQL au fost create cu scopul de a rezolva aceste probleme, fiind flexibile și având posibilitatea de a se extinde la dorința utilizatorului. Acest tip de sistem informatic este suficient de puternic pentru a administra un număr mare de date neorganizate, și poate ține pasul cu multitudinea de schimbări în programele cu schimbare de date intensivă. Această schimbare reprezintă un mare avans pentru tehnologia bazelor de date, și ne arată cât de important și folositoare sunt bazele de date NoSQL.

Avantaje și Limite ale Bazelor de Date NoSQL

Bazele de date NoSQL au numeroase avantaje, în special în ceea ce privește big data, cloud computing și analiza datelor în timp real.

Extensibilitatea

Bazele de date NoSQL sunt foarte ușor de extins, în special dacă utilizatorul dorește să utilizeze mai multe servere pentru a completa mai multă muncă. Extensibilitatea este importantă pentru aplicațiile big data, în care cantitatea de informație și numărul de utilizatori crește exponențial în fiecare zi. În comparație cu bazele de date tradiționale care cresc cu greu într-o singură direcție, bazele de date NoSQL se pot extinde foarte ușor pentru a administra mai multe date în același timp, ceea ce ajută aplicațiile cu seturi mari de date.

Modelarea flexibilă a datelor

Bazele de date NoSQL au multă flexibilitate în administrarea diferitor date datorită abilității sale unice numită "schema-less model" (model de date fara schema). Această caracteristică este deosebit de folositoare pentru administrarea datelor ce nu sunt structurate, care sunt comune în aplicațiile moderne. Flexibilitatea acestor baze de date se extinde și la modificarea structurilor de date și adaptarea cerințelor aplicației fără a avea nevoie de un mare redesign.

Optimizarea performanței

Bazele de date NoSQL sunt proiectate în așa fel ca să fie foarte rapide, mai ales în mediile care necesită procesarea rapidă a datelor. Ele sunt eficiente în operațiile de citire și scriere la viteza mare, ceea ce le face o alegere excelentă pentru analizele în real-time și pentru aplicațiile care trebuie să proceseze și să retragă date rapid. Acest avantaj în performanțe este posibil datorită optimizării salvării și retragerii datelor, precum și abilității de a distribui operațiuni pe mai multe noduri.

Gestionarea datelor diverse și complexe

Bazele de date NoSQL pot salva și procesa diverse tipuri de date, de la perechi key-value simple la grafuri complexe. Această varietate le permite să execute mai multe lucruri diferite, cum ar fi salvarea și analiza datelor, acesta fiind important în lumea de astăzi unde totul este conectat online.

Limitele Bazelor de Date NoSQL

Pe când bazele de date NoSQL au avantaje mari, acestea au și câteva seturi de dezavantaje și limitări. Multe baze de date NoSQL optează pentru consistență diferită de modelul ACID. Aceasta abordare poate duce la niște inconsistențe temporare ale datelor, mai ales în

medii de distribuire ale acestora. Acest lucru poate fi o problemă mare în situații în care este importantă precizia datelor în timp real.

NoSQL des nu are aceleași puteri de interogare ca bazele de date SQL tradiționale. Această problemă poate duce la niște dificultăți mari când utilizatorul vrea să combine și să afișeze diferite informații, mai ales atunci când trebuie de căutat multe tipuri de informații diferite.

Bazele de date NoSQL nu sunt tot așa de securizate și eficiente ca bazele de date tradiționale. Asta poate dificulta dezvoltarea aplicațiilor care au ca scop păstrarea datelor în siguranță. Pentru a folosi soluțiile NoSQL, programatorii au nevoie de niște abilități și cunoștințe noi.

Deoarece există atât de multe tipuri diferite baze de date NoSQL, fiecare având caracteristici și proporții a lor unice, procesul de învățare este mai dificil pentru programatori și administratori.

Concluzii

NoSQL este folositor pentru mari cantități de informații și mai sunt potrivite pentru soluții cloud computing, deoarece pot administra multe diferite tipuri de date, funcționând bine în proces. Chiar și cu toate acestea, utilizarea bazelor de date NoSQL poate aduce și unele probleme, ca de exemplu menținerea consistenței datelor, combinarea datelor în moduri diferite și diferite aspecte legate de securitate acestora. Este important de echilibrat aceste probleme dacă sunt folosite tehnologiile NoSQL. Studiind bazele de date NoSQL arată cât de important este să înțelegem atât puterea acestora, cât și limitele abilităților acestora, și să ne asigurăm că le folosim în mod strategic. Pe măsură cu cât se mărește dimensiunea și complexitatea datelor, bazele de date NoSQL rămân un instrument important pentru inginerii de date.

Referințe:

- [1] „The Evolution of Data Engineering: From Traditional Databases to NoSQL and Beyond” [Online]. Available: <https://iabac.org/blog/the-evolution-of-data-engineering-from-traditional-databases-to-nosql-and-beyond>
- [2] „The Evolution of Databases from SQL to NoSQL” [Online]. Available: <https://bangdb.com/blog/the-evolution-of-databases-from-sql-to-nosql/>
- [3] „SQL vs NoSQL Databases: Key Differences and Practical Insights” [Online]. Available: <https://www.datacamp.com/blog/sql-vs-nosql-databases>
- [4] „NoSQL Databases: Advantages and Disadvantages” [Online]. Available: <https://www.dataversity.net/nosql-databases-advantages-and-disadvantages/>
- [5] „NoSQL Databases: A Key Component in Modern Data Engineering” [Online]. Available: <https://iabac.org/blog/nosql-databases-a-key-component-in-modern-data-engineering>