

Temă 2 DATC

Document no-sql database. Azure Cosmos DB.

Din ce în ce mai multe companii întâmpină dificultăți în stocarea volumului mare de date pe serverele relaționale MySQL, a căror structură se schimbă destul de des, fapt care duce la perioada lungi de nefuncționare a bazei de date și la costuri mari. O soluție eficientă a problemei este ***no-sql database***, care nu stochează datele în tabele relaționate, ci sub formă de pachete de informații, care nu trebuie să respecte o anumită schemă. Fiecare pachet de informații se deosebește de un alt pachet de informații printr-un id unic.

Prin ***document no-sql database (non-relational database)*** se înțelege că datele sunt stocate sub formă de “documente”, care de fapt sunt obiecte JSON. Fiecare rând al unei tabele va fi un obiect JSON, iar o tabelă va fi o colecție de obiecte JSON. Dacă un câmp este NULL într-o tabelă, atunci în obiectul JSON acest câmp nu va apărea, fiind considerat opțional.

De exemplu:

Structura tabelii relaționate sql:

| FirstName | LastName | Gender | Address Id (FK) | |
|-----------|-----------------|--------|-----------------|---------|
| Minal | Wad | F | Null | |
| Sid | Atreya | M | 1 | |
| Roma | Kole | Null | 2 | |
| AddressId | AddressLine1 | State | Country | Zipcode |
| 1 | 123 Land Rd | NJ | USA | 33345 |
| 2 | 456 Sunset Blvd | VA | USA | 22278 |

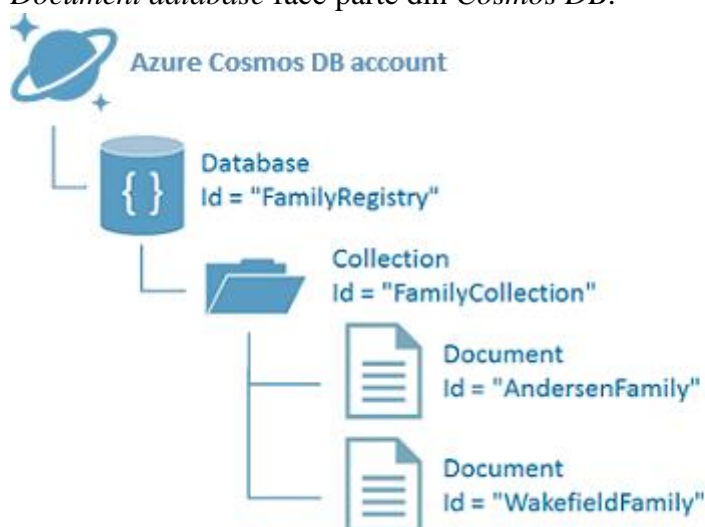
Echivalentul primei înregistrări din tablele relaționate în Document no-sql database, format JSON:

```
{
  "FirstName": "Minal",
  "LastName": "Wad",
  "Gender": "F",
  "Address": {
    "AddressLine1" : "123 Land Rd",
    "State": "MD",
    "Country": "USA",
    "Zipcode": "21703"
  }
}
```

| Document no-sql Database | Relational Database |
|------------------------------------|---|
| Datele nu sunt normalizate | Date normalizate |
| Integritate referențială neforțată | Integritatea referențială forțată prin normalizare și relaționare |
| Datele din colecție sunt mixte | Datele din table sunt la fel |
| Schemă flexibilă | Schema este setată prin schimbare |
| Limbaj SQL, precum și JavaScript | Limbaj SQL |

Printre primii utilizatorii a tehnologiei no-sql database au fost Google, Amazon și Facebook, înlocuind implementările bazelor de date relaționale pentru a obține flexibilitate și scalabilitate.

Document database face parte din Cosmos DB.



Azure Cosmos DB este un serviciu de date multi-dimensional, global distribuit, care permite o capacitate mare de stocare a datelor pe regiuni, utilizând latență redusă și disponibilitate ridicată.

Avantajele folosirii **Azure Cosmos DB** sunt:

- ✓ Distribuție globală: datele sunt distribuite în regiunile dorite cu ajutorul API-urilor, care permit ca datele să fie stocate unde sunt utilizatorii, asigurând o latență scăzută.
- ✓ Modele multiple de date și API-uri populare pentru accesarea și interogarea datelor: SQL API, MongoDB API, Cassandra API, Gremlin API, Table API.
- ✓ Elasticitate și autonomie la scară de producție și stocare la cerere și la nivel mondial prin mărirea spațiului de stocare și extinderea cu ușurință a capacității bazei de date.
- ✓ Dezvoltarea aplicațiilor extrem de receptive.
- ✓ Asigurarea disponibilității.
- ✓ Dezvoltarea aplicației, distribuită global în mod corect.
- ✓ Nu ține cont de o schema a bazei de date sau de gestionarea indexului.
- ✓ Costul redus.

Azure Cosmos DB garantează toate aceste servicii ca o soluție pentru orice aplicație web, mobilă, pentru jocuri sau IoT, care lucrează cu o masă mare de date, citește și scrie date la nivel global, necesitând un timp de răspuns aproape real.

Cele mai comune cazuri în care este folosit Azure Cosmos DB:

- ✓ *industria comerțului*, deoarece Azure Cosmos DB suportă schemele flexibile și datele ierarhice, fiind potrivit pentru stocarea produselor a căror atribute se pot modifica în timp. Un exemplu practic de utilizare: un catalogul de produse auto cu componente care au atât atribute comune cu celelalte componente cât și atribute specifice, care se pot modifica în timp.
- ✓ *industria jocurilor*, deoarece acestea necesită latențe de o milisecundă pentru a citi și scrie date. Azure Cosmos DB oferă o performanță de elasticitate mărită pentru actualizarea profilului și a statisticilor la zeci de milioane de jucători simultan prin efectuarea unui singur apel API. Indexarea automată permite localizarea jucătorilor prin ID-urile interne ale acestora, fără construirea unei infrastructurii complexe de indexare.
- ✓ *industria aplicațiilor web și mobile* pentru modelarea interacțiunilor sociale și integrarea cu serviciile terților. Conținutul chat-urilor, comentariilor sau a mesajelor pot fi stocate în Cosmos DB fără a necesita transformări. Aplicațiile care se integrează cu rețelele sociale terțe trebuie să răspundă schimbării schemelor din aceste rețele. Dat fiind că datele sunt indexate automat în mod implicit în Cosmos DB, datele sunt gata de a fi interogate în orice moment. Prin urmare, aceste aplicații au flexibilitatea de a recupera datele conform nevoilor lor.

Comparând din punct de vedere al capabilităților, Non-sql database împreună cu Azure Cosmos DB oferă servicii de *scalare orizontală*, *latență garantată* și *disponibilitate mărită*, servicii pe care bazele de date relaționate nu le oferă.