

Tendințe actuale în evoluția bazelor de date

Conf.dr. Manole VELICANU, prof.dr. Ion LUNGU,
 prof.dr. Gheorghe SABĂU, conf.dr. Constanța BODEA,
 Catedra de Informatică Economică, A.S.E., București

Datorită limitelor apărute la realizarea și utilizarea bazelor de date, au fost abordate noi direcții de cercetare. Acestea vizează dezvoltarea unei noi generații de baze de date. Sunt prezentate condițiile în care se face trecerea la generația a treia de baze de date și obiectivele acesteia.

Cuvinte cheie: baze de date, sisteme de gestiune a bazelor de date, aplicații informatice.

Sistemele de gestiune a bazelor de date (SGBD) au evoluat de la sisteme de generația întâi, care utilizau în principal modele de date ierarhice și rețea, spre sisteme de generația a doua, care utilizează abordarea relațională. Această evoluție a permis îmbunătățirea independenței fizice (complet) și logice (parțial) a datelor, oferirea de limbaje de interogare declarative, îmbunătățirea integrității și securității datelor, optimizarea accesului la date. În plus, abordarea relațională a adus o fundamentare teoretică a bazelor de date deschizând un domeniu de cercetare distinct.

Pe de altă parte, sistemele relaționale au anumite limite care fac dificilă utilizarea lor sau chiar imposibilă pentru anumite tipuri de aplicații. Printre slăbiciunile cele mai importante ale relaționalului, sunt: prea marea simplitate a conceptului de relație (care nu satisface totdeauna ca structură de date) și existența unor limbaje de manipulare a datelor (LMD) prea limitate, care necesită utilizarea comenzilor acestora într-un limbaj universal (exemplu Oracle). În ultimii ani au apărut SGBD relaționale, cum este Foxpro, care au un limbaj propriu LMD foarte puternic, apropiat de cele universale. Datorită limitelor apărute, o nouă generație de baze de date va urma celor relaționale. Această **a treia generație de baze de date** va evolua din relațional, păstrând ceea ce este mai important și adăugând concepte semantice mai elaborate decât relația, precum

și integrând conceptele de baze de date cu cele de limbaje de programare. Condițiile în care se trece de la generația a doua la a treia, sunt total diferite de condițiile în care s-a trecut de la generația întâi la a doua. Aceste condiții sunt legate de: contextul informatic, noile tipuri de aplicații și limbajele de programare.

Contextul informatic

Situația diferită dată de contextul informatic în care se face trecerea la generația a treia de baze de date este dată de evoluția sistemelor de calcul, a sistemelor de operare și a interfeței om-mașină.

a) - *Sistemele de calcul.* Sistemele relaționale au apărut în momentul în care s-a generalizat utilizarea discurilor magnetice, iar capacitatea acestora a crescut mult. La începutul anilor '80 arhitectura unui sistem informatic era centralizată, cu un server legat la un număr de terminale alfanumerice. Aceasta a influențat definirea centralizată a modelului relațional, în care o schema conceptuală globală era partajată între mai mulți utilizatori prin intermediul viziunilor. Prețul și gabariturile mari ale memoriei centrale au condus la dezvoltarea de tehnici de optimizare a acesteia. La începutul anilor '90 acest context este radical schimbat. Prețul și gabaritul componentelor de calcul au scăzut tot timpul, în

timp ce performanțele lor au crescut. Un impact deosebit pentru bazele de date l-a avut apariția stațiilor de lucru grafice interconectate în rețea. Fiecare stație grafică are puterea de calcul și performanțele unui server din anii '80, ba și mai mult. Noțiunea de stație grafică modifică arhitectura clasică a unui SGBD: puterea de calcul care era concentrată în server se găsește repartizată în stațiile de lucru. Noile sisteme vor ține cont de această descentralizare a puterii de calcul pentru o mai bună repartizare a prelucrărilor.

b) - *Sistemele de operare*. Domeniul sistemelor de operare evoluează spre funcționalități care ne amintesc de cele ale SGBD-urilor. Ele tind tot mai mult să preia din sarcinile SGBD-urilor (exemplu: sistemul de operare de pe mașinile IBM AS/400, elemente de tratare a bazelor de date în Windows). Rețeaua este invizibilă pentru utilizatori datorită folosirii unor protocoale adecvate (exemplu NTS - Network File System). La apariția SGBD-urilor în anii '60 se spunea că funcțiile acestora vor fi rapid integrate funcțiilor sistemelor de operare, ceea ce nu s-a întâmplat. Noile generații de SGBD vor ține cont de evoluția sistemelor de operare iar proiectarea și arhitectura acestor sisteme vor avea în vedere factori relevanți ai SO, cum ar fi gestiunea memoriei interne și externe, a rețelei etc.

c) - *Interfața om-mașină* a devenit tot mai performantă și mai importantă în aplicații, ca o consecință a folosirii stațiilor de lucru grafice. Multe aplicații și produse program sunt judecate după calitățile și defectele interfețelor de dialog. Aceste interfețe schimbă considerabil poziția și sarcinile utilizatorului, permițând nespecialiștilor să utilizeze aplicații complexe. În momentul de față aplicațiile conțin, de cele mai multe ori, peste jumătate din cod pentru gestiunea dialogului cu utilizatorul. Sistemele relaționale nu au fost concepute nici

pentru date grafice, nici pentru a suporta interfețe cu utilizatorul foarte laborioase. Ultimele versiuni de SGBDR au făcut totuși progrese importante în acest sens, fie prin componente generatoare specializate (Oracle, Access, Paradox), fie prin generatoare alături de un limbaj specific puternic și pentru interfața cu utilizatorul (Foxpro, dBase).

Noile tipuri de aplicații cu baze de date

Sistemele relaționale satisfac cerințele aplicațiilor tradiționale, pentru care, de fapt, au fost concepute. Descrierea datelor sub formă de tabele corespunde tipului de informații manipulate de aceste aplicații. O dată cu scăderea costului calculatoarelor și creșterea puterii lor de calcul, au apărut noi aplicații care manipulează cantități urișe de date. Printre acestea: 1) proiectarea asistată de calculator, 2) aplicațiile multimediei, 3) sistemele deschise.

Aplicațiile de proiectare asistată de calculator generează un ansamblu de faze (activități) pentru realizarea unui produs. Datele manipulate sunt adesea foarte complexe, descrierea unei componente fiind dependentă în mare măsură de celelalte componente ale aceluiași produs. Se regăsește aici și o parte importantă a informației de tip regăsim documentară.

Caracteristici ale aplicațiilor de proiectare:

a) Datele din bază reprezintă direct realul: există o legătură directă, uneori explicită, între entitățile generate de aplicație și obiectele lumii reale. Întotdeauna între datele generate și un sistem fizic din realitate va exista o legătură directă (exemplu: între datele referitoare la un circuit electronic și circuitul fizic). În aplicațiile tradiționale datele reprezintă obiecte din lumea reală, dar într-un mod mult mai puțin direct (exemplu:

datele referitoare la o clădire și clădirea propriu-zisă).

b) Datele sunt de cele mai multe ori ierarhizate. Un sistem de proiectare asistată de calculator manipulează legături de componente la subcomponente, deci o ierarhie de compoziție. Performanțele globale ale sistemului sunt, de cele mai multe ori, influențate de modul în care sunt manipulate ierarhiile.

c) Proiectarea se face numai într-un proces interactiv. Proiectarea asistată se realizează pe mașini interactive, iar rolul interfeței utilizator este de mare importanță în aceste aplicații.

Pentru a se putea face o aplicație de proiectare, trebuie să se poată modifica nu numai datele din bază, dar chiar și metodele, adică schema. Această situație este diferită de generațiile de până acum, în care schema este, dacă nu invariabilă, cel puțin relativ stabilă iar evoluția schemei este în sarcina administratorului bazei de date (care este un expert). Pentru aplicațiile de proiectare, schema trebuie să poată evolua adesea și să fie manipulată fără a apela la un expert uman.

O altă consecință a interactivității este că sistemul trebuie să furnizeze un mecanism care să permită definirea și folosirea de alternative de proiectare sau să poată alege din variantele anterioare. Ducând raționamentul la extrem, se poate spune că o aplicație de proiectare nu suprimă niciodată datele, ci acumulează versiunile succesive pe toată durata procesului de proiectare. Acest lucru este analog cu ceea ce se întâmplă deja în aplicațiile cu baze de cunoștințe din domeniul inteligenței artificiale.

d) Mai mulți proiectanți trebuie să poată lucra simultan. Aceasta caracteristică a aplicațiilor de proiectare justifică, pe lângă volumul important de date utilizat, folosirea unui SGBD ca suport. Partajarea datelor într-un mod coerent între mai mulți utilizatori este una din facilitățile oricărui SGBD. Natura acestor aplicații, face ca partajarea și contro-

lul concurenței între diferiți proiectanți, să aibă aspecte diferite față de aplicațiile clasice.

Aplicațiile multimedia au drept caracteristică aceea că administrează datele într-un mod netradițional. Exemplele cele mai cunoscute sunt aplicațiile care administrează imagini și sunet pe lângă text și grafică. Există deja acum aplicații comerciale care folosesc astfel de date, cum ar fi, de exemplu, aplicațiile meteorologice. Acest tip de aplicații se caracterizează printr-un volum foarte mare de date tratate. Imaginile sunt date foarte voluminoase care necesită un suport de stocare și prelucrare performant. Tehnologia discurilor optice numerice este adaptată acestor aplicații. Un SGBD care suportă aplicații multimedia, trebuie să folosească tratamentul clasic asupra imaginilor și să administreze legăturile multiple dintre acestea. De exemplu, într-o aplicație meteorologică trebuie căutat printre imaginile stocate, pe toate cele care ajută la detectarea unui ciclon. Pentru o astfel de operație trebuie folosită tehnica de căutare și acces clasic într-o bază de date, precum și tehnica specifică de tratare a imaginilor. Aceleași observații sunt valabile și pentru tehnica specifică de tratare a sunetului. Pentru aplicațiile multimedia este nevoie de integrarea tehnologiilor noi cu cele tradiționale. Aplicațiile multimedia există deja și reprezintă o piață potențială importantă pentru SGBD. Majoritatea dintre ele nu utilizează SGBD, ci sunt construite cu sisteme dedicate. Acest lucru se datorează faptului că un SGBDR nu oferă funcționalitățile necesare. Noile generații de baze de date vor trebui să țină cont nu numai de aplicațiile tradiționale, dar și de noile aplicații. Utilizarea unui SGBD standard în locul unui sistem dedicat, va permite reducerea considerabilă a costului de punere în funcțiune a noilor aplicații. Noile generații de SGBD vor trebui să aibă implementat conceptul de extensibilitate, adică să fie

capabile să administreze nu numai tipurile de aplicații identificate la un moment dat, ci să se adapteze la alte tipuri de aplicații care nu au fost prevăzute inițial.

Sistemele deschise corespund unui concept vag. Ideea este de a aduce un plus de flexibilitate aplicațiilor care se dezvoltă în diverse societăți.

a) *Suplețea* (flexibilitatea) în dezvoltarea și exploatarea sistemelor aferente unei unități se realizează prin:

- numeroasele periferice și platforme de diferite tipuri ce pot fi interconectate;
- ușurința cu care proiectantul poate trece de la un instrument la altul; acest lucru este facilitat de instrumentele care obligă proiectantul să utilizeze limbaje și baze de date standard;
- interconectarea, integrarea și portabilitatea aplicațiilor.

b) *Standardele (normele)* sunt deosebit de importante la sistemele deschise. Normele sunt stabilite de comitete naționale și internaționale, de grupurile de proprietari sau vânzatori sau există, pur și simplu, pentru anumite produse larg utilizate. Standardele proprietarilor sau producătorilor sunt acceptate mai ușor dacă produsul oferă facilități de interconectare cu alte produse standard.

c) *Evaluarea nivelului de deschidere* a unui produs informatic se face pe baza unor punctaje, ca urmare a răspunsului la următoarele întrebări principale:

- Este scris într-un limbaj standard sau propriu?
- Utilizează o notație a unei metode standard sau o notație proprie?
- Este exploatat cu o interfață standard sau nu?
- Permite legături cu alte instrumente sau standarde?
- Care este natura aplicațiilor pentru care se poate folosi?

Conform acestor punctaje instrumentul ideal va fi exploatabil în medii multiple și autorizat de standarde multiple.

Singurele instrumente cu adevărat deschise sunt limbajele de programare. Toți ofertanții vând instrumente care facilitează dezvoltarea de aplicații. Acestea vor introduce elemente non-standard. Nici un instrument nu este în întregime deschis decât dacă produce aplicații deschise.

d) *Relativitatea* normelelor de deschidere este datorată evoluării rapide a sistemelor. Deschiderea este o noțiune relativă și societățile tind să impună diverse tipuri de deschideri. Iată câteva caracteristici ale sistemelor deschise, în acest sens:

Punct de vedere	Nivel de deschidere al instrumentului	Grad de deschidere al aplicației produse
Mașina (hard)	Pe ce hard este exploatat instrumentul	Pe ce hard este exploatată aplicația
Sistem operare	SO utilizat Interfețe utilizate	De ce SO are nevoie aplicația Ce interfețe sunt oferite
Notație și limbaj de programare	În ce limbaj este scris instrumentul Ce limbaj utilizează proiectantul Ce notație oferă instrumentul	În ce mod utilizatorul dialoghează cu aplicația În cadrul cărui limbaj Pe baza căror notații
Biblioteci de clase	Ce biblioteci de clase poate utiliza instrumentul	Ce programe de comunicație folosește aplicația
Baze de date	Care sunt conexiunile cu bazele de date	Ce BD utilizează aplicația Cum se conectează aplicația la bazele de date

Limbajele de programare

Scrierea programelor de aplicații cu baze de date este întâi de toate o activitate de programare. Programarea aplicațiilor relaționale necesită utilizarea a două limbaje: unul de manipulare a datelor și altul pentru a programa alte funcții. Fuziunea celor două limbaje creează numeroase probleme pentru productivitatea programării, independența datelor etc. Sunt și excepții în acest sens, de exemplu, SGBDR Fox. Foxpro face o simbioză între programare și baze de date, ceea ce constituie tendința pentru bazele de date de generația a treia. De fapt tendința este de revenire la modul de programare de la generația întâi de baze de date. Multe din SGBD-uri aveau un limbaj de programare gazdă puternic, adaptat dintr-un limbaj universal (de obicei Cobol). Sistemele relaționale s-au depărtat de la această abordare, pentru ca acum să se revină. Noua generație de baze de date orientate obiect, va trebui, din punct de vedere al limbajelor de programare, să îndeplinească anumite cerințe.

În primul rând se va depăși cadrul strict al sistemelor orientate obiect. Sistemele vor combina programarea generală, esențial imperativă, cu manipularea colecțiilor de date, esențial declarativă.

În al doilea rând se remarcă un lucru specific abordării orientate obiect și anume faptul că operațiile sunt atașate anumitor tipuri de date. În SOO compunerea cererilor este dificilă atunci când rezultatul nu are aceeași natură cu datele la care sunt atașate operațiile. Noile sisteme vor trebui să ofere limbaje de cereri care să permită consultarea unei baze de date la fel de ușor ca în sistemele relaționale.

Al treilea aspect este și cel mai important, pentru că diferențiază net bazele de date de generația a doua (relaționale) de cele de generația întâi (arborescente și rețea). În bazele de date de generația întâi erau privilegiate anu-

mite legături între date și anumite drumuri de acces, noile sisteme 3G nu vor privilegia anumite accese la date. Va trebui să fie posibil să se definească legături complexe și neregulate între date, ceea ce este simplu dacă se manipulează relații, dar mai dificil dacă se manipulează înregistrări și atribute.

Conform cerințelor anterioare, noile sisteme vor trebui să preia anumite funcționalități din tehnologia relațională, alături de unele facilități din primele două generații de baze de date.

Se constată că noile SGBD vor trebui să ofere un limbaj integrat și complet, după cum noile limbaje de programare tind să manipuleze tot mai multe date.

Concluzii

Conceptul care apare cel mai des la noua generație de baze de date este integrarea: integrarea tehnologiilor de limbaje, a sistemelor de operare cu bazele de date, integrarea tehnologiilor de interfețe om-mașina cu bazele de date, integrarea instrumentelor specifice noilor tipuri de aplicații cu bazele de date. SGBD 3G vor fi sisteme integrate și extensibile cu următoarele obiective:

- o nouă abordare a programării aplicațiilor: integrarea SGBD cu un limbaj de programare
- o gestiune mai simplă a datelor și meta-datelor
- un mediu integrat care să utilizeze ce e mai bun din interfața om-mașină
- luarea în considerare a noilor arhitecturi de sisteme de calcul și de operare
- o extensibilitate care să permită adaptarea la noi tipuri de aplicații.

Bibliografie

1. Delobel C.- *Bases de donnees: des systemes relationnels aux systemes a objets*, Paris, 1991
2. Gardarin G.- *Bases de donnees les systemes et leurs langages*, Paris, 1994
3. Cattell R.- *ODMG-93 Le standard des bases de donnees objet*, 1995
4. *** - *Genie Logiciel*, martie, 1995