

BAZE DE DATE DISTRIBUITE ȘI ETEROGENE

mat. Alexandra Galatescu

Institutul de Cercetări în Informatică

Rezumat: Acest articol sintetizează cele mai importante aspecte din domeniul bazelor de date distribuite și eterogene, care este foarte puțin cunoscut și abordat în țara noastră, atât de cercetători, cât și de proiectanții de aplicații complexe. Acest domeniu poate fi abordat, atât din punct de vedere teoretic, ca un domeniu important în cercetarea fundamentală a bazelor de date care interferează cu alte domenii importante ca: bazele de date de noua generație, inteligența artificială, sisteme paralele etc., cât și din punct de vedere practic, ca un set de instrumente, medii și arhitecturi care facilitează distribuția datelor și aplicațiilor din domenii foarte variate ca: sisteme industriale, sisteme informatice de întreprindere sau macroeconomice, sisteme tranzacționale și în timp real, contabilitate bancară transnațională, sisteme de apărare națională, poștă, telecomunicații etc. Articolul prezintă succint stadiul actual în cele două direcții: cercetare și producție, încercând pe de o parte să informeze, iar pe de altă parte să polarizeze atenția și preocupările cercetătorilor și ale potențialilor utilizatori asupra acestui domeniu.

Cuvinte cheie: Baze de date distribuite și eterogene, sistem de multibaze de date, baze de date federate, eterogenitatea resurselor, autonomia locală, SGBD-uri distribuite, medii și arhitecturi distribuite, monitoare tranzacționale, arhitecturi client-server, standardizare.

1. Introducere

Bazele de date de viitoarea generație (orientate pe obiect, deductive, spațiale, temporale etc.) și bazele de date distribuite și eterogene reprezintă direcțiile actuale, principale în preocupările cercetătorilor, producătorilor și utilizatorilor bazelor de date și ale sistemelor de administrare a acestora.

În România, preocupările și rezultatele din acest domeniu au rămas la nivelul anilor '80 când s-a studiat exclusiv problema distribuției datelor, de mare actualitate în acea perioadă. În anii '90, complexitatea aplicațiilor și diversitatea resurselor conduc la necesitatea tratării distribuției datelor simultan cu eterogenitatea resurselor și autonomia locală a componentelor (aplicațiilor) distribuite.

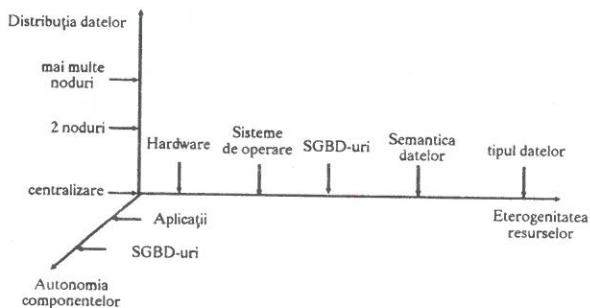


Figura 1. Caracteristicile aplicațiilor actuale

Problemele introduse de coexistența acestor trei caracteristici în aplicațiile actuale sînt reunite și rezolvate în domeniul bazelor de date distribuite și eterogene pentru care în țara noastră există o insuficientă documentare, preocupare și experiență, sursele de informare și resursele tehnice fiind încă limitate.

Acest articol își propune să motiveze utilitatea și necesitatea bazelor de date distribuite și eterogene, în perspectiva dezvoltării informaticii în țara noastră. De asemenea, își propune să prezinte succint preocupările mondiale de cercetare și de producție din domeniul bazelor de date distribuite și eterogene și din domenii conexe, precum și posibilitățile actuale de realizare a acestor tipuri de baze de date.

2. Definirea domeniului

Pentru o delimitare corectă a domeniului bazelor de date distribuite și eterogene, se vor prezenta în continuare, într-un mod simplificat, câteva definiții și concepte de bază.

O **bază de date distribuită** este compusă din mai multe baze de date, repartizate fizic pe mai multe sisteme de calcul dintr-o rețea, dar corelate din punct de vedere logic. Pentru utilizatorul din orice nod al rețelei, ea apare ca o singură bază de date, accesibilă în mod transparent (fără necesitatea cunoașterii suportului fizic al datelor).

Se va numi **sistem de bază de date (SBD)**, cuplul (SGBD, baza de date) în care SGBD este sistemul de gestiune al bazei de date. După cum baza de date este centralizată sau distribuită, SBD-ul se numește **centralizat** sau **distribuit**.

În cazul general, sistemul de gestiune al unei baze de date distribuite este cuplul $\{SBD_1, \dots, SBD_n\}$, $F(SMBD)$ numit **sistem de multibaze de date (SMBD)** sau de multiSGBD-uri, care integrează două sau mai multe SBD-uri (centralizate sau distribuite), cu diferite nivele de dependențe între ele. Integrarea SBD-urilor locale și dependențele între ele sînt realizate prin funcții și instrumente proprii SMBD-ului (notate $F(SMBD)$).

Principala deosebire dintre un SBD distribuit și un SMBD consta în faptul că, în primul datele sînt gestionate de un singur SGBD, iar în cel de-al doilea, datele sînt gestionate cu ajutorul celor "n" SGBD-uri care aparțin SBD-urilor din sistem.

O **bază de date distribuită** se numește **eterogenă**, dacă bazele de date care o compun sînt repartizate pe sisteme de calcul eterogene și/sau sînt gestionate de SGBD-uri eterogene (numai în cazul SMBD-urilor).

Un SMBD este **eterogen**, dacă SBD-urile care îl compun sînt eterogene, adică sînt gestionate de SGBD-uri diferite, sînt rezidente pe calculatoare diferite sau sînt operaționale sub sisteme de operare

diferite. El va gestiona, deci, o bază de date distribuită și eterogenă.

Pentru crearea, administrarea și utilizarea unei baze de date distribuite, SMBD-ul trebuie să asigure facilități și instrumente pentru: integrarea (totală sau parțială) a schemei globale a bazei de date, gestiunea cererilor distribuite, gestiunea tranzacțiilor distribuite, administrarea globală, transparența localizării datelor, securitatea globală, asigurarea consistenței și integrității datelor globale.

Pentru fiecare din aceste funcții, SMBD-urile eterogene trebuie să rezolve, în plus, problemele introduse de eterogenitatea calculatoarelor, a sistemelor de operare, a SGBD-urilor sau de eterogenitatea semantică a datelor (dezacordul între interpretarea, reprezentarea sau modul de utilizare a datelor cu aceeași semnificație, existente în diferite baze de date locale).

În [4] și [7] se realizează o clasificare a SMBD-urilor, după **gradul de autonomie a SBD-urilor** care le compun (vezi figura 2).

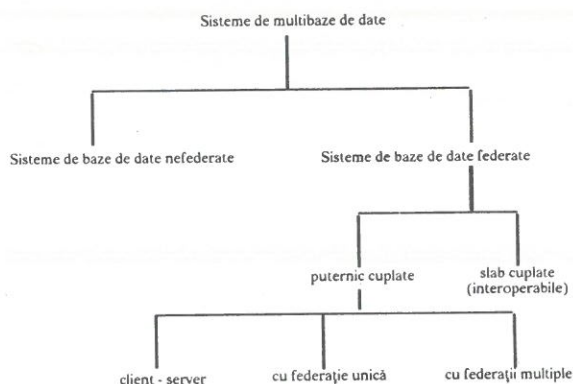


Figura 2. Clasificarea sistemelor de multibaze de date

Sistemele de multibaze de date se impart în sisteme de baze de date nefederate și sisteme de baze de date federate.

Un **sistem de baze de date nefederate** reprezintă o integrare a mai multor SBD-uri care își pierd autonomia locală. El se bazează pe o singură schemă globală a datelor, permite numai administrarea globală și accesul global și uniform la bazele de date care îl compun. Deci, accesul unui utilizator la baza de date locală nu se poate realiza direct, prin SGBD-ul local, ci numai prin interfața globală a acestui tip de sistem.

Spre deosebire de acesta, un **sistem de baze de date federate** este compus din SBD-uri parțial integrate, care își păstrează autonomia locală, dar participă la o federație care asigură partajarea parțială și controlată a bazelor lor de date.

Sistemele federate sînt împărțite, în continuare, în

sisteme **slab cuplate** (sau **interoperabile**) și sisteme **puternic cuplate**, în funcție de gradul de integrare a schemelor bazelor de date locale și de modul de administrare și de control în cadrul federației.

Aproape unanim acceptată este ideea că viitorul va aparține sistemelor federate, mai flexibile și evolutive în comparație cu sistemele nefederate. Dintre sistemele federate, pe măsură ce se vor găsi sau optimiza soluțiile de realizare, vor fi preferate cele slab cuplate, care minimizează integrarea și controlul global în cadrul federației, mărind gradul de autonomie locală a SBD-urilor participante la federație.

3. Actualitatea domeniului

Preocupările, atât teoretice cît și practice, în domeniul bazelor de date distribuite și eterogene, s-au intensificat în lumea informaticii spre sfîrșitul anilor '80. Ele au fost stimulate de necesitatea diversificării și extinderii schimburilor informaționale, precum și de progresele înregistrate în tehnologiile hardware și software privind: memoria, discurile, procesoarele, rețelele și sistemele de comunicație, sistemele de operare, software-ul de baze de date, interfețele grafice etc.

Așa cum s-a remarcat anterior, aplicațiile complexe ale deceniului '90 au noi coordonate pe plan mondial. Este tot mai mult necesar ca ele să fie distribuite (chiar multinaționale), eterogene din punct de vedere hardware și software, interoperabile, evolutive, portabile, fiabile și performante.

Ele gestionează volume mari de date și permit, în același timp, utilizatorilor nespecialiști accesul la aceste date prin interfețe interactive, prietenoase, bazate pe meniuri, ferestre, videoformate, spreadsheet-uri etc.

În plus, aplicațiile acestui deceniu vor extinde gestiunea datelor neconvenționale (hipertexte, hiperșiruri, imagini, grafice, sunete, proiecte etc.), precum și utilizarea rezultatelor inteligenței artificiale (baze de cunoștințe, sisteme expert, tehnici inteligente de regăsire etc.).

Toate aceste motive au determinat intensificarea cercetării fundamentale și aplicative în domeniul bazelor de date distribuite și eterogene, singurele capabile să cuprindă complexitatea cerințelor aplicațiilor actuale. Cercetarea este urmată îndeaproape de realizările deosebite ale producătorilor de SGBD-uri, în implementarea de funcții și instrumente utile pentru prelucrările distribuite și pentru interconectarea cu alte SGBD-uri.

Atît în cercetare, cît și în producție, acest domeniu interferează cu alte domenii de mare actualitate, cum sînt cel al bazelor de date orientate pe obiect, al inteligenței artificiale, al sistemelor paralele, al standardelor și al sistemelor deschise.

4. Direcții în cercetarea fundamentală și aplicativă

În ultimul deceniu, cercetătorii în domeniul bazelor de date distribuite și eterogene și-au conturat obiectivele, detașând aspectele critice ale domeniului și analizând resursele disponibile.

Deși în multe centre universitare și de cercetare din lume, și în primul rând din SUA, există preocupări încă de la sfârșitul anilor '70, totuși nu există rezultate remarcabile în acest domeniu care este încă nedefinitiv din punct de vedere teoretic și este parțial soluționat din punct de vedere practic (vezi [1-3]).

Problemele introduse în acest domeniu de eterogenitatea calculatoarelor, a sistemelor de operare, a rețelelor și sistemelor de comunicație au fost într-o mare măsură depășite, ca urmare a tendinței de portabilitate a SGBD-urilor și de diversificare și standardizare a interconectărilor hardware și software. Sînt încă nerezolvate sau parțial rezolvate problemele introduse de eterogenitatea SGBD-urilor datorită deosebirilor existente între modelele logice (ierarhic, rețea, relațional, obiect) implementate de SGBD-urile eterogene, precum și între mecanismele utilizate de acestea pentru accesul la date, gestiunea cererilor, gestiunea tranzacțiilor, administrare etc.

Încă în studiu sînt și problemele introduse de eterogenitatea semantică a datelor prin care, pentru valori ale atributelor cu aceeași semnificație, asociate entităților din diferite baze de date locale, se utilizează convenții diferite de reprezentare, de măsură sau de calcul, lungimi sau precizii diferite.

În contextul acestor două tipuri de eterogenități precum și al problemelor ce trebuie rezolvate o dată cu creșterea gradului de autonomie a componentelor locale, cercetătorii în domeniul bazelor de date distribuite își propun să găsească sau să-și optimizeze soluțiile pentru: integrarea schemei bazei de date distribuite, gestiunea cererilor distribuite, gestiunea tranzacțiilor distribuite, administrarea globală, asigurarea transparenței localizării datelor, standardizarea interfețelor, asigurarea securității datelor, optimizarea accesului distribuit și concurrent.

În cercetarea fundamentală mondială pentru acest domeniu principalele preocupări sînt pentru:

- definirea unui model de date comun ca interfață între modelele de date existente (ierarhic, rețea, relațional) sau de perspectivă (obiect, temporal, spațial etc.). El trebuie completat cu algoritmi de translație bidirecțională între modelele de date locale și acest model comun, precum și cu un limbaj de cereri asociat acestuia;
- definirea de algoritmi pentru integrarea schemei bazei de date distribuite;
- definirea de algoritmi pentru rezolvarea eterogenității semantice a datelor distribuite;
- definirea de algoritmi pentru asigurarea accesului concurrent și pentru prevenirea, evitarea sau

detectarea deadlock-urilor;

- definirea de algoritmi pentru optimizarea cererilor distribuite;
- definirea de algoritmi pentru gestiunea tranzacțiilor tolerante la erori.

În [9-15] sînt prezentate o mică parte dintre aceste preocupări și rezultatele obținute de cîțiva dintre cei mai activi cercetători în acest domeniu.

Dintre tipurile de SGBD-uri prezentate anterior, o preocupare deosebită în cercetarea aplicativă este pentru realizarea de sisteme federate. Instituții importante au realizat și utilizează prototipuri pentru astfel de sisteme, cum sînt: DATAPLEX (General Motors Corporation), ADDS (Amoco Production Company), IMDAS (National Institute of Standard and Technology, Florida), MERMAID (Data Integration), MRDSM (INRIA, Franța), MULTIBASE (XEROX Advanced Information Technology), arhitecturile Datacycle și OSCA (Bellcore) și multe altele prezentate în [4-7] și în [9].

Dintre sistemele federate, în practică (prototipuri, SGBD-uri, medii și arhitecturi distribuite) s-au implementat și se utilizează mai mult funcțiile sistemelor de tip client-server, din categoria sistemelor federate, puternic cuplate.

În cercetarea fundamentală însă, actuala orientare este spre sistemele slab cuplate sau interoperabile, bazate pe limbaje multibaze de date și pe tehnici ale inteligenței artificiale. Un astfel de sistem ar fi ideal prin gradul ridicat de autonomie acordată componentelor sale și prin limitarea constrîngerilor de orice tip în cadrul federației, dar este și cel mai greu de realizat. Dintre prototipurile de sisteme interoperabile realizate pînă în prezent, cel mai important este MRDSM realizat în Franța și prezentat succint în [6]. Trebuie remarcate, în mod deosebit, și cercetările actuale pentru standardizare, ale căror rezultate ar garanta succesul și perenitatea sistemelor de administrare a bazelor de date distribuite și eterogene. S-au obținut deja rezultate în standardizarea limbajului de manipulare a datelor (SQL), a prelucrării tranzacțiilor (TP), a comunicării între procese aflate la distanță (RPC), a interfețelor programabile (API), a protocolului de acces la distanță (RDA), a protocolului de transfer de date (FAP). Acestea reprezintă, însă, numai o parte din necesitățile de standardizare și nici nu rezolvă complet funcțiile pentru care sînt destinate.

5. Interferența cu alte domenii

Cercetările în domeniul bazelor de date distribuite și eterogene se intersectează cu domeniile rețelelor, sistemelor de comunicație, sistemelor de operare, precum și cu domenii mai noi ca: baze de date orientate pe obiect, sisteme paralele, inteligență artificială, standarde și sisteme deschise.

Bazele de date orientate pe obiect sînt considerate a fi potrivite pentru distribuție și interoperabilitate, datorită caracteristicilor modelului logic utilizat, mecanismelor implementate și eterogenității datelor pe care le suportă.

Deocamdată însă, cererea nu se îndreaptă către utilizarea de sisteme de multibaze de date exclusiv orientate pe obiect, ci mai degrabă către sisteme mixte, care combină SBD-uri orientate pe obiect cu SBD-uri relaționale, acestea captînd încă interesul utilizatorilor.

Apropierea între ele se realizează din ambele direcții. Pe de o parte, producătorii SGBD-urilor orientate pe obiect își îmbogățesc produsele cu interfețe SQL pentru interconectarea cu SGBD-uri relaționale. Se pot exemplifica SGBD-urile ONTOS (realizat de firma Ontologic) și O2 (realizat de firma Altair).

Pe de altă parte, firme ca Software AG și ASK/INGRES Products Division își îmbogățesc SGBD-urile relaționale ADABAS și INGRES cu interfețe orientate pe obiect.

Cercetătorii studiază, atît integrarea SGBD-urilor obiect în SGBD-uri eterogene, mixte, cît și realizarea de SGBD-uri distribuite, orientate pe obiect. Această ultimă posibilitate nu este neglijată nici de producătorii de SGBD-uri orientate pe obiect care, deja, își transportă produsele pe mai multe calculatoare și le înzestreză cu facilități pentru prelucrarea distribuită a datelor.

O altă importantă preocupare a producătorilor de SGBD-uri este transportul produselor lor pe sisteme **paralele**. Această activitate are mai multe consecințe pozitive: preîntîmpină extinderea cererii pentru calculatoare multiprocesor, îmbunătățește performanțele SGBD-urilor, mai ales în medii distribuite și asigură instrumente puternice pentru serverele de baze de date.

Dintre cele mai active firme cu preocupări în această direcție este firma ORACLE, care și-a testat deja produsul ORACLE Parallel Server pe sistemul paralel, slab cuplat de tip VAX Cluster, pe sistemul paralel masiv nCube, pe sistemul NonStop al lui Tandem și intenționează să își transporte produsul și pe alte mașini paralele.

Firme ca: Microsoft, Borland/Ashton-Tate, Sybase, ASK/INGRES, INFORMIX, UNIFY și altele își propun, de asemenea, să își transporte SGBD-urile pe sisteme multiprocesor.

Conexiunea tehnologiilor de baze de date cu tehnicile **inteligentei artificiale** este un nou aspect în cercetarea mondială, cu impact asupra realizării și optimizării sistemelor eterogene de multibaze de date. Ea ar asigura prelucrarea inteligentă și optimă a unor volume mari de date distribuite (inclusiv a celor nestructurate și neconvenționale), gestiunea eficientă a datelor și cunoștințelor existente pe mai multe noduri într-o rețea, cuprinderea și utilizarea informațiilor semantice,

optimizarea localizării datelor distribuite etc..

În [8] se prezintă succint etapele parcurse și rezultatele obținute de cercetători în încercarea de apropiere între cele două domenii. Ca cel mai recent și important rezultat, se prezintă sistemul BrAID (Bridge between Artificial Intelligence and Database Management Systems), un sistem experimental de integrare eficientă a logicii inteligenței artificiale cu tehnologiile de baze de date. Este considerat o etapă superioară față de încercările anterioare de extindere a sistemelor inteligenței artificiale cu facilități ale SGBD-urilor, de extindere a SGBD-urilor cu facilități de reprezentare a cunoștințelor sau de integrare slabă a celor două domenii. Integrarea bazelor de date distribuite și eterogene în sisteme deschise este o nouă cerință a aplicațiilor acestui deceniu. Producătorii de SGBD-uri colaborează cu producătorii de calculatoare, sisteme de operare, rețele și sisteme de comunicație pentru integrarea produselor lor în sisteme deschise. Un exemplu în acest sens este colaborarea firmei DIGITAL cu firme ca IBM, Microsoft, APPLE/Macintosh și altele pentru realizarea sistemului deschis, distribuit NAS (Network Application Support), considerat cel mai important sistem deschis, operațional în acest moment.

Integrarea într-un sistem eterogen de multibaze de date, a diferite tipuri de calculatoare, sisteme de operare, SGBD-uri, rețele și sisteme de comunicație presupune utilizarea standardelor de interconexiune între acestea. Această cerință constituie, alături de realizarea sistemelor deschise, o nouă motivare a intensificării preocupărilor actuale de **standardizare** a produselor hardware și software și a interfețelor între acestea. Alături de organizațiile ISO, ANSI, EIA, IEEE au apărut noi organizații pentru standardizare: OSF (Open Software Foundation), UNIX International, X/Open, SQL Access.

Preocupat exclusiv de standardizarea în domeniul bazelor de date, SQL Access își propune, ca principal obiectiv, crearea unor standarde pentru o reală interoperabilitate între SGBD-uri de origini diferite, funcționînd pe calculatoare diferite și sub sisteme de operare diferite.

6. Instrumente și medii distribuite

O prezentare sintetică, dar completă, a acestui domeniu trebuie să conțină și o evaluare a posibilităților actuale de realizare practică a bazelor de date distribuite și eterogene.

Referitor la resursele hardware, se poate afirma că interconexiunile între principalele tipuri de calculatoare și de rețele sînt în acest moment disponibile, deși suportă încă îmbunătățiri și diversificări. Același lucru se poate spune despre resursele software și interconexiunile între diverse

sisteme de operare și sisteme de comunicație. Se remarcă însă tendința de standardizare a sistemelor de tip UNIX și de dezvoltare a interfețelor cu acesta.

Principalele instrumente pentru realizarea, administrarea și utilizarea bazelor de date distribuite și eterogene sînt în acest moment SGBD-urile distribuite. Practic, toate SGBD-urile importante (ORACLE, INGRES, Rdb, Sybase/Microsoft, SQL Server, INFORMIX, EMPRESS, PARADOX, dBASE, FOXPRO etc.) dispun de funcții și instrumente, fie pentru prelucrări distribuite reale, fie numai pentru simple interconectări cu alte SGBD-uri, fie pentru ambele tipuri de prelucrări.

Dintre prelucrările distribuite, mai accesibile în acest moment sînt cele de tip **client-server**, deoarece majoritatea SGBD-urilor au facilități și instrumente care le permit utilizarea într-o astfel de arhitectură, pe poziția de SGBD frontal, de SGBD server sau în ambele situații. Trebuie remarcată proliferarea serverelor de baze de date cum sînt: SQL Server (Sybase/Microsoft), NetWare SQL (Novell), SQLBase (Gupta Technologies) și altele, precum și realizările unor firme importante ca ORACLE, ASK/INGRES, INFORMIX, DIGITAL, care și-au creat propriile servere de baze de date.

Ultimele versiuni ale SGBD-urilor ORACLE, INGRES, Rdb, EMPRESS dispun și de facilități și instrumente pentru vizualizări, inserări și actualizări distribuite, dar numai în cadrul unor sisteme de multibaze de date **omogene** din punctul de vedere al SGBD-urilor utilizate (în care toate SBD-urile sînt gestionate de același SGBD, indiferent de calculatorul sau sistemul de operare pe care este rezident).

Aproape toate SGBD-urile importante dispun de **pasarele** pentru accesul la baze de date create cu alte SGBD-uri, eventual pe alte tipuri de calculatoare. Pasarelele nu asigură o prelucrare distribuită reală (adică accesul simultan la baze de date eterogene), ci realizează un acces consecutiv la mai multe baze, care trebuie urmat de prelucrarea și combinarea ulterioară a datelor extrase, de către programator sau utilizator.

Utilizarea pasarelelor anticipează crearea unor baze de date distribuite și eterogene veritabile, către care se îndreaptă în acest moment eforturile cercetătorilor. Dar ele sînt neperformante și practic inutilizabile în aplicații complexe, cu un volum mare de date și prelucrări.

Efortul utilizatorului de integrare a mai multor SGBD-uri pentru realizarea, cu ajutorul pasarelelor, a unei baze de date distribuite și eterogene, poate fi micșorat prin utilizarea serviciilor oferite de **medii distribuite** ca PROGRESS, RECITAL, UDMS, SMARTSTAR și altele sau de **arhitecturi distribuite** ca NAS (DIGITAL) și SAA (IBM). Acestea, împreună cu instrumentele de interfață pe care le integrează, asigură o relativă optimizare și transparența accesului la baze de date eterogene.

Un pas important în direcția realizării sistemelor eterogene de multibaze de date îl constituie preocupările actuale ale unor firme (ex. AT&T și IBM) pentru realizarea de **monitoare tranzacționale**, deocamdată de tip client-server. Realizatorii lor își propun să integreze funcții de prelucrare tranzacțională distribuită între SGBD-uri eterogene, rezidente pe platforme eterogene. Firma IBM își propune finalizarea unui astfel de monitor pentru anul 1993, în cadrul arhitecturii sale SAA (System Application Architecture), iar AT&T a realizat TUXEDO 4.0 integrînd sisteme de tip UNIX, servere multiprocesor și o nouă versiune a SGBD-ului ORACLE. Intenționează să integreze și SGBD-urile INFORMIX, Sybase, UNISYS.

7. Domenii de aplicabilitate

Domeniile în care își găsesc utilitatea bazele de date distribuite și eterogene sînt numeroase și variate. Spectrul lor se lărgțește o dată cu creșterea complexității aplicațiilor și a instrumentelor de interfață, cu extinderea geografică a aplicațiilor, cu diversificarea tipurilor de date gestionate și cu creșterea volumului acestora.

Cele mai multe sisteme și aplicații distribuite realizate în lume pînă în prezent sînt pentru:

- informatizarea marilor întreprinderi și conerene;
- sisteme industriale care includ: automatizarea proceselor industriale, proiectarea asistată de calculator, fabricația asistată de calculator;
- sisteme tranzacționale și în timp real;
- poștă, telecomunicații și teletransmisiuni;
- sisteme de apărare națională;
- contabilitate bancară transnațională;
- documentare automată transnațională;
- birotică;
- învățămînt, educație, cultură.

În [5] și [9] sînt prezentate mai multe proiecte și prototipuri de sisteme de multibaze de date, pentru o parte din aceste domenii de aplicabilitate, iar în [16-20] se regăsesc multe din preocupările producătorilor și utilizatorilor pentru distribuire, descentralizare, interconectare și standardizare.

8. Concluzii

Acest articol a prezentat sumar aspectele principale, teoretice și practice în domeniul bazelor de date distribuite și eterogene. Scopul lui a fost să semnaleze actualitatea și importanța practică a domeniului, să atenționeze proiectanții de aplicații complexe asupra posibilităților actuale de distribuire și interconectare a bazelor de date, precum și asupra perspectivelor în acest domeniu.

Pe lângă necesitatea ei obiectivă, impusă deseori de eterogenitatea resurselor, distribuția bazelor de date are drept consecințe și **avantaje**:

- îmbunătățirea timpului de acces la volume mari de date a unui număr mare de utilizatori, prin posibilitatea separării gestiunii datelor de prelucrarea acestora;
- creșterea fiabilității bazei de date prin asigurarea disponibilității restului bazei în cazul avariei unei baze de date locale;
- facilitarea evoluției aplicației prin extinderea bazei de date distribuite sau înlocuirea sistemelor de calcul pe care este rezidentă;
- asigurarea posibilității selectării resurselor hardware și software în funcție de necesitățile aplicațiilor și de resursele financiare disponibile;
- asigurarea posibilității reutilizării resurselor hardware și software, a datelor și aplicațiilor existente deja în unitatea beneficiară.

Din păcate nu există în acest moment o metodologie coerentă de realizare a bazelor de date distribuite și eterogene, deși începând cu anul 1990 s-au intensificat în lume preocupările de ordonare și sistematizare a cunoștințelor și rezultatelor în acest domeniu.

De asemenea, prototipurile de sisteme de multibaze de date realizate pînă în prezent, deși implementează cele mai valoroase rezultate ale cercetării fundamentale, sînt destul de specializate și restrictive, fiind destinate unor tipuri particulare de aplicații cerute de beneficiarii lor.

În concluzie, în România ar fi util să se abordeze acest domeniu în paralel în două planuri:

- în cercetare, de către un colectiv multidisciplinar de specialiști cu scopul realizării unui prototip românesc de sistem de administrare a bazelor de date distribuite și eterogene, general, performant și flexibil;
- în proiectarea aplicațiilor complexe, distribuite de către colective mixte compuse din proiectanți specialiști în domeniile de aplicabilitate și cercetători sau specialiști informaticieni în calitate de consultanți în domeniul bazelor de date distribuite.

Bibliografie

1. SILBERSCHATZ, A., STONEBRAKER, M., ULLMAN, J.D.: DataBase Systems: Achievements and opportunities. În: SIGMOD RECORD, vol.19, nr.4, dec.1990, pp.6-22.
2. SCHEUERMANN, P., ELMAGARMID, A., MANOLA, F., ș.a.: Report on Workshop on Heterogeneous Database Systems. În: SIGMOD RECORD, vol.19, nr.4, dec.1990, pp.23-31.
3. GARCIA-MOLINA, H., LINDSAY, B.: Research Directions for Distributed Databases. În: SIGMOD RECORD, vol.19, nr.4, dec.1990, pp.98-103.
4. SHETH, A.P., LARSON, J.A.: Federated Database Systems for Managing Distributed, Heterogeneous and Autonomous Databases. În: ACM Computing Surveys, vol.22, nr.3, sept.1990, pp.183-236.
5. THOMAS, G., THOMPSON, G.R., CHUNG CHIN-WANG, ș.a.: Heterogeneous Distributed Database Systems for Production Use. În: ACM Computing Surveys, vol.22, nr.3, sept.1990, pp.237-266.
6. LITWIN, W., MARK, L., ROUSSOPOULOS, N.: Interoperability of Multiple Autonomous Databases. În: ACM Computing Surveys, vol.22, nr.3, sept.1990, pp.267-293.
7. SHETH, A.P.: Integrating Heterogeneous Distributed Databases: Requirements, Concepts and Solutions. În: The Sixth International Conference on Data Engineering, February 1990.
8. SHETH, A.P., O'HARE, A.B.: The Architecture of BrAID: A System for Bridging AI/DB Systems. În: The 7th Intl. Conf. on Data Engineering, Kobe, Japan, April, 1991.
9. SCHMIDT, J.W., CERI, S., MISSIKOFF, M.: Advances in Database Technology, EDBT'88. În: The International Conf. on Extending Database Technology, Italy, 1991.
10. BREITBART, Y.: Multidatabase Interoperability. În: SIGMOD RECORD, vol.19, nr.3, sept.1990, pp.53-60.
11. THOMPSON, G.R.: Multidatabase Concurrency Control. Lucrare de doctorat la Oklahoma State University, 1987.
12. GELLER, J., MEHTA, A., PERL, Y., ș.a.: Algorithms for Structural Schema Integration. În: Research Report at Institute for Integrated Systems, 1991.
13. MEHROTRA, S., RASTOGI, R., KORTH, H., SILBERSCHATZ, A.: Relaxing Serializability in Multidatabase Systems. În: The 2nd Intl. Workshop on Research Issues in Data Eng., febr.1992.
14. BREITBART, Y., SILBERSCHATZ, A., THOMPSON, G.R.: Transaction Management Issues in a Failure-prone Multidatabase System Environment. În: The VLDB Journal, 1992.
15. GEORGAKOPOULOS, D., RUSINKIEVICZ, N., SHETH, A.: On Serializability of Multidatabase Transactions through Forced Local Conflicts. În: The 7th Intl. Conf. on Data Eng., Japan, apr.1991.
16. Colecția revistei "Le monde informatique", 1991.
17. Colecția revistei "DATAMATION", 1991.
18. Colecția revistei "01 Informatique", 1990.
19. Colecția revistei "BYTE", 1991.
20. Colecția revistei "DEC Professional", 1990.