

Concept voor een volledig nieuwe architectuur

Het Elementair DataBase Management Systeem

Hans van Bakkum

Het Elementair DataBase Management Systeem, verder EDBMS te noemen, is een volledig nieuw soort Relationeel Database Management Systeem (RDBMS), geschoeid op het object-relatieve model. Deze volledig nieuwe architectuur heet Elementair, omdat alle attributen en objectwaarden als een Elementaire waarde in de verschillende elementen van de eigenlijke database worden opgenomen, waarbij de elementen van deze database een vaste grootte in bytes (of eventueel bits) hebben in de orde van grootte van 4 bytes.

In vergelijking met een ander RDBMS bedraagt het ruimtebeslag van de eigenlijke database zo'n 20 procent, zodat opname van de volledige database in het geheugen zeer goed mogelijk is.

De database kan, naar behoefte, op elk willekeurig moment zowel naar boven als naar beneden geherdimensioneerd worden (vergroting met 1 byte levert daarbij een capaciteitstoename op met een faktor 256, verkleining met 1 byte een capaciteitsafname met een faktor 256).

Logisch gezien is de eigenlijke database een 3-dimensionale array met als coördinaten Tabelnummer, Regelnummer en Attribuutnummer (zie afbeelding 1), waarbij zowel het maximale Regelnummer als het maximale Attribuutnummer voor alle tabellen hetzelfde zijn.

Technisch gezien is deze database te beschouwen als een 3-dimensionale array binnen het logische model met als coördinaten Tabelnummer, Regelnummer en Kolomnummer (zie afbeelding 2), waarbij zowel het maximale Regelnummer als het maximale Kolomnummer voor elke tabel hun eigen waarden hebben.

BEHEER

Wat het EDBMS verder onderscheidt van elk ander RDBMS, is de opbouw van de database in een aantal Beheergedeelten, namelijk een Databasebeheergedeelte, een Systeembeheergedeelte en afhankelijk van de organisatie een of meer Gebruiksbeheergedeelten (zie afbeelding 3); door deze indeling is een zeer goed en afgescheiden beheer op alle niveaus mogelijk. Elk beheer-

gedeelte omvat een of meer tabelgroepen, waarbij het beheer per tabelgroep afzonderlijk kan worden geregeld.

SELECTIES

Door de eerdergenoemde 3-dimensionale opbouw zijn 3-dimensionale selecties op de database mogelijk. We onderscheiden daarbij de punt-selectie, de lijn-selectie, de vlak-selectie en de balk-selectie.

Bij een punt-selectie zijn alle drie de coördinaten (Tabelnummer, Regelnummer en Attribuut- c.q. Kolomnummer) bekend (zie afbeelding 4).

Bij een lijn-selectie zijn twee van de drie coördinaten bekend, Tabelnummer en Regelnummer, Tabelnummer en Attribuut- c.q. Kolomnummer (zie afbeelding 5) of Regelnummer en Attribuut- c.q. Kolomnummer; de onbekende coördinaat is variabel.

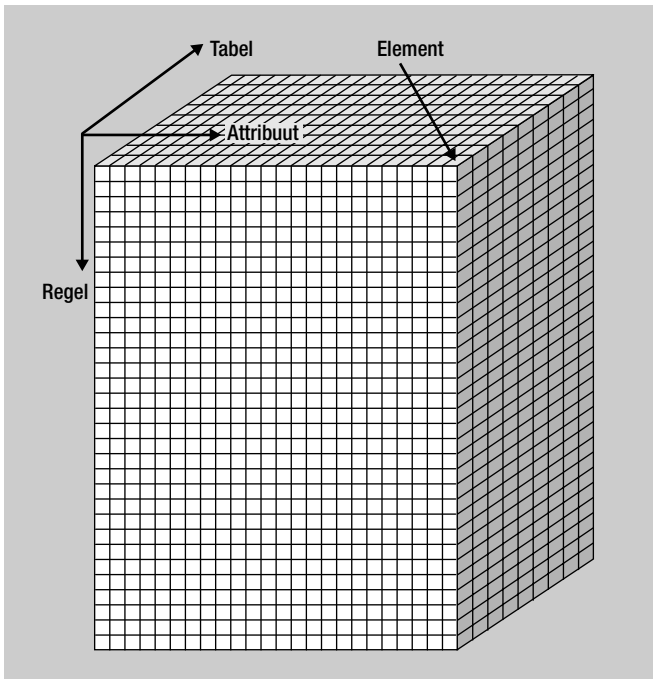
Bij een vlak-selectie is slechts een van de drie coördinaten bekend, Tabelnummer (zie afbeelding 6), Regelnummer of Attribuut- c.q. Kolomnummer; de onbekende coördinaten zijn variabel.

Bij een balk-selectie is geen van de drie coördinaten bekend; zowel Tabelnummer als Regelnummer als Attribuut- c.q. Kolomnummer zijn variabel (zie afbeelding 7).

Hierbij dient echter wel bedacht te worden, dat elke variabele niet als lineair sequentieel gezien behoeft te worden, maar ook lineair gebroken kan zijn, bijvoorbeeld Tabelnummer = 3 tot en

Het Elementair DBMS

Hans van Bakkum is zelfstandig consultant en werkt sinds de jaren zeventig in de IT. Hij heeft vele databases gebouwd, zowel in het pré-relatieve tijdperk als daarna. Het gevoel groeide dat de RDBMS'en niet optimaal worden benut. Hij zette zijn ideeën op papier en ontwikkelde een concept dat hij het Elementair DBMS noemde. Het EDBMS zou het RDBMS overal kunnen vervangen. Hij kreeg de laatste jaren veel positieve reacties op zijn ideeën; DB/M biedt hem de kans een groter publiek te bereiken.

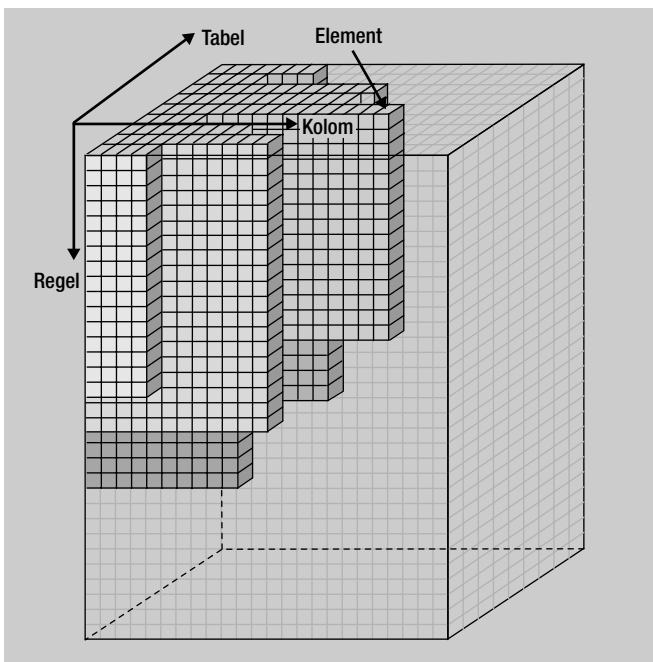


AFBEELDING 1: LOGISCH DATAMODEL EDBMS-DATABASE.

met 5, 7 tot en met 11, 15 of 46 tot en met 50 in plaats van Tabelnummer = 3 tot en met 50; bovendien kan daarbij Tabelnummer ook afgeleid zijn van Tabelnaam.

DATATYPEN

De Elementaire waarden in de elementen van de database worden voor elk Datatype verkregen via het daarbij behorende algoritme. Voor numerieke waarden, zoals bijvoorbeeld numerieke waarde zonder teken, numerieke waarde met teken, datumwaarde,



AFBEELDING 2: TECHNISCH DATAMODEL EDBMS-DATABASE.

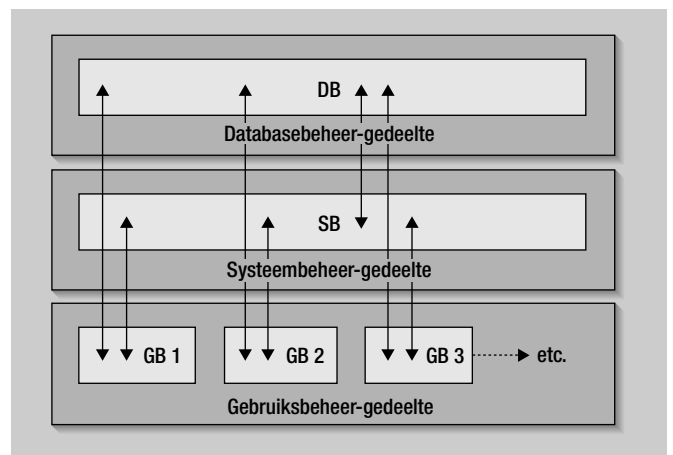
tijdwaarde enzovoort, is dit een eenvoudig algoritme, waarbij de aangeboden waarde rechtstreeks wordt omgezet in een Elementaire waarde. Hierbij wordt tevens gebruik gemaakt van een Basiswaarde (standaard 0) en een Waardegrondslag (standaard 1) om een zo klein mogelijke Elementaire waarde te verkrijgen. Dit wordt verzorgd door de bijbehorende Inputprocessor (zie afbeelding 8), waarbij geldt:

$$\text{Elementaire waarde voor opslag} = (\text{omgerekende numerieke Attribuutwaarde} - \text{omgerekende Basiswaarde}) / \text{omgerekende Waardegrondslag} \text{ ROUNDED.}$$

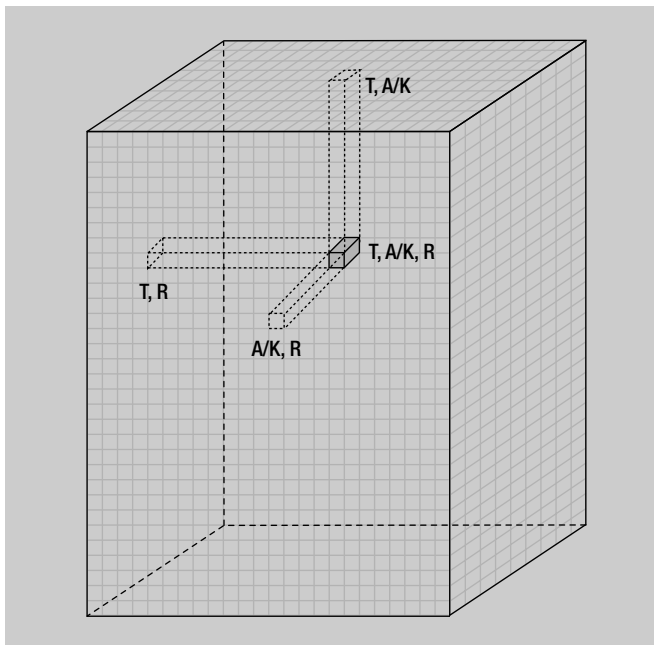
Voor niet-numerieke attribuutwaarden wordt de betreffende waarde eerst via de bijbehorende Vertaaltabel omgezet (vanwege de door de gebruiker gewenste alfabetische volgorde) en met behulp van de omgezette waarde wordt de bijbehorende Elementaire waarde opgezocht in de bijbehorende (Attribuut)Waarde-tabel, dan wel wordt eerst de nieuwe omgezette attribuutwaarde opgenomen in deze Waarde-tabel en wordt een Elementaire waarde toegekend overeenkomstig de alfabetische plaats van deze waarde ten opzichte van de dichtstbijliggende hogere en de dichtstbijliggende lagere waarde in de betreffende Waarde-tabel.

Voor objectwaarden wordt de betreffende waarde eerst via de bijbehorende Objectgenerator omgezet en wordt de omgezette waarde opgezocht in de bijbehorende (Object)Waarde-tabel, dan wel wordt eerst de nieuwe omgezette objectwaarde opgenomen in deze Waarde-tabel en wordt een Elementaire waarde toegekend, overeenkomstig de gewenste plaats van deze waarde ten opzichte van de dichtstbijliggende hogere en de dichtstbijliggende lagere waarde in de betreffende Waarde-tabel, of wordt eventueel de eerstvolgende vrije Elementaire waarde toegekend, een en ander afhankelijk van het bijbehorende algoritme.

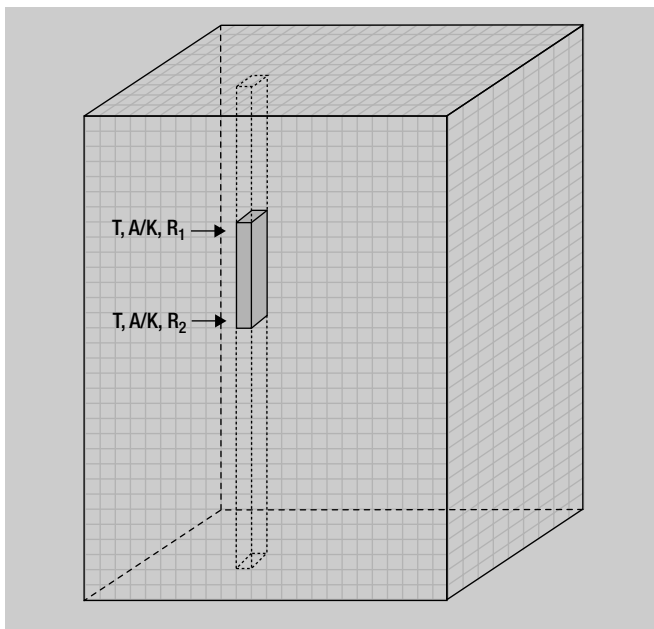
Het database-systeem is onbeperkt uit te breiden met nieuwe datatypen, zoals bijvoorbeeld numerieke waarden met teken met onbekende delen, datumwaarden met onbekende delen, niet-numerieke waarden met een andere Vertaal tabel en alle mogelijke soorten objecten, zoals bijvoorbeeld XML-documenten.



AFBEELDING 3: BEHEERGEDEELTEN EDBMS-DATABASE.



AFBEELDING 4: VOORBEELD VAN EEN PUNT-SELECTIE.



AFBEELDING 5: VOORBEELD VAN EEN LIJN-SELECTIE MET VASTE TABEL EN ATTRIBUUT/KOLOM EN VARIABLE REGEL.

FLEXIBILITEIT

Het EDBMS kan slechts een database omvatten met bijbehorende (Attribuut)Waarde-tabel(len) en (Object)Waarde-tabel(len), zie afbeelding 9a.

Ook kan het EDBMS meerdere databases omvatten met bijbehorende (Attribuut)Waarde-tabel(len) en (Object)Waarde-tabel(len), waarbij een of meer van deze Waarde-tabellen gedeeld kunnen worden door een aantal van deze databases; de overige Waarde-tabel(len) zijn dan database-gebonden (zie afbeelding 9b).

TIJDSAFHANKELIJKHEID

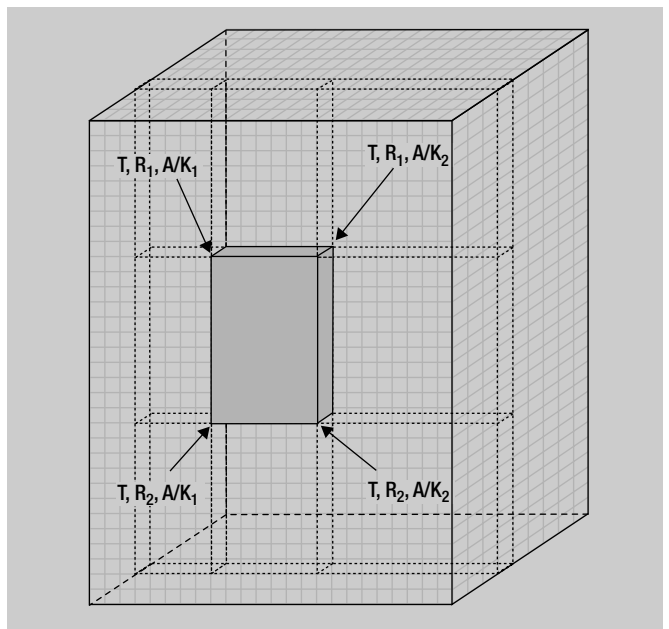
Ook volledig nieuw is de tijdsafhankelijkheid van alle tabelregels in de database(s). Deze tijdsafhankelijkheid wordt verkregen, door aan elke tabelregel een datum en tijd ingang geldigheid en een datum en tijd einde geldigheid mee te geven, zodat van te voren kan worden vastgelegd, vanaf wanneer tot wanneer een bepaalde tabelregel actueel is in de database.

Uiteraard wordt de integriteit binnen de database tussen deze twee tijdstippen streng gecontroleerd. Door dit mechanisme, dat volledig in de architectuur van het EDBMS is ingebouwd, wordt het database-systeem dynamisch in de tijd; dit in tegenstelling tot de huidige RDBMS-en, die allen statisch in de tijd zijn.

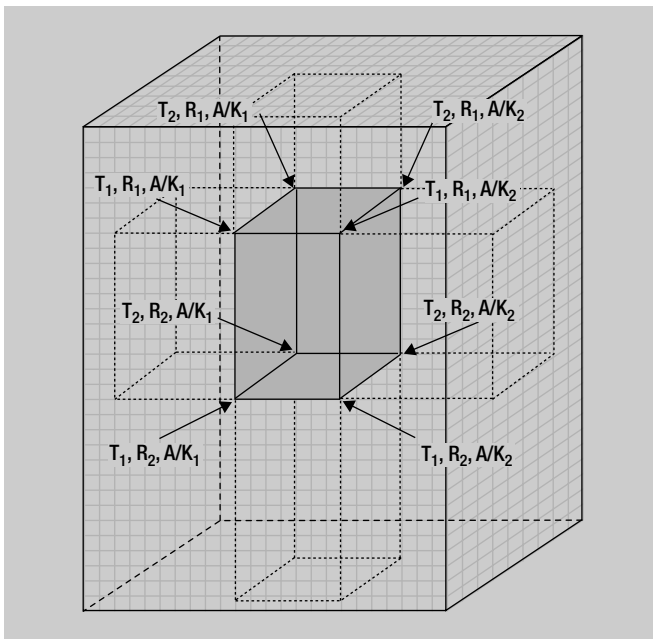
We onderscheiden nu dan ook verschillende soorten tabelregels, namelijk *actuele* tabelregels, *niet meer actuele* tabelregels en *nog niet actuele* tabelregels.

Actuele tabelregels zijn tabelregels, die op de Systeemdatum en -tijd hun geldigheid bezitten; hiervoor geldt dat datum en tijd einde geldigheid > Systeemdatum en -tijd >= datum en tijd ingang geldigheid. Deze tabelregels verliezen bij het verstrijken van de tijd automatisch hun geldigheid. Niet meer actuele tabelregels zijn tabelregels, die op de Systeemdatum en -tijd geen geldigheid meer hebben; hiervoor geldt, dat Systeemdatum en -tijd >= datum en tijd einde geldigheid. Deze tabelregels worden na een zekere bewaarperiode automatisch uit de database verwijderd. Nog niet actuele tabelregels zijn tabelregels, die op de Systeemdatum en -tijd nog geen geldigheid hebben; hiervoor geldt, dat datum en tijd begin geldigheid > Systeemdatum en -tijd. Deze tabelregels worden bij het verstrijken van de tijd automatisch geldig.

Aangezien de integriteit tussen begin en einde geldigheid van tevoren reeds volledig is gecontroleerd, heeft noch het geldig



AFBEELDING 6: VOORBEELD VAN EEN VLAK-SELECTIE MET VASTE TABEL EN VARIABLE REGEL EN ATTRIBUUT/KOLOM.



AFBEELDING 7: VOORBEELD VAN EEN BALK-SELECTIE.

worden in de tijd noch het ongeldig worden in de tijd, invloed op de totale integriteit van de database. Om flexibel om te kunnen gaan met deze tijdsaspecten zijn twee nieuwe database-commando's in het leven geroepen, te weten het HISTORIZE-commando en het FUTURIZE-commando.

Met behulp van het HISTORIZE-commando kunnen niet meer actuele tabelregels weer actueel worden gemaakt dan wel actuele tabelregels niet meer actueel. Met behulp van het FUTURIZE-commando kunnen nog niet actuele tabelregels actueel worden gemaakt dan wel actuele tabelregels nog niet actueel. Daarbij blijft de integriteit van de database in de tijd volledig gehandhaafd.

Door deze ingebouwde tijdsafhankelijkheid zijn selectiemogelijkheden in de tijd bovenop de 3-dimensionale selectiemogelijkheden in de ruimte mogelijk geworden. Zo zijn zowel actuele selecties (selectietijdstip in het heden), historische selecties (selectietijdstip in het verleden) als futuristische selecties (selectietijdstip in de toekomst) mogelijk.

Bij actuele selecties (zie afbeelding 10) geldt de Systeemdatum en -tijd op het moment, dat de selectie begint; alleen tabelregels waarvoor geldt, dat deze actueel zijn op dit begintijdstip, worden bij het verdere selectieproces meegenomen (een zelfde selectie kan iets later al een ander resultaat geven, omdat de database tijdens de selectie gewijzigd kan zijn door toevoeging, wijziging of verwijdering van tabelregels).

Bij historische selecties geldt het opgegeven tijdstip (datum en tijd) in het verleden; alleen tabelregels, die op dat tijdstip actueel waren worden bij het verdere selectieproces meegenomen (het resultaat van een zelfde selectie zal op elk willekeurig tijdstip altijd hetzelfde resultaat geven, omdat het verleden nu eenmaal vastligt en niet meer veranderd kan worden).

Bij futuristische selecties geldt het opgegeven tijdstip (datum en tijd) in de toekomst; alleen tabelregels, die op dat tijdstip actueel

zullen zijn, worden bij het verdere selectieproces meegenomen (het resultaat van een zelfde selectie zal op elk willekeurig tijdstip meestal een ander resultaat geven, omdat de toekomst nog onzeker en veranderlijk is; bovendien zal deze selectie door het verstrijken van de tijd op een zeker tijdstip actueel worden en daarna zelfs historisch).

Om de actuele toestand van de database op het opgegeven tijdstip exact te kunnen bepalen zal ook de database-logging bij het selectieproces worden betrokken, want de actuele toestand van dat tijdstip is wellicht niet of niet meer in de database zelf aanwezig (dit geldt overigens voor alle hierboven vermelde selecties).

STATUS

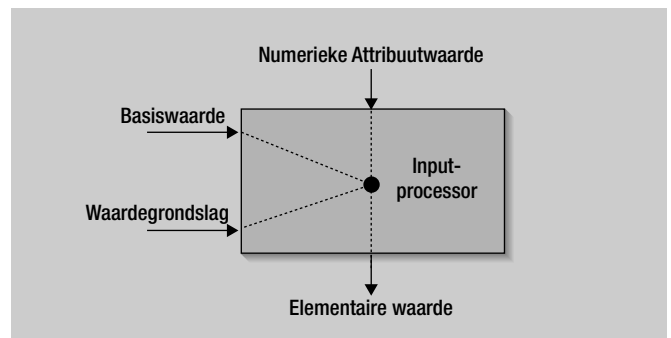
Alle selecties hebben standaard betrekking op actieve en tijdelijk inactieve tabelregels; deze status hoeft daarom niet te worden opgegeven bij een selectie. Worden bij een selectie een of meer statuswaarden opgegeven, dan geldt(t)(en) de opgegeven waarde(n).

De normale status van de tabelregels is de actieve status, dat wil zeggen ze kunnen normaal gewijzigd of verwijderd worden en ze gaan standaard mee met het selectieproces. Ook bevat de database tabelregels, die bij de initialisatie van de database door het database-systeem zelf automatisch zijn gegenereerd; deze kunnen noch gewijzigd noch verwijderd worden, maar kunnen wel normaal worden geselecteerd. Het is mogelijk tabelregels tijdelijk of permanent te inactiveren in de database, dan wel tabelregels direct uit de database te verwijderen.

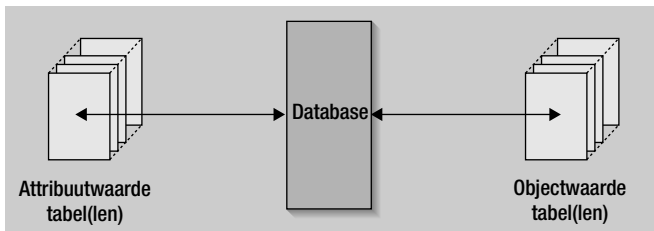
Met behulp van het DEACTIVATE-commando worden een of meer tabelregels tijdelijk geïnactiveerd in de database, dat wil zeggen dat deze tabelregels niet meer gewijzigd of verwijderd kunnen worden zolang deze status gehandhaafd blijft, maar ze gaan wel standaard mee met het selectieproces.

Met behulp van het ACTIVATE-commando kunnen dergelijke tabelregels weer geactiveerd worden in de database, waarna alle andere bewerkingen daarop weer normaal mogelijk zijn.

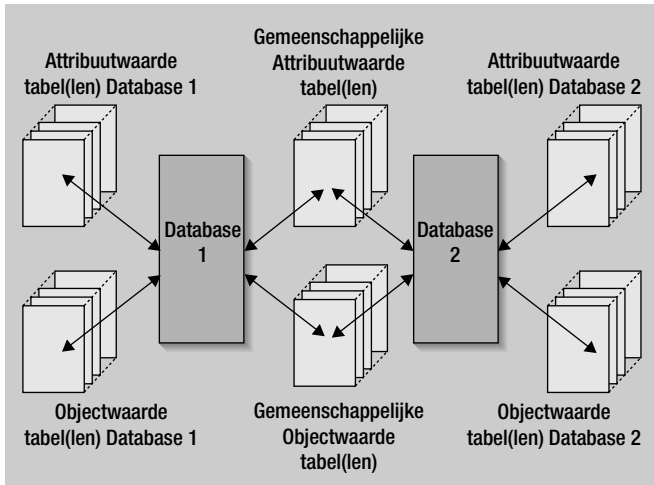
Met behulp van het DELETE-commando worden een of meer tabelregels permanent geïnactiveerd in de database, dat wil zeggen dat deze tabelregels niet meer gewijzigd of verwijderd



AFBEELDING 8: INPUTPROCESSOR VOOR NUMERIEKE ATTRIBUUT- WAARDEN.



AFBEELDING 9A: VOORBEELD VAN EEN EDBMS MET 1 DATABASE EN BIJBEHORENDE ATTRIBUUT- EN OBJECTWAARDE TABEL(LEN).



AFBEELDING 9B: VOORBEELD VAN EEN EDBMS MET 2 DATABASES, GEMEENSCHAPPELIJKE ATTRIBUUT- EN OBJECTWAARDE TABEL(LEN) EN 'EIGEN' ATTRIBUUT- EN OBJECTWAARDE TABEL(LEN).

kunnen worden zolang deze status gehandhaafd blijft en ze gaan bovendien niet mee met het selectieproces, tenzij daar uitdrukkelijk om wordt gevraagd. Deze tabelregels worden na een zekere bewaarperiode automatisch door het database-systeem uit de database verwijderd en in de database-logging opgenomen.

Met behulp van het UNDELETE-commando kunnen dergelijke tabelregels weer geactiveerd worden in de database, als ze tenminste nog niet automatisch daaruit zijn verwijderd, waarna andere bewerkingen weer normaal mogelijk zijn.

Met behulp van het REMOVE-commando worden een of meer tabelregels direct uit de database verwijderd en in de database-logging opgenomen.

Ook is het mogelijk met behulp van het UNDO-commando bepaalde acties of alle acties, die in een bepaalde periode op de database zijn uitgevoerd, weer ongedaan te maken. In alle gevallen wordt er streng op toegezien, dat de integriteit van de database gehandhaafd blijft.

Door de fysieke scheiding tussen de database(s) enerzijds en de Waarde-tabel(len) anderzijds is het database-systeem optimaal te beveiligen. Bovendien is de inhoud van de database zelf voor buitenstaanders niet of nauwelijks te interpreteren vanwege de elementaire opslag van de verschillende datatypen. Ook is de toegang tot de verschillende Beheergedeelten van de database en de daarin te onderscheiden tabelgroepen binnen het database-systeem zeer streng geregeld.

CONCLUSIE

Het EDBMS is een zeer flexibel, uitbreidbaar en dimensioneerbaar RDBMS, met vrijwel ongekende selectie-mogelijkheden. De bijbehorende database is bovendien zeer compact van omvang ten opzichte van die van een gelijksoortig RDBMS, zodat het in zijn geheel of grotendeels permanent in het geheugen kan worden opgenomen. Dit laatste geldt eveneens voor de bijbehorende Waarde-tabel(len).

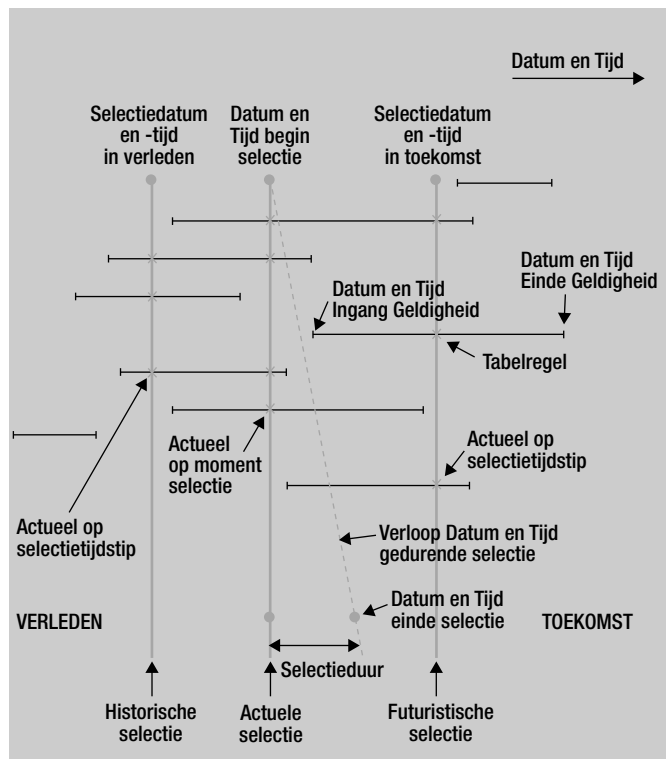
Vanwege de in de architectuur ingebakken tijdsafhankelijkheid is het bovendien een echt dynamisch, want 'levend', RDBMS, omdat het verstrijken van de tijd constant van invloed is op de actualiteit van de aanwezige tabelregels.

Hoewel elke tabel, onder andere vanwege de ingebouwde tijdsafhankelijkheid, een aantal attributen extra heeft ten opzichte van een normaal RDBMS en ondanks het feit dat attribuutwaarden en objectwaarden altijd worden omgezet naar elementaire waarden en omgekeerd, kan toch een uitstekende performance worden verkregen, onder meer doordat vrijwel alles permanent in het geheugen aanwezig is.

Doordat beheeraspecten volledig zijn ingebouwd, is een optimaal beheer mogelijk op alle beheerniveaus. Bovendien is het EDBMS zeer geschikt voor velerlei uiteenlopende toepassingen. Een (Attribuut)Waarde-tabel is bijvoorbeeld ook zeer geschikt als thesaurus.

Ing. H.N.M. van Bakkum (hansvanbakkum@hotmail.com)

is zelfstandig consultant.



AFBEELDING 10: VOORBEELD VAN EEN HISTORISCHE, ACTUELE EN FUTURISTISCHE SELECTIE.