

Utfärdare
2540 Henrik Holst/ed 031-667640

Datum
89-06-27 I330025 Sida 1

Ärende

INFORMATION

Dokument: Referat av Yankee Group Europe-rapport, jan 1989
Avser: DDBMS - A European Perspective

Sammanfattning av "Executive Summary"

"Cynikern kan avfärda distribuerade databaser (DDBMS) som ännu en lösning i behov av ett problem."

YG håller inte med om denna beskrivning, som ska framgå av nedanstående.

Först en distinktion: distribuerad datakraft skall inte förväxlas med DDBMS. Det har länge funnits mer eller mindre distribuerad databehandling, men som Chris Date uttryckt saken:

"Sömmarna syns", dvs användaren, eller applikationsprogrammet, är på något sätt medveten om att datat är distribuerat. Så är inte fallet med DDBMS; För användaren ser en distribuerad databas exakt likadan ut som vore den inte distribuerad.

Teknologin för att etablera DDBMS finns nu eller är på väg:

Relationsdatabasen (RDBMS), ett aktivt data dictionary för att 'separera' användaren från datat, standardiserat språk (SQL) för att manipulera data samt nätverk, utgörande de 'motorvägar' på vilka datat skall transporteras. Trycket från den affärsmässiga sidan att etablera DDBMS utgörs främst av en allt mer tilltagande globalisering av företagens verksamhet.

Den 31 december 1992, då EG har rivit alla interna handels hinder, kommer att innebära en hemmamarknad vilken på sikt kommer att bli större än USA.

Men vi är inte framme vid DDBMS än. Diskussionen för dagen förs till största delen i tekniska termer, och detta är lite grand av att spänna vagnen framför hästen. Fler stora frågor är fortfarande kvar att lösa innan företagen kan utnyttja DDBMS fullt ut.

De fundamentala frågor som måste ersätta nuvarande mer tekniska resonemang är:

- vem äger vilket data?
- hur struktureras applikationsövergripande data?
- vilken teknik skall användas för att bygga de nödvändiga datastrukturerna?
- etc.

Dvs, en "corporate data architecture" måste etableras.

Det finns ingen enkel väg från nuvarande strukturer till RDBMS och DDBMS, trots att en del migrationshjälpmedel har sakta börjat se dagens ljus. YG förutspår en blivande jättemarknad för migrationskonsulter och dito hjälpmedel.

Sammanfattningsvis anser YG att en massiv etablering av DDBMS är osannolik på kort sikt, beroende på de tekniska och framför allt organisatoriska problem som finns. På kort sikt är det mer troligt att DDBMS implementeras delvis, typ read-only-access till distribuerat data.

Tekniska frågor

Ted Codd, "uppfinnaren" av relationsdatabasen, har definierat 12 regler som ett DBMS skall uppfylla för att kallas 'relational'. Hans kollega Chris Date har tagit fram ytterligare 12 regler som definierar en distribuerad databas (DDBMS). Date framhåller dock att "... det finns inget 'slutgiltigt distribuerat databassystem'. Olika användare kommer att framhålla olika kriterier som viktiga."

Bakom dessa 12 regler finns 'regel 0', som säger att ett distribuerat system skall upplevas av användaren som ett icke distribuerat system.

De 12 reglerna är:

1. Lokal autonomi
2. Inget beroende av en central datormiljö
3. Avbrottsfri drift
4. Transparent dataplacering
5. Oberoende av fragmentering av data
6. Oberoende av replikation av data
7. Distribuerad query-hantering
8. Distribuerad transaktionshantering
9. Hårdvaru-oberoende
10. Operativsystem-oberoende
11. Nätverk-oberoende
12. RDBMS-oberoende

Beståndsdelarna i ett DDBMS

Följande grundläggande byggstenar utgör grunden för ett DDBMS:

1. Ett relationsdatabassystem (RDBMS)
2. Ett aktivt data dictionary

3. Språk, Structured Query Language (SQL)
Det finns både för och nackdelar med SQL:
 - + 'Set-at-a-time processing'
 - + Inbyggd säkerhetshantering
 - Strukturen uppmuntrar till flera, redundanta sätt att ställa en fråga
 - Ej stöd för primary-key/foreign-key eller tillhörande integritetsproblem
4. Nätverk
Peer-to-peer-kommunikation (APPC, LU6.2 kallas det också ibland) har accepterats bland sw-företagen som standard.
5. Mekanism för att hantera en distribuerad databas
Olika tekniker finns för att 'dela upp' data:
 - kopior i olika former (snapshots, extrakt o.d.),
 - partitionering
 - = horisontell, dvs uppdelning map rader i en tabell
 - = vertikal, dvs uppdelning map kolumner i en tabell
6. Transaktionshanterare
För att kunna säkerställa data-integriteten i en transaktionsomgivning krävs funktioner för bl.a synkronisering av accesser och backup/recovery.

Hårdvaruplattformar

DDBMS implementeras på stordatorer (framför allt IBM/370), mid-range (VAX o.d. dominerar), och PC-baserade LANs. Stordatormiljön domineras av IBM med DB2, men flera oberoende softvaruföretag förekommer här också. DEC gör inbrytningar på denna plattformen med VAX-clusters.

Mid-range med VMS som ledande sw-plattform (men även UNIX) domineras av företag som Oracle, RTI, Sybase och Tandem.

Den tredje plattformen, PC-baserade LANs, har haft liten betydelse än, men där bedömer YG att PS/2 med OS/2 EE, Presentation Manager och MS-LAN kommer att vara nyckelprodukter.

Den nya europeiska informations-infrastrukturen

1992, dvs EG's inre marknad, innebär en stor stimulans för internationalisering. Detta innebär i sin tur ökad konkurrens som medför ett ökat behov av att kunna knyta ihop geografiskt spridd information, dvs organisera dataresurserna, dvs skapa en ny data-arkitektur.

Rapporten dyker sedan ner i en historisk återblick över det elände vi alla befinner oss i för närvarande. Där berörs applikationer vilka specialbyggts för varje funktion (tex orderhantering) och hur dåligt dessa är anpassade till

morgondagens krav på information.

YG Europe anser att pga att tekniken för att distribuera databaser nu finns (dock mycket rudimentär/öv. anm.) är det ett av dom bästa argumenten för att börja bygga en övergripande dataarkitektur och en dito data-administration.

Utvecklingsvägen, hur DDBMS skall byggas, följer enligt YG väl den modell över informations-systemens utveckling, som beskrivits av Nolan 1982. Sex stadier finns beskrivna:

1. Automatisera grundläggande affärsfunktioner
2. De flesta operationella funktioner har automatiserats var för sig med stor dataredundans som ett resultat
3. Försök att kontrollera det kaos som uppstått efter 1 och 2 ovan.
4. Existerande applikationer börjar struktureras om för att ta till vara databas-teknologin
5. Alla datasystem delar en central kärna av 'corporate data'.
6. Total separation av data och de datoriserade processer de deltar i.

Problemet för många företag är att komma förbi steg 4. Detta kan endast göras genom att etablera en övergripande funktion för data-administration, som skall implementera och sköta en 'corporate data structure'.

Data Administration

YG anser att behovet att etablera en företagsgemensam datastruktur (corporate data structure) gör det nödvändigt att centralisera kontroll och styrning (management) av data under en "Data Administration Group", och den bästa platsen organisatoriskt för en sådan grupp är UTANFÖR data avdelningen, inom vad YG kallar "business planning".

Data-administratörens roll

Huvudarbetsuppgifterna för en Data Administratör skall vara:

1. ansvara för data-administrations policy
2. förankra en medvetenhet om vikten av data inom HELA organisationen
3. data analys; välja metod och ta fram verktyg/hjälpmedel
4. data definition standards
5. data dictionary control
6. samarbeta med databas-administratören/erna
7. data access; auktoritetsregler
8. implementera strategier för säkerhet/integritet

Användarfrågor och migreringsproblem

Ytterst få företag har kommit in i en stabil miljö med RDBMS, ännu mindre med DDBMS. En eventuell konvertering ligger i framtiden för de flesta.

YG-rapporten beskriver i en fallstudie, hur Londons Metropolitan Police Force håller på att implementera ett DDBMS med hårdvara från DEC, MicroVAX till VAX 8800. Mjukvaran är Ingres från Relational Technology. Systemet går under beteckningen CRIS.

Utan att gå in på detaljer, vilka det finns gott om i rapporten, kan två saker nämnas:

- pga teknologiska begränsningar har man fått göra stora kompromisser, systemet är knappast ett DDBMS längre utan snarare en samling RDBMS som samverkar i ett nätverk.
- trots detta vill YG beskriva projektet som succé (än så länge) många och nyttiga lärdomar har dragits och kommer att dragas framöver.

Produktionsstart i begränsad skala är planerad till tidigt 1989, med färdigtid ca 2 år senare, alltså 1991.

Sedan följer en diskussion om distribuerad datakraft vs. distribuerade databaser, DDBMS.

Fördelarna med det enklare alternativet distribuerad datakraft är enligt YG:

- få organisationer har idag den dataarkitektur som behövs för att implementera DDBMS.
- det är enklare att distribuera datakraft än data
- många organisationer är idag sysselsatta med att fundera ut (och implementera) hur man ska gå från pre-relational till relational, RDBMS. Att tänka på DDBMS i det läget är inte möjligt.
- det är fortfarande oklart till vilken grad datat inom en organisation behöver synkroniseras, dvs sekund-, minut-, tim-synk eller annat.

YG drar slutsatsen att den äkta distribuerade databasmiljön inte kommer att accepteras. Företagen kommer i allmänhet att gå in för hybridlösningar, tex distribuerad read-only. Detta inom den överblickbara framtiden i alla fall.

Prestanda

Prestandamässigt, anser YG, finns det ingen anledning att ett RDBMS inte skulle kunna klara alla situationer (bortsett från vissa extrema). Men innebär det att även DDBMS skulle vara 'i hamn'? Svaret är NEJ, det största enskilda problemet för DDBMS idag är prestanda. YG karaktäriserar problemet så här:

Global DDBMS prestanda

=

Lokalt RDBMS prestanda + CPU-prestanda + Nätverkprestanda

Metoder och verktyg

Stor brist på verktyg råder på metodfronten för DDBMS. Som exempel nämner YG att det idag inte finns något hjälpmedel för att bestämma den optimala placeringen av data inom ett DDBMS nätverk. Detta är bara ett exempel. Dock, framhåller YG, är det på gång metoder och hjälpmedel inom bl. a. området strategisk planering.

Migration - möjlig...?

Här framhåller YG 'Reverse Engineering' som ett alternativ. Även om denna teknik ännu är i sin linda tror YG att den kommer att spela en stor roll i framtida migrations-projekt. Men, säger YG, förmodligen kommer inte många dataavdelningar inom företagen att kunna hantera denna sortens frågor, varför marknaden för fristående konverteringsexperten/firmor kommer att bli mycket luckrativ.

Slutligen vill rapporten framhålla några saker:

- leverantörer med produkter inom minidatorområdet, tex DEC, Oracle och Sybase, ligger långt framme när det gäller att möta de krav kunderna ställer avseende DDBMS.
- användarna av minidatorer har det följdaktligen bättre förspänt än stordatoranvändarna avseende en framtida migration till DDBMS
- minidatoranvändare har förmodligen färre problem på nätverkssidan än deras stordatorkollegor.

Minidatorbaserade nätverk är ofta mycket mer komplexa än stordatorbaserade nätverk, vilka ofta har formen av ekrar i ett hjul och där stordatorn utgör navet.

Leverantörspositionering

Sista delen av YG-rapporten går igenom de leverantörer som finns på marknaden, och försöker peka ut vinnare och förlorare i det framtida DDBMS-loppet. Här följer en MYCKET summarisk sammanställning:

ADR (Computer Associates CA)	DBMS-produkt: DATACOM/DB. Knappast.
Computer Corp of America (CCA)	DBMS-produkt: Model 204 Knappast. CCA har inte helt köpt idén med relationsdatabaser, och kommer bl a därför att bara spela en mindre roll inom DBMS.
Cincom	DBMS-produkt: Supra En bra produkt som tekniskt kommer att utrustas för att klara en DDBMS-omgivning. Dock ligger Cincom lågt och avvaktar IBM's steg.
Cullinet DBMS-produkt	Enterprise: DB (f.d. IDMS/R) En av de bästa produktstrategierna på marknaden. YG håller sig dock avvaktande till dess produkt(er) som lever upp till dessa strategier sett dagens ljus.
Digital Equipment DEC	DBMS-produkt: Rdb/VMS Bra produkt och en strategi som väl svarar mot marknads behov. DEC kommer att vara med i loppet.
Groupe BULL	DBMS-produkt: Oracle (!) Obs inget eget DBMS. BULLs attityd till DDBMS är ganska negativ, varför dom får dåliga odds.
IBM	DBMS-produkt: DB2, OS/2 Ext Ed OS/400 SAA utlovar vissa saker inom DDBMS, annonseringen av DB2 ver 2 rel 2 innehåller en del 'distribuerat'. Självfallet kommer IBM att röra sig åt DDBMS-hållet, men man kommer

att fördröja så länge som möjligt, delvis av rent marknadsmässiga orsaker. IBM vill ha ut mer försäljning av nuvarande version av DB2 innan det är dags att satsa på DDBMS.

(Principen känns igen/ö.a.)
Det ständigt fördröjda 'repositoryt' påverkar IBM's satsning på DDBMS negativt.

ICL

DBMS-produkt: Oracle
RTI/Ingres

Inget eget DBMS som synes.
Vilken sw-leverantör som kommer att lyckas bäst inom ICL-världen när det gäller DDBMS vågar YG inte sia om. Inget kommer att hända snabbt i alla fall.

Olivetti

DBMS-produkt: Oracle
RTI/Ingres. Inget eget DBMS.
YG tror att Olivetti kommer att var bland de första hw-leverantörerna som implementerar DDBMS.

Oracle Corp.

DBMS-produkt: Oracle
I frontlinjen avseende DDBMS. Oracle har vissa tekniska problem, bl a har deras 'friserade' benchmarks kritiserats mycket hårt, men trots detta är produkten mycket bra enligt YG. Oracles utvecklingshjälpmedel har rönt mycket beröm.

Siemens

DBMS-produkt: SESAM (!) UDS
Förutom dessa egna produkter kommer Siemens att annonsera support för Oracle på sina stordatorer tror YG.
Det faktum att Siemens egna DBMS saknar SQL-stöd kommer att innebära problem för dessa DBMS i framtiden.

Software AG

DBMS-produkt: ADABAS
SAG anser inte att en distribuerad databas måste vara relational! Detta kommer att ligga SAG i fatet om de inte lyckas omvända världen till sin ståndpunkt. Trots detta en bra produkt, och en stor kundbas som säkrar framtida DDBMS-satsningar inom ADABAS.

Sybase

DBMS-produkt: SQL Server
Senaste tillskottet på DBMS-himlen och ett av de mest intressanta. Sybase grundläggande hw-plattform är DEC, men har också genom avtal med Microsoft och Ashton-Tate etablerat sig på PC-sidan. Största framtidsutsikterna på PC-sidan enligt YG.

Slutsatser

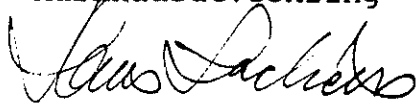
YG anser att, på kort sikt i alla fall, alla försök till att etablera en DDBMS-miljö skall övervägas MYCKET noga.

Av hw-leverantörerna förväntas DEC leverera DDBMS snabbast. Av oberoende vendors anses Cullinet ha den bästa produktstrategin, men ingen produkt som svarar upp mot denna strategi.

CCA och Software AG kämpar i uppförsbacke pga att de inte har någon relationsdatabas.

Göteborg den 27 juni 1989

VOLVO DATA AB
Marknadsutveckling



Hans Lackéus

Distribution

- 33 Kostnadsställeansvariga
- 101 AU-ansvariga
- 102 Metod- och teknikansvariga
- 103 CAE-kontakter

FORSBERG INGE
02510
DA2