

# Open Source databaser: Modne for større oppgaver?

*“Databaser? Det tar vi alvorlig. Der ligger organisasjonens viktigste juveler. Om noe er virksomhetskritisk, så er det databasene.”* Sitatet er ikke autentisk, men en oppsummering av responsen fra 8 av 10 IT-ledere på spørsmål om hvordan de vurderer databasenes rolle i organisasjonen.

## Hva stoler vi på og hvorfor?

Responsen er som ventet, men likevel interessant. Er det riktig at databasene inneholder virksomhetenes mest verdifulle data? Hva er det de typisk inneholder? Er regnskapsdata, økonomisk informasjon, produksjonsdata, ordrer, bestillinger og så videre virksomhetskritiske? Uten tvil. Men er det riktig at databasene huser organisasjonens viktigste juveler? Det kommer naturligvis an på virksomheten og ikke minst på øynene som ser. De fleste organisasjoner arkiverer applikasjonsdata i databaser, mens dokumenter – brev, fakser, regneark, maler, tegninger, produktinformasjon, kontrakter og så videre – lagres i filsystemer. Ved nærmere ettersyn kommer det gjerne for en dag at vel så mye av organisasjonens verdier er å finne i vanlige filer som i databaser.

Dette er naturlig. Selv om databaseleverandørene har arbeidet hardt og lenge for å få oss til å flytte all lagring over til databaser, er filsystemer fortsatt den foretrukne lagringsmekanisme for alle tenkelige typer data. Årsaken er like mye tradisjon som oppfatningen av at dette er sikkert nok. Dessuten betraktes filsystemer – med rette – som enklere å forholde seg til enn databaser.

## Databaser og filsystemer

Filsystemene stiller vi sjelden spørsmålstegn ved. Deres pålitelighet er forankret i erfaring, og ingen andre enn driftspersonell og systemarkitekter bryr seg om hvor de befinner seg – lokalt, på en tjener, på en NAS-løsning eller kanskje et SAN. Heller ikke når vi får vite at filsystemene er implementert på en Open Source plattform finner vi grunn til å reagere.

I teknisk forstand har databaser og filsystemer tallrike likhetspunkter. De er mekanismer for lagring, organisering og manipulering av dataobjekter på et masselager av vilkårlig type. Vår avhengighet av begge er total. Likevel møter forslag om bruk av Open Source databaser gjerne hevede øyenbryn og skeptisk hoderystelse. Hvorfor denne forskjellsbehandlingen? Er Open Source databaser mindre å stole på og dermed farligere enn filsystemer?

Mange vil oppfatte spørsmålet som underlig. Open Source databaser (OSDB) – hva er det? Mens de fleste har fått et forhold til eller en oppfatning av Open Source gjennom operativsystemer (Linux), Web-tje-

nerer (Apache), epost-programmer (Sendmail, Postfix) og nettlesere (Mozilla), og tar deres funksjon og kvalitet for gitt, er databaser et relativt ubeskrevet blad i en slik sammenheng. Noen har hørt om MySQL og aner at det er et databasesystem, men kunnskapen stopper der.

Denne situasjonen passer naturligvis de tradisjonelle databaseleverandørene utmerket. De har all interesse av å bevare status quo så lenge som mulig, og har lite positivt å si om de nye konkurrentene. Deres situasjon er helt analog med forholdet mellom Microsoft og Linux, og håndteres på tilsvarende måte – med FUD (*Fear, Uncertainty, Doubt*) som viktig ingrediens.

Siden det generelle kunnskapsnivået om OSDB-systemer er såvidt lavt, kommer utbredelsen av disse produktene – med PostgreSQL og MySQL i spissen – gjerne som en overraskelse. De brukes i tusenvis av virksomhetskritiske anvendelser over hele verden, og viser større og raskere vekst enn kommersielle alternativer. Videre tar 'nykommerne' raskt inn på de kommersielle alternativene med hensyn til funksjonalitet. Følgende fakta taler for seg:

- ✓ Bruken av MySQL vokste med 30% i 2003 ifølge en undersøkelse fra amerikanske Evans Data. I samme periode vokste MS SQL Server og Access med 6%.
- ✓ En ny versjon av PostgreSQL (7.5), som kommer i juni måned, vil for første gang støtte Windows. I tillegg til en lang rekke ytelsesmessige forbedringer, vil den nye versjonen også ha støtte for såkalt *two phase commit*, som er kritisk i forbindelse med transaksjons-prosessering.
- ✓ MySQL kommer i inneværende måned med en klynge-utgave som tar vare på kritiske anvendelsers krav til oppetid.
- ✓ Firebird,<sup>4</sup> hvis opprinnelse er 'gode, gamle' Interbase fra Borland, kommer i ny utgave i løpet av sommeren, og hevdes å være klar for virksomhetskritiske oppgaver.

Og dette er kun toppen av isfjellet. "I løpet av de to neste årene vil Open Source Databaser oppnå teknisk likeverdighet med sine tradisjonelle (kommersielle) konkurrenter på de fleste nøkkelområder" skriver AMR Research i en fersk rapport, og fortsetter: "Dette vil skape et vell av nye muligheter for brukermiljøer som vil senke kostnadene knyttet til utrulling av nye applikasjoner."

### **Først Linux og Apache, så OSDB**

Igjen er likhetene med Linux påfallende. For 5-6 år siden hadde mange hørt om Linux, men få visste hva det egentlig var og enda færre hvilken utbredelse systemet hadde. Situasjonen er eksakt den samme for OSDB i dag, med ett viktig unntak: Open Source er akseptert som konsept og utviklingsmodell – med den følge at avstanden fra markedsignorans til markedsaksept blir vesentlig kortere. Derfor er 2-års horisonten fra AMR Research rimelig.

<sup>4</sup> Firebird har også vært navnet på den selvstendige nettleseren utviklet fra og ved siden av Mozilla. I februar i år bestemte imidlertid Mozilla-teamet seg for å omdøpe sin Firebird nettleser til Firefox.

Også argumentene markedet bruker for å velge inn de nye alternativene er som hentet fra Linux-historien:

- ✓ Lavere kostnader
- ✓ Høyere fleksibilitet – ikke minst i forhold til plattformer
- ✓ Større effektivitet
- ✓ Enklere drift
- ✓ Bred utviklerstøtte
- ✓ Høy support-kvalitet

Faktorene er interessant nok også overlappende med hva de tradisjonelle (kommersielle) database-leverandørene bruker – ikke minst for å distingvere seg fra Open Source. Et viktig element i forklaringen – og ikke minst i legitimeringen av produktene – er at flere av OSDB-produktene er halvkommerielle. De kan lastes ned gratis – og vi får den støtte og hjelp Open Source-miljøene er villige til og interesserte i å gi. Alternativt kan vi kjøpe kommersielle lisenser av det samme produktet, med tilleggstjenester – og i noen tilfeller tilleggsfunksjonalitet. Prinsipielt er dette det samme som vi får fra RedHat, SuSE og andre Linux-leverandører. Forskjellen i forhold til tradisjonelle databasealternativer er at både lisenser og support-avtaler har dramatisk lavere pris. Dessuten har praktisk erfaring nok en gang vist at supporten er både raskere og av høyere kvalitet enn det som er vanlig i segmentet.

Også måten produktene kommer inn på markedet på, er analog med Linux. Miljøer med egen kompetanse går foran, etablerer systemer og løsninger som svarer til målsettingene og deretter blir referanse-installasjoner som brukes både av pressen og av andre miljøer. Massefarten tiltar, og i løpet av noen måneder eller år flytter alternativene seg fra å være esoteriske til å bli først rollemodeller og deretter selvfølgelige.

Dit er vi ikke kommet med OSDB enda, men antall historier om store og små vellykkede prosjekter øker, og i takt med dem interessen – ikke minst fra presse og analytikere. Allerede nå merker de tradisjonelle databaseleverandørene et begynnende press på priser og betingelser. “Vi skulle utvide, men kunne ikke leve med betingelsene fra leverandør XYZ. Derfor valgte vi MySQL” er en beskrivelse som stadig dukker opp.

### **LAMPer og pakker**

En annen faktor som drar i samme retning, er at databaser i voksende grad blir en del av applikasjons- eller verktøy-pakker. Trenden er synlig på tvers av IT-segmentet, med Microsoft som drivkraft. Selskapet har i løpet av de 3-4 siste årene ‘bundlet’ sin SQL Server med det ene produktet etter det andre, i noen tilfeller med god grunn, og minst like ofte for å fremtvinge en tilstedeværelse på databasesiden. Tilsvarende bundling gjøres på Open Source siden, men med en kostnad og en grad av valgfrihet som unngår å forårsake de negative reaksjonene Microsofts ‘tvangsinnføring’ har skapt.

LAMP er en typisk slik bundling – en relativt fersk betegnelse på en komplett Web-tjener/e-handel/portal-pakke bestående av **L**inux (OS), **A**pache (Web-tjener med tilbehør), **M**ySQL (database) og **P**HP/Python/Perl (utviklings/scriptspråk). Hvilket script-språk som velges er naturligvis valgfritt, og det er en heldig tilfeldighet at alle har P som forbokstav. I mange tilfeller brukes en kombinasjon av to eller alle tre.

En annen og ferskere variant over samme tema – og konkurrent – er 'Brighther LAMP', med utgangspunkt i PostgreSQL-databasen, der M står for *Middleware* og P for PostgreSQL. Mellomvaren inneholder typisk en Java applikasjonsserver (for eksempel *jboss*, som vi nevner avslutningsvis) og funksjoner for meldingsformidling, og reflekterer blant annet den oppfatning at PostgreSQL er bedre egnet enn MySQL for store/tunge databaser og ditto transaksjonsvolumer. Hvorvidt denne oppfatningen er korrekt er en annen sak, og kommer nok mer an på typen transaksjoner enn på volum og størrelse.

## Nykommere?

Hvem er de, disse nykommerne og hva er de gode for? Felles for dem er at de knapt kan kalles nykommere. **PostgreSQL** [www.postgresql.org] er et godt eksempel, med aner tilbake til Project Ingres og universitetet i Berkeley i 1972. Forskningsprosjektet Ingres la grunnlaget for moderne relasjonsdatabaser og en tilgjengelig Open Source database på 80-tallet. Senere ble den kommersialisert og deretter overtatt av Computer Associates (CA) tidlig på nittitallet. Studenter og forskere med fartstid fra Project Ingres var involvert i etablering og utvikling av samtlige store databaseprosjekter på 70- og 80-tallet, inklusive Oracle, Sybase/SQL Server, DB 2, Informix, Unify og flere. Den akademiske/åpne delen av Ingres ble omdøpt til Postgres da Ingres ble kommersiell i 1986. Etter en omskriving til SQL-standarden fikk navnet sin siste justering.

Som navnet forteller, har **Berkeley DB** [www.sleepycat.com] samme geografiske opprinnelse som PostgreSQL, og aner tilbake til tidlig på 1980-tallet. Når Berkeley DB (BDB) kun unntaksvis nevnes i samme åndedrag som MySQL og PostgreSQL, er årsaken at BDB verken er en relasjonsdatabase eller har SQL-grensesnitt. Systemet er laget for høy-effektiv lagring hvor en relasjonsdatabase er unødig komplisert, og har spesielt stor popularitet (og markedsandel) i industrielle (*embedded*) systemer. Videre har de aller fleste Linux- og Unix-systemer minst et titalls anvendelser av BDB – langt ute av syne for de fleste av oss.

**MySQL** [www.mysql.com] er mest kjent og utbredt av Open Source databasene. Den har sin opprinnelse fra Sverige i 1995, hvor selskapet med samme navn fortsatt har tilhold. Produktets popularitet har nødvendigvis etablering i USA, for support såvel som for videreutvikling, og selskapet har totalt 100 ansatte 'spredt over hele kloden', som selskapet selv sier. I tillegg til produktets tekniske egenskaper, har produktets alder og tilblivelse hatt stor betydning for suksessen. MySQL har 'ridd' Open Source bølgen fra første dag, og har lyktes på en bemerkelsesverdig måte, både kommersielt og teknisk. Produktets

mest profilerte referanser er det verdensomspennende flybooking-systemet Sabre og det Internett-baserte reisebyrået Travelocity. MySQL AB overtok i 2003 et utviklingsprosjekt fra Ericsson som blant annet inneholdt klynge-teknologi. Denne teknologien blir nå inkorporert i MySQL, og bringer produktet over til et nytt nivå i virksomhetskritisk sammenheng.

**Firebird** [www.firebird.org] er på overflaten den ferskeste av de kjente Open Source databasene, men har like fullt historie tilbake til 1980. Mest kjent under navnet Interbase, vandret produktet fra leverandør til leverandør i PC-segmentet på slutten av 80- og begynnelsen av 90-tallet, for til slutt å ende opp hos Borland. Utviklingen stoppet etter hvert helt opp, men ble revitalisert da Borland frigjorde kildekoden for snart 2 år siden. Firebird er en full relasjonsdatabase med SQL-grensesnitt som dekker Windows- og Linux-plattformene. Den har spesielt stor utbredelse i Tyskland, der 'Firebird Verdenskonferanse' avholdes for 2. gang i mai måned.<sup>5</sup>

### Plattform-agnostisk

I motsetning til hva mange fortsatt synes å tro, har verken OSDB-systemer eller andre Open Source produkter noen spesiell tilknytning til Linux. Tvert imot, er det fristende å si: De fleste av dem har sin opprinnelse på andre plattformer, og har ingen preferanser i så henseende. På OSDB-siden har vi allerede nevnt at PostgreSQL i disse dager foreligger for Windows, etter å ha trått sine barnesko i Unix-miljø og vokst opp med Unix og Linux. Interbase/Firebird kommer fra Windows, og har fortsatt sin hovedbase der. Det samme gjelder for MS Office-konkurrenten OpenOffice.org, som vi har diskutert tidligere i Mellvik-Rapporten.

Denne plattform-agnostiske filosofien har i seg selv positive konsekvenser – i første omgang for innstillingen i markedet, og dernest for fokuseringen på optimale i stedet for tilgjengelige plattformer. For eksempel kjører noen av de største MySQL-installasjonene i verden på Apple-tjenere (XServe) med Unix-baserte Mac OS X, mens de største finansmiljøene i verden – i New York og Chicago – i løpet av de siste årene har konvertert fra Unix til Linux som fundament for sine systemer.

## Konklusjon

Den kraftige fremgangen for OSDB-systemer kan ikke unngå å få konsekvenser for IT-utviklingen i sin alminnelighet – på samme måten som Linux har fått det. Utbredelsen og veksten har forlenget passert det som kan kalles støynivå i markedet, og produktenes egenskaper nærmer seg raskt paritet med kommersielle produkter på de fleste områder. Det betyr ikke at de tradisjonelle leverandørene er truet på livet, like lite som Microsoft trues på livet av Linux. Her er det plass for om ikke alle, så i alle fall mange.

<sup>5</sup> Også de andre produktene har tilsvarende 'verdenskonferanser' – for MySQLs vedkommende går den av stabelen i Florida rett etter påske.

Den umiddelbare konsekvensen er for det første korreksjoner av prisnivå som følge av konkurransen, og for det andre en voksende bevissthet rundt valg og bruk av optimale verktøy til riktig oppgave. Tilbakemeldingene fra tallrike brukermiljøer understreker nettopp dette forholdet. De har ønsket forenkling, og finner at OSDB-systemer kan overta en rekke oppgaver som tidligere ble ivaretatt av kommersielle systemer med langt høyere kompleksitet og ressurskrav.

Slike signaler bør vekke betydelig oppmerksomhet hos tradisjonelle database-leverandører, som anser seg for å levere drifts- og styringsverktøy i toppklasse. Hva som er toppklasse er i høyeste grad avhengig av øynene som ser. En trailer er neppe optimalt arbeidsverktøy verken for en maler eller en rørlegger.

### Søkefaktoren

En annen faktor som påvirker bildet utenfra, er fremveksten av avanserte, universelle søkemotorer – med norske Fast Search and Transfer som nærliggende eksponent – som søker like godt i alle datatyper, strukturerte og ustrukturerte. At disse produktene overgår database-leverandørenes egne, spesialsyddde søkemekanismer med hensyn til effektivitet og bruk, betyr at et viktig argument for valg av slike systemer faller bort. Datatilgangen er ikke lenger kontrollert av databasелеverandørene, og valg av andre alternativer kan være positivt, også for søkehastighet og tilgjengelighet.

### Trendeffekter

Trendeffekten av at OSDB-systemer i økende grad velges selv til virksomhetskritiske anvendelser, er interessant. At databasemarkedet dermed gjennomgår en forandring, er innlysende. Videre bidrar utviklingen til en ytterligere styrking av Open Source-alternativer generelt. Linux-syndromet – *The Open Source Express*, som enkelte analytikere kaller det – ruller videre med voksende styrke, og spørsmålet er ikke lenger om Linux eller OSDB vil lykkes, men i hvilken grad og hva som blir det neste.

Svaret aner vi når vi ser at både Java-baserte applikasjons-tjenere (med jboss [www.jboss.org] som primær eksponent) og ERP/CRM-systemer (for eksempel Compiere<sup>6</sup>) i løpet av de siste månedene har fått stor og positiv oppmerksomhet – og spennende referanser. Begge tilhører Open Source kategorien, har tallrike fellestrekk med OSDB-produktene i utviklingsmessig og supportmessig forstand.

I løpet av året er det med andre ord sannsynlig at en ny bølge avtegner seg, mens Linux-bølgen ruller – og vokser – videre, med OSDB hakk i hæl. Dermed blir det ekstra spennende å høre hva Oracle har å si om Linux-utviklingen – i juni-utgaven av Mellvik-Rapporten. Følg med! ■

<sup>6</sup> Compiere [www.compiere.org] er database-agnostisk, og har i følge utviklingsmiljøet støtt på hemmende problemer eller begrensninger i de fleste databasesystemer, Open Source og kommersielle.