

# Lagringssystemer: Snarvei til kaos?

*Masselager er billig. Ufattelig billig – så billig at vi uten å tenke oss om slenger inn et par hundre gigabytes ekstra dersom et kapasitetsproblem oppstår. Så går verden videre og vi tenker ikke mer på den siden – før neste problem dukker opp.*

Situasjonen har mye til felles med tilsvarende på nettverkssiden: Prisene på båndbredde har falt så raskt og så lenge at å kaste mer kapasitet på problemene uten unntak har vært enklere enn å gripe fatt i de egentlige utfordringene. Å kurere symptomene i stedet for sykdommen er blitt regelen i stedet for unntaket.

## Kostbart og naivt kaos

I begge tilfeller nærmer vi oss imidlertid en grense for hva som er mulig – ikke i kapasitetsmessig forstand, men i praksis. Følgeskadene av å feie problemene under teppet i år etter år kommer smått om senn for en dag. For masselager og lagringssystemer betyr det – blant annet – at brukerne har samlet så mye rot at det stadig oftere er vanskelig eller umulig å finne de riktige data når de trengs. Hyppigheten med hvilken det tas beslutninger basert på gamle eller feilaktige data er faretruende stor. Situasjonen er analog med et overfylt og kaotisk skrivebord – eller en garasje der bilen ikke har fått plass på mange år. Mengden av verdiløst skrot er så stor at det ikke lar seg gjøre å finne det som har verdi med rimelig innsats.

Myten om viktigheten av å beskytte medarbeidernes verdifulle tid slår med andre ord tilbake. Tallrike organisasjoner har unnlatt å stille krav til sine brukere så lenge at det ikke lenger er mulig verken å skape orden eller finne frem i kaoset. Her må det annen lut til.

### Gi brukerne ansvar – for seg selv

Vi aksepterer lett viktigheten av at alle kjører på høyre side av veien, går på fortauet når dette finnes, eller setter oss på tildelte plasser på kino og i flyet. Reglene er velkjent og akseptert, og vi innser at alternativet – ingen regler – er det samme som kaos. Likeledes holder vi en viss orden rundt oss – på kontoret, hjemme, i bilen – riktignok i varierende grad etter smak og behov, men en viss orden er vi alle avhengige av.

Det finnes ingen rimelig grunn til at situasjonen skulle være annerledes i IT-sammenheng. Vi er i stand til å generere vilkårlige datamengder på minimal tid, og vet i det fleste tilfeller hva som bør lagres og hva som bør slettes. Likeledes har vi et forhold til verdien av de data vi omgir oss med – om de er historiske (kan arkiveres), nyttige i nær fremtid (tilgjengelige) eller aktive (under arbeid). Denne kunnskapen er ferskvare. Det tar kort tid før vi glemmer hva som er hva – med mindre vi sørger for merking av dataene – via navn, metadata eller på annen måte. Derfor må det gjøres til den enkelte medarbeiders ansvar å holde

orden, i like stor grad på masselagerområder som i fysisk forstand. Årsaken til at brukere flest vegrer seg for slike oppgaver er ikke først og fremst at oppgavene er ulystbetonte, men at etablert praksis er å la det skure.

Å stille brukerne til rette for situasjonen blir imidlertid som å skyte pianisten fordi vi ikke liker musikkstykket. Brukerne gjør som de får beskjed om, og får de ingen beskjed, gjør de som det passer dem. Stiller IT-organisasjonen mer lagringsplass til disposisjon hver gang fyllingsgraden blir for stor, uten å stille krav tilbake til den enkelte, gir resultatet seg selv. Den primære mangelen er ofte en kombinasjon av regler, rutiner og opplæring – med andre ord: Motivasjon.

## Tid for erkjennelse – av virkeligheten

Utviklingen har forårsaket betydelig innsats fra både leverandør- og kundesiden i markedet de siste to årene, hvilket har frembrakt mye interessant informasjon om den faktiske situasjonen. Ikke minst har kostnadsanalyser bidratt til å bringe problemstillingen inn på toppnivå i mange organisasjoner. Blant de mest generelle av forholdene som er avdekket, finner vi:

- ✓ Det eksplosive behovet for mer lagringskapasitet står ikke i forhold til veksten i data med reell nytteverdi.
- ✓ Kostnadene knyttet til tidstap i forbindelse med søking etter data samt anvendelse av feilaktige eller utrangerte datasett, er vanskelig å kvantifisere, men uten tvil betydelige.
- ✓ Kostnader og komplikasjoner knyttet til sikkerhetskopiering vokser raskt, ved siden av at etablerte rutiner for tilbakerulling av data ikke lenger fungerer. Det lar seg simpelthen ikke gjøre med dagens typiske backup-verktøy å rulle de enorme datamengdene tilbake til aktiv drift på rimelig tid.
- ✓ Mangelen på regler og rutiner skaper en umulig situasjon for driftsavdelingen som ikke har mulighet til å skille mellom data av ulik kvalitet. I mange tilfeller er det både enklere og raskere å regenerere enorme datasett enn å hente dem inn fra en sikkerhetskopi.<sup>6</sup>
- ✓ Manglende policy for bruk av masselager er blitt et juridisk problem for mange organisasjoner. At brukerne lagrer pornografi, stjålet materiale, kopibeskyttet materiale (video, musikk etc.) blir organisasjonens ansvar dersom slike policies ikke finnes.<sup>7</sup>

Spesielt siste punkt har tatt en rekke organisasjoner på sengen, ikke minst her hjemme. Påminnelsene burde være mer enn nok til umiddelbart å sende samtlige våkne IT-avdelinger til møterommet for å snekre sammen grunnleggende regler for bruk av alle IT-ressurser – fra mobiltelefoner til lagringskapasitet.

<sup>6</sup> Fremtredende eksempler er oljebransjen og video/animasjonsbransjen.

<sup>7</sup> Problemstillingene gjelder ikke bare lagring, men også bruk av andre ressurser, for eksempel Internett-forbindelsen. Det som skal til er regler som for det første er kjent av alle brukere og for det andre rimelige i forhold til virksomhetens og den enkeltes behov.

### Fra sløsing til kontroll

Også en rekke andre interessante observasjoner er kommet frem i undersøkelsene vi refererte til ovenfor. For eksempel bekrefter de at den årlige behovsveksten er i størrelsesorden 40-50%, med omtrent tilsvarende tall for bemanningssituasjonen. Videre foreligger det en del interessante tall med hensyn til effekten av optimaliserende tiltak:

- ✓ Konsolidering på lagringssiden, fra generelle systemer (NT, W2k, Novell, Unix/Linux/NFS) til lagringsnettverk, gir en kortsiktig kostnadsreduksjon på minst 15%.
- ✓ Oppryddingskampanjer – som gjerne arrangeres avdelingsvis eller for hele organisasjoner, med innlagte konkurransemessige elementer – gir i de fleste tilfeller store uttelling, ofte i størrelsesorden 25-50% redusert forbruk per bruker.
- ✓ Det er nødvendig å skille lagringsrelaterte kostnader fra øvrige IT-kostnader for å få kunnskap om situasjonen og utviklingen, og dermed motivasjon for tiltak.

De kontrollerende tiltakene som nå synes påkrevet, var interessant nok alminnelige for 15 år siden. Kvotemekanismer som begrenset den enkelte brukers uttak av CPU-tid, hukommelse, masselager og andre ressurser, var en forutsetning for 70- og 80-tallets flerbruker-systemer, og finnes fortsatt i dagens stormaskiner. Deres relevans for brukere flest ble borte under PC-revolusjonen, der hver enkelt for en periode ble ansvarlige for sine egne data. Etter at kaoset var et faktum, dukket funksjonaliteten opp igjen da Microsoft for alvor kastet seg over tjenermarkedet – med Windows NT 4.0.

Utfordringen vi møter i dag, er imidlertid en helt annen enn tilsvarende for over 10 år siden. Datidens ressursstyrings-mekanismer og -systemer hadde til oppgave å fordele, begrense og optimalisere bruken av kostbare fysiske ressurser. I dag er det andre ressurser som direkte eller indirekte koster penger og derfor må optimaliseres:

- ✓ Tid – for brukere som ikke finner det de trenger når de trenger det, og driftspersonale som administrerer og sikkerhetskopierer verdiløse data.
- ✓ Kvalitet – på dataene som brukes: Hvilke datasett er oppdaterte og hvor befinner de seg?
- ✓ Driftssikkerhet – kravene til oppetid forutsetter at vi vet hvilke data som er viktige når ulykken er ute.
- ✓ Ansvar – fordi manglende regler eksponerer organisasjonen i juridisk forstand for medarbeidernes overlagte eller uaktsomme handlinger.

### Hvor er verktøyene?

Det er innlysende at vi trenger verktøy for å få situasjonen under kontroll. Videre – i og med at vi nå er langt mer opptatt av data enn av kapasitet, kan ikke tradisjonelle SRM – *Storage Resource Management* – systemer dekke våre behov. Ikke desto mindre er det fortsatt res-

surs-administrasjon som står i fokus hos de fleste verktøyene som finnes, og som typisk leveres av hardware-leverandørene i segmentet.

Det betyr ikke at slike systemer er unyttige. For å fremtvinge et minimum av orden og disiplin, har organisasjoner flest behov for å sette tak på hvor store lagringsressurser den enkelte bruker kan legge beslag på. Poenget er at vi ønsker mer fra slike verktøy. Vi trenger hjelp til kategorisering av data, til oppryddinger, intelligent rapportering. Videre vil vi i takt med veksten i nettverkenes størrelse og omfang ha voksende nytte av å kunne identifisere bruksmønstre, slik at vi på et tidlig tidspunkt kan se hvor flaskehalsen vil oppstå og komme i forkant av utviklingen.

### **Fokus på informasjon, ikke lagringsplass**

Dit er det langt frem, tilstrekkelig langt til at vi har god tid til å omfokusere våre prioriteringer mens leverandørsiden gjør det samme – og får frem produkter som er tilpasset virkeligheten i stedet for historien.

I mellomtiden benytter vi det som finnes av verktøy – utilstrekkeligheten til tross, fordi de bringer oss et betydelig stykke på vei. Erkjennelsen av at det er informasjon, ikke lagringsressurser som fordrer vår oppmerksomhet, eliminerer som nevnt ikke behovet for styring av de ressursene som allerede finnes.

Denne erkjennelsen er første trinn på veien mot et nytt og mer virkelighetstilpasset forhold til lagring kontra informasjon. Neste trinn er en kvantifisering av informasjonens verdi for virksomheten. En slik kvantifisering lar seg ikke nødvendigvis oversette til kroner og øre, men vil alltid kunne rangeres i forhold til andre ressurser og verdier i virksomheten. Forandringen må innarbeides i eksisterende eller nye strategier for organisasjonen og for dens IT-anvendelse.

## **Konklusjon**

Vårt hovedpoeng i denne sammenhengen er at en omfokusering fra kapasitet til innhold er nødvendig for å kunne håndtere utfordringer på datalagringssiden. Mens det for 10 år siden var rimelig å behandle alle data under én kam – fordi totalmengden tillot en slik forenkling, er situasjonen i dag en annen.

Vi har siden 'tidens morgen' hatt omgang med ulike informasjonstyper i det daglige, og håndtert dem i henhold til realverdi med den største selvfølge: Aviser, magasiner, bøker, notater og så videre. Den samme selvfølgen må inn i våre datalagringssystemer: Ferskvareinformasjon (for eksempel status for et prosjekt) behandles forskjellig fra historisk informasjon (prosjekttegninger, sluttrapport, økonomi, osv.).

Denne implisitte kategoriseringen av informasjon må inn i våre systemer og rutiner, for eksempel i tre nivåer:

- ✓ Høy prioritet: Kritisk for organisasjonens virke fra time til time.
- ✓ Middels: Nødvendig referanseinformasjon som vi kan klare oss uten i noen timer.

- ✓ Lav: Informasjon som må finnes, men som vi kan klare oss uten i flere dager, kanskje uker.

En av nøklene for å få en slik kategorisering til å fungere i praksis er å definere regler for hvordan omklassifiseringen foregår. Dagens informasjon kan være historisk eller i alle fall mindre viktig i morgen. Likeledes kan det være behov for å skille mellom dynamisk og statisk informasjon innenfor hvert nivå. Organisasjonen kan være totalt avhengig av at varekatalogen er tilgjengelig til enhver tid, men den forandrer seg ikke daglig. Blant annet sikkerhetskoperings-rutiner må reflektere slike forhold.

*Lagringssystemene eksisterer ikke for sin egen del, men for å ta vare på og gi tilgang til informasjon – som representerer den egentlige verdien. Det faktum at få organisasjoner har gjort jobben med å kvantifisere verdien av sine data, er årsaken til at vi fortsatt ofte henfaller til en diskusjon om kapasitet når utfordringer knyttet til lagringssystemer skal diskuteres.*

Veien frem mot å bringe situasjonen under kontroll, består – ved siden av omfokuseringen fra lagring til informasjon – av følgende punkter:

- ✓ Etablering av enkle og velfunderte regler for bruk (policy) – en grunnleggende forutsetning for å få den eksplosive utviklingen under kontroll – og ikke minst for å ta vare på juridiske ansvarsforhold. Brukerne skal opplæres og tilbys verktøy som gir grunnlag for å praktisere gjeldende policy.
- ✓ Restriksjoner på bruken av lagringsressurser er en nødvendig oppfølging av en policy i de fleste organisasjoner.
- ✓ Kategoriseringen av informasjonstyper kan begynne umiddelbart, og vil få konsekvenser for blant annet sikkerhetskoperings-prosedyrer og policy overfor brukerne.

Vi dro innledningsvis frem likheten i utvikling mellom lagringskapasitet og nettverks-båndbredde. Det er knapt overraskende at tiltakene for å få situasjonen under kontroll har mange fellestrekk. I begge tilfeller må det gjøres bevisste og godt publiserte prioriteringer, og brukerne må aktiviseres med opplæring og forståelse av problemer, løsninger og prioriteringer. Det er påtagelig hvilke negative effekter som følger i kjølvannet av forandringer og omdisponeringer som få eller ingen forstår hensikten med. På samme måte er det utrolig hvilke positive effekter forståelse og engasjement kan gi. ■

#### Mer om lagring i Mellvik-Rapporten

I månedene fremover vil vi følge opp fokuseringen på lagringsproblematikk – med praktisk orienterte artikler om sikkerhetskopering, policy og lagringsarkitekturer.