

Lagring: Prispres og nye arkitekturer

Gjennomgangen av grunnleggende lagringsteknologi i forrige utgave etterlot et åpent spørsmål: Hva skal vi velge – NAS eller SAN? Spørsmålet har vært innom de fleste IT-lederes bord i løpet av de siste 3 årene, og fasitsvaret mangler fortsatt. Likeledes har faktorene som må veies inn i svaret, forandret seg i løpet av perioden, og er blant elementene vi skal innom i denne oppfølgingsartikkelen.

SAN kontra NAS: Ja takk, begge deler?

Utviklingen er ikke til å ta feil av. Mens NAS var anerkjent som den optimale løsningen – spesielt for mellomstore miljøer – mot slutten av 90-tallet og frem til 2002, har oppmerksomheten de to siste årene skiftet over mot SAN. Årsaken er delvis at SAN-prisene har falt betydelig – i takt med at nye aktører har forsøkt seg i segmentet. Videre har fokuseringen på stabilitet og tilgjengelighet i markedet tiltatt betraktelig, og ikke minst har behovet for lagringskapasitet vokst eksponensielt. Spesielt det siste har vært utslagsgivende fordi NAS-løsninger ikke har like gode skaleringsegenskaper som SAN.

Samtidig har SAN-teknologi vært både misforstått og misbrukt. Den lovede påliteligheten og robustheten har i mange tilfeller uteblitt på grunn av feilkonfigurasjoner og lite optimale utstyrvalg – forårsaket av manglende kompetanse både på leverandør- og kundesiden.

I kampens hete er det ofte blitt glemt at uten 'arkiveringsmekanismer' – som filsystemene representerer – kan ingen ting lagres. Sagt på en annen måte: Filtjenere trenger vi uansett om vi har lokal lagring (DAS) eller lagringsnettverk (SAN). Spørsmålet om hvilke ressurser som er optimale for oppgaven, forsvinner ikke om vi introduserer et SAN i stedet for lokale diskere.

Store utfordringer for NAS

Når NAS-produkter nå har kommet tilbake i fokus, kan vi observere at produktene har gjennomgått en betydelig oppgradering i perioden. Konsolidering på leverandørsiden, med Network Appliances oppkjøp av Spinnaker Networks som eksempel, signaliserer både høy energi og fokus på de riktige utfordringene: Skalerbarhet, styringsverktøy og tilgjengelighet.

Spesielt har skalerbarhet vist seg å være en langt større utfordring for NAS-løsninger enn selv leverandørene var klar over. Den universelle sannhet at 'å administrere 100 av hva som helst er bortimot umulig' dukket pussig nok opp som en overraskelse for begge sider av markedet. Horisontal skalering betyr flere enheter å holde styr på – inntil styringsverktøy og virtualisering gjemmer mengden bak et skall av effektiv programvare. Alternativet er vertikal skalering – mer lagringskapasitet bak hver enkelt boks – hvilket tar vare på skaleringsproble-

met i første omgang, men samtidig introduserer nye utfordringer. Den ene er at selve filsystemet alltid har begrensninger, både med hensyn til størrelse og antall filer. Slike begrensninger kan være absolutte eller ytelsesmessige – at ytelsen påvirkes vesentlig i negativ retning når gitte terskler passerer.

En annen utfordring har med båndbredde å gjøre. NAS-bokser har begrenset CPU-kapasitet og endelig båndbredde mellom prosessor og fysisk lager. Båndbredden kan håndteres via en oppjustering av hardware-arkitekturen, for eksempel med FibreChannel-svitsjer bygget sammen med styringslogikken (prosessor og hukommelse). Hvorvidt og hvor lenge dette hjelper er imidlertid et åpent spørsmål, fordi prosessorkapasiteten fortsatt er endelig. Vi kan kaste megahertz, gigahertz eller parallelle arkitekturer etter problemet, og fjerne slike flaskehalsar – hvilket NAS-leverandørene legger ned store summer i. Samtidig øker imidlertid kostnadene, og våre tidligere svært så attraktive regnestykker ramler sammen. Faren for at vertikal skalering av NAS-systemer skaper flaskehalsar i stedet for forenkling og pålitelighet, er med andre ord betydelig.

Sist, men ikke minst har NAS som konsept et troverdighets-problem i forhold til høy tilgjengelighet og klynger. Mens sistnevnte per definisjon betyr å gjøre identiske ressurser tilgjengelige fra flere kanter, hvilket er overkommelig for grunnleggende ressurser som prosessorkapasitet, hukommelse og masselager (datablokker), er det en langt større utfordring å tilby tilsvarende muligheter på filsystem- og filnivå. Vi kan få en rekke ulike tilgjengelighetsforsikringer fra dagens NAS-løsninger – med *hardware failover*, dublering av utstyr, RAID og så videre. I mange tilfeller er dette tilstrekkelig, men må likevel ikke oppfattes som ekvivalent med klynger.

En ideell arkitektur

Til tross for disse utfordringene, er en kombinasjon av NAS og SAN ønskelig. Som vi antydte innledningsvis, er kombinasjonen sågar logisk: NAS er en naturlig *front end* som leverer tjenester på filnivå, mens SAN er en naturlig *back end*, en leverandør av rå lagringstjenester – på blokknivå. For at kombinasjonen skal bli optimal, må imidlertid *front end* utstyret få en arkitektur som tar vare på skaleringsutfordringene både horisontalt og vertikalt (problemstillingen skal gjemmes bak effektiv teknologi). Dernest må kapasitetsutfordringene håndteres – blant annet gjennom parallellisering. Til slutt må graden av robusthet og pålitelighet heves gjennom en ny generasjon av hardware og programvare.

Om markedsføring og regnestykker

Leverandørenes vidløftige kalkulasjoner av TCO og ROI er sjelden relevante som praktiske eksempler. De tar utgangspunkt i virkelighetsbilder som sjelden forekommer andre steder enn på laboratoriet, og utelater dessuten vesentlige deler av virkeligheten.

ROI – *Return on Investment*

TCO – *Total Cost of Ownership*

SCSI – *Small Computer Systems Interface*

SATA – *Serial ATA*

ATA – *AT (Advanced Technology Attachment)* (en annen betegnelse for IDE – *Integrated Device Electronics*)

RAID – *Redundant Array of Inexpensive Disks*

FC – *FibreChannel*

Fibre channel – grunnlag for forvirring

I motsetning til hva navnet indikerer, har Fibre Channel (FC) ingen direkte tilknytning til fiber. FC er en protokoll – eller mer spesifikt et sett av protokoller – laget for lagringsnettverk, med opprinnelse tilbake til ca. 1990 (se Mellvik-Rapporten nr. 4). Årsaken til navnet er at teknologien i sin tid ble utviklet for fiber-basert infrastruktur, og ingen anså det som mulig at slike hastigheter kunne kjøres via andre fysiske media. I dag er kobberkabel like vanlig som fiber i små lagringsnettverk – og vesentlig rimeligere.

FC er en standard, men en standard med problemer. Ikke bare er den kostbar å implementere – et forhold som har sin opprinnelse i dens alder, og som stimulerer markedets søken etter rimeligere alternativer. Videre er den lite fleksibel i forhold til dagens behov. Aller viktigst er imidlertid at samspillet mellom produkter fra ulike leverandører ofte er problematisk. To produkter som beviselig begge følger standarden, kan fortsatt være umulige å få til å spille sammen, fordi standarden er for 'løs i kantene' og overlater for mange kritiske valg til implementasjonen.

Dette er noen av faktorene som driver frem interessen for alternativer til FC i disse dager, som vi kommer inn på avslutningsvis i denne artikkelen.

I bakkant må en slik *front end* kunne håndtere enten sitt eget direkte-koblede masselager – for miljøer med begrenset kapasitetsbehov, via SCSI, SATA og RAID, eller ha IP- og FC-baserte grensesnitt for tilknytning til en tradisjonell SAN-løsning.

Teknologien for å lage en slik arkitektur finnes, og den eneste årsaken til at produktene uteblir, er at NAS-leverandørene ikke ser seg tjent med å tilby slike løsninger. Så lenge det lar seg gjøre, ser de seg best tjent med proprietære arkitekturer som gir maksimal kontroll over kundens lagringsløsning. En åpning i bakkant – til SAN-løsninger – kan potensielt kannibalisere deres salg av fysisk lagringskapasitet, hvilket vil si mer enn halvparten av omsetningen. Så lenge størrelse måles mer i omsetning enn i marginer og overskudd, er en slik fokusering naturlig, kanskje sågar nødvendig. På den andre siden er det sannsynlig at en arkitekturoendring vil presse seg fram på sikt, katalysert av overgangen til IP-basert lagring, som vi skal se nærmere på nedenfor.

Valgets kval

Hvilke konklusjoner kan vi trekke fra disse observasjonene? Først og fremst at en behovsanalyse fortsatt er en grunnleggende forutsetning for evaluering og valg av lagringsløsning. Tabellen nedenfor oppsummerer noen av de viktigste fordeler og ulemper knyttet til henholdsvis SAN og NAS, og svaret er slett ikke alltid 'enten/eller'.

Som vi har vært inne på, er krav til kapasitet og tilgjengelighet to nøkkelparametre. Likeledes er sikkerhetskopiering en faktor som påvirker valget. I små og oversiktlige forhold kan SAN være det vi kaller *overkill*. Et SAN krever størst investeringer, men kan ha driftsmessige fordeler som mer enn oppveier disse kostnadene – for eksempel i tilknytning til nettopp sikkerhetskopiering.

Tabell 1 SAN og NAS har tradisjonelt tatt vare på ulike behov – med NAS' fokus på små og mellomstore miljøer, og SANs fokus på store miljøer og ditto datamengder. Forskjellene reduseres etterhvert som de to vokser mot hverandre, men er fortsatt signifikante. Tabellen påpeker noen av fordelene og ulempene knyttet til de to. Se også diskusjonen om konsolidering og virtualisering på side 7.

	Fordeler	Ulemper
SAN	<ul style="list-style-type: none"> • Typisk raskere • Enklere sikkerhetskopiering • Større fleksibilitet og skalerbarhet • Større robusthet/høyere tilgjengelighet 	<ul style="list-style-type: none"> • Forholdsvis kostbart • Betydelig forvirring med hensyn til standarder – kompatibilitetsproblemer på FC-siden, iSCSI fortsatt fersk for mange leverandører • Komplisert konfigurasjon
NAS	<ul style="list-style-type: none"> • Relativt lav kostnad • Enkel drift • Sømløs installasjon • Ingen tilleggsprogramvare nødvendig (utover administrasjonsprogrammer) 	<ul style="list-style-type: none"> • Ytelsesmessig svakere enn SAN • Komplisert sikkerhetskopiering som kan forstyrre driften • Begrenset skalerbarhet • Ingen reelle klynge-muligheter (ifm. høy tilgjengelighet)

Ethernet og IP vs. FibreChannel

Om iSCSI og IP-lagring i Mellvik-Rapporten:

- "iSCSI i din fremtid" – nr. 91 s. 27
- "Nettverkstrender" – nr. 96 s. 4
- "Fleksible tjenerarkitekturer" – nr. 99 s. 14

Om FibreChannel er gammel i teknologisk forstand, er Ethernet nærmest for en oldsak å regne. Likevel er det først og fremst Ethernet som utfordrer FC i lagringssammenheng, og forklaringen er innlysende enkel. Mens Ethernet ikke bare har hengt med i utviklingen, men sågar vært toneangivende, har FC i beste fall 'luntet etter'. Ethernet har dannet skole for enkelhet, kompatibilitet og standardisering, mens FC har vært det motsatte, som vi også er inne på i faktarammen på side 5. Og om FC har vært bærebjelken i SAN-sammenheng så langt, er Ethernet på full fart inn – i første omgang som støttespiller, i neste omgang som erstatning for FC.

Utviklingen ønskes velkommen av en hel IT-verden – med unntak av leverandørene som har investert stort i FC-kompetanse og utvikling. Ethernet appellerer til alle med sin enkelhet, aggressive utvikling, allsidige bruksområder, bred kompetanse i markedet, lav pris og så videre. Spesielt har Gigabit Ethernet og den gryende bruken av 10 Gigabit Ethernet vært viktige katalysatorer i lagrings-sammenheng. En oppsummering av de viktigste drivkreftene bak Ethernet kan se slik ut:

- ✓ Gigabit-hastigheter og høy pålitelighet
- ✓ Flytkontroll (IEEE 802.3x) – som er viktig for anvendelser som antar at pakker mottas i samme rekkefølge som de sendes, for eksempel iSCSI

Utvikling, konsolidering og virtualisering

Når vi betrakter utviklingen – innen teknologi og på de fleste andre områder – fra et visst perspektiv, vil vi som regel oppdage pendel-effekter: Metoder, oppfatninger og prinsipper blir avløst, for så å dukke opp igjen noen år senere, gjerne noe modifisert, men likevel lett gjenkjennelige. Fra en annen og tredimensjonal synsvinkel ser pendelen ut som en spiral. Det går hele veien fremover, men rundt en akse som gjør at vi kjenner igjen viktige elementer fra generasjon til generasjon.

Dagens trend bort fra titalls eller hundretalls tjenere med hver sine oppgaver, passer godt inn i et slikt bilde. For noen år siden var mange bra, fordi det var billig og ga robusthet. I dag vet vi bedre. Det visste vi også for 15 år siden, men 'glemte' det i en periode.

Konsolideringen som herjer i markedet, er med andre ord ikke bare sunn, men helt nødvendig. Det er ironisk at floraen av tjenere for alle tenkelige formål som vi nå søker å kvitte oss med, oppsto fordi 90-tallets operativsystemer var vesentlig svakere og mindre pålitelige enn 80-tallets. I dag har vi behov for å få situasjonen tilbake under kontroll, og konsoliderer ikke bare flere tjenester inn i samme maskin, men mange tjenere inn i samme hardware. Mens konseptet er det samme som IBM benyttet i stormaskiner på 70- og 80-tallet, har teknologien og leverandørene gjennomgått et par omdreininger på utviklingsskruen.

I dag er VMware den ledende leverandøren, mens plattformen er Intel-basert – og markedsaktørene er 'alle som betyr

noe' – i PC-server segmentet, med IBM, Dell og HP i spissen. Samtlige tilbyr konsolidering og virtualisering, sammen med systemer (hardware) og verktøy tilpasset slike løsninger. Denne konsolideringen av mange logiske tjenere inn i få fysiske, har en rekke fordeler utover å forenkle driften og redusere kostnadene. En av dem er større stabilitet, delvis på grunn av at de simulerte omgivelsene bidrar til mer forutsigbare omgivelser for klientsystemene, som typisk er Windows-tjenere. Den andre siden av stabilitetsbildet er at lagringen naturlig frikobles fra systemene, slik at en logisk tjener med et håndgrep kan flyttes fra én fysisk maskin til en annen. Med VMwares ESX-produkt og tilhørende administrasjonsverktøy, er det sågar mulig å flytte en virtuell tjener fra én fysisk tjener til en annen uten at brukerne merker det. En slik fleksibilitet og robusthet forutsetter at de fysiske systemene deler en felles SAN-løsning, som gjerne i sin tur har sine egne mekanismer som sikrer høy tilgjengelighet. Resultatet blir en særdeles attraktivt kombinasjon av konsolidering, forenkling og pålitelighet.

Årsaken til at konseptet forutsetter et SAN, er at de virtuelle tjenernes masselager i sin helhet må være tilgjengelige – blokk for blokk og alltid oppdaterte – fra alle (typisk to eller tre) fysiske tjenere som skal kunne kjøre nettopp disse virtuelle tjenerne. Selv i relativt små miljøer – med 10 til 30 tjenere, har slik konsolidering gitt bemerkelsesverdige resultater. Med andre ord et alternativ som er vel verdt å undersøke nærmere.

- ✓ Virtuelle kanaler (på nivå 2 – IEEE 802.1Q, også kjent som VLAN tagging)
- ✓ Prioriteringsmekanismer (IEEE 802.1p/Q)
- ✓ Aggregering av fysiske linker (IEEE 802.3ad)
- ✓ Hardware-basert protokollstøtte (grensesnitt med innebygget TCP/IP-processor, såkalte *TCP Offload Engines – TOE*)

Som vi ser, er tilveksten av Ethernet-relaterte standarder på en rekke ulike områder en viktig forutsetning for den utviklingen vi nå ser konturene av. Videre er overføring av protokoll-håndtering fra sentral prosessor til grensesnitt (NIC), nødvendig for i første omgang å sikre høy effektiv hastighet, og i andre omgang å opprettholde det raske priset på systemer og komponenter. For eksempel er iSCSI utenkelig dersom protokoll-håndteringen skal foregå på tjenersiden, mens en NIC med *Offload Engine* gjør protokollhåndteringen transparent, både funksjonelt og ytelsesmessig.

Første runde: Støttespiller for FC

Den første praktiske anvendelsen for Ethernet i SAN-sammenheng var som støttespiller for FibreChannel. Ingen kjenner FCs svakheter bedre enn leverandørene i segmentet, og behovet for større fleksibilitet i SAN-løsningen sørget for at en *FibreChannel over IP* protokoll ble utviklet i siste halvdel av 90-årene.

Protokollen muliggjør sammenkobling av FC-baserte SAN-installasjoner via Ethernet og TCP/IP, over avstander som tidligere var utenkelige. Sammenkoblingen skjer via FC-svitsjer som ved hjelp av FCIP oppretter transparente tunneller seg imellom. Teknologien fungerer, men har mange svakheter – blant annet fordi både FC-protokollen og FCIP-spesifikasjonene er løse i kantene. Videre blir effektiviteten skadelidende fordi pakkestørrelsen i et FC-nett er større enn i Ethernet. Derfor er det vanskelig å finne noen som mener at FCIP er noe annet enn en midlertidig nødløsning.

Andre runde: IP-lagring

Neste runde – noen vil kalle det generasjon – er en variant over samme tema, fortsatt med sterke koblinger til FC, men med vesentlig bedre karakteristika. “Storage over IP” (SoIP) ble ‘oppfunnet’ og promotert av teknologiselskapet Nishan, en pionér i lagringssegmentet som i 2003 ble oppkjøpt av langt mer kjente McData. SoIP introduserer mer intelligens i håndteringen av FC, blant annet ved å koordinere pakkestørrelsene, slik at ineffektivitetene fra FCIP blir eliminert.

SoIP fungerer rimelig godt og har betydelig utbredelse – egentlig en banebrytende teknologi i den forstand at den åpnet vei for kombinasjonen IP og Ethernet i et segment som var temmelig stivnet i teknologisk forstand.

SoIP er én av flere teknologiske bidrag fra Nishan. En annen er iFCP/mFCP, som legger komplette, umodifiserte FC-pakker i henholdsvis TCP- og UDP-pakker (TCP for lange, UDP for korte distanser)

for transport over Gigabit Ethernet. Teknologisk er dette både enklere og mer effektivt enn SoIP, men løsningen fremsto som truende for de andre aktørene i bransjen, og vant derfor aldri gjennomslag.

Tredje runde: Finito FC

Mens FC er velkjent og veletablert, har det siden 2001 vært åpenbart for de fleste med tilknytning til lagringssegmentet at teknologiens dager er talte. Den er for kostbar, lite fleksibel og uten et nytt ytelses- eller utviklingsnivå på horisonten. Infiniband har i noen år våket på horisonten som et mulig alternativ, og er på ingen måte borte. Enkelte leverandører arbeider fortsatt hardt for å promotere nettopp Infiniband som neste generasjons FC – av innlysende årsaker: I motsetning til Ethernet og IP er Infiniband spesialisert lagringsteknologi, som – dersom den får sitt gjennombrudd – vil beskytte leverandørene i lagringssegmentet mot det teknologiske og prismessige råkjøret som hersker i resten av nettverks-verden.

Sjansene for at satsingen skal lykkes er imidlertid beskjeden. Riktignok fremstår Infiniband som en selvskreven vinner rent teknologisk – skreddersydd som den er for nettopp disse oppgavene. Økonomien og fleksibiliteten i kombinasjonen Ethernet/IP er imidlertid vanskelig å hanskkes med, selv for spesialistene i segmentet, og Infiniband blir i beste fall en nisje i markedet – i verste fall et kort mellomspill.

Dermed seiler en ren Ethernet-basert lagrings-infrastruktur opp som det mest sannsynlige og mest optimale scenario for fremtiden. Disker, disk-kontrollere og RAID-kontrollere med varierende grad av innebygget intelligens og direkte Ethernet-tilkobling, er allerede tilgjengelige og raserer oppfatningen i markedet av hva slike komponenter skal koste. En første utstyrsgenerasjon med proprietære, UDP-baserte protokoller fungerte både som et *proof of concept* og som en *teaser* for leverandører som enda ikke hadde bestemt seg.

Så kom iSCSI-spesifikasjonen på banen og fjernet enhver tvil om hvilken vei det bar. De første iSCSI-produktene ble demonstrert i 2002, og teknologien står i dag på startstreken – klar til å snu opp-ned på vante forestillinger om hva, hvor og hvor mye i et stadig mer spennende SAN-segment.

Neste utgave

I neste utgave avslutter vi denne miniserien om lagringsteknologi med en teknisk gjennomgang av iSCSI og en diskusjon om hvordan veien videre ser ut. ■