



”– Qbranch är hjärtat i vår driftmiljö och håller tummen i ögat på Telenor som har våra WAN-länkar. De håller samma tumme i ögat på maskinvaruleverantörerna, och så är de ovärderliga för vår tekniska utveckling.”

Thomas Ericson Teknikchef på Riksbyggen



Riksbyggens nya park är inget svartbygge

Ännu en packe servrar konsoliderade. Fastighetsförvaltaren Riksbyggen kom fram till att man helt enkelt hade för många servrar och att det var för osäkert med lokala diskar. Vi har tittat på deras lagringslösning och den heta striden mellan Fibre Channel och iSCSI. Av Jörgen Städje

Allt fler företag strävar mot högre driftsäkerhet och bättre skalbarhet. En flaskhals är alla hårddiskar som går sönder hela tiden. För frågan är ju inte om en hårddisk ska krascha, utan när. Ju fler servrar, desto större risker. Lagringsnäten, SAN, med centrala, redundanta diskar är lösningen. Det finns två konkurrerande metoder för dataöverföring till det centrala datalagret: Fibre Channel (FC) och Internet SCSI (iSCSI). De presterar ungefär lika, men prisskillnaden är stor. Den svenska fastighetsförvaltaren Riksbyggen (RB) valde det billigare alternativet iSCSI – vi lade vår näsa i blöt och ville veta varför.

Token Ring-fäste i forntiden

Vi tog en skinkmacka med RB:s IT-chef Christer Wallin, teknikanvarige Thomas Ericson och outsourcingfolket från Qbranch på Lilla Essingen i Stockholm. Hur detta hänger ihop kommer vi till senare. Medan smörgåsarna långsamt transfererades till de olika magarnas djup skärptes pannornas veck och samtalet kom in på data-lagringsfrågorna. Christer Wallin började med den grå forntid:

– RB var ett Token Ring-fäste fram till 2001, när vi definitivt tröttnade på dyr nätutrustning och långsamma WAN-länkar. Vi hade 160 servrar utspridda på 60 platser och all backup skedde lokalt. I den mån den över huvud taget skedde. Det var mycket problem med magnetbanden, övervakning, återläsning, montering med mera. Vi konsoliderade servrar i många omgångar och började skaffa snabbare länkar.

– Vi såg hur sårbara vi var. En incident fick oss att vakna till och slutligen ta tag i frågan om ny lagringsteknik. Det var en server som havererade och den tog mycket lång tid att få igång. Behovet av att lagra filerna fristående från maskinvaran uppstod. Det ledde fram till att vi började titta på central fillagring och leverantören Proact föreslog iSCSI som protokoll. Vi började med en mindre modell av Netapps lagringsserver, en FAS 810, och fann att den var tillförlitlig och att det var mycket lättare att konsolidera till en lagringsserver än till en massa NT-servrar. Men FAS 810 var lite för liten i längden och idag har vi två klustrade FAS 920 C.

Det är alltså iSCSI som gäller. Men om det fungerar så bra med iSCSI – varför köper folk FC då?

– Den frågan ställde vi oss också.



50 procent av Riksbyggens IT-avdelning, IT-chefen Christer Wallin till vänster och teknikanvarige Thomas Ericson till höger, och i mitten "Miss Q" herself: Qbranch:s kundansvariga Sofia Svanholt. Gettingen får symbolisera stinget i iSCSI.

DRASTISKT:

160

servrar blev 70 när Riksbyggen konsoliderade. Alla de 160 hårddiskarna blev till noll.

Vi var lite oroliga i början eftersom det inte fanns särskilda iSCSI-kort till serverna.

Måste man ha särskilda iSCSI-kort?

– Vi trodde det, svarar Thomas, men det visade sig att det räckte alldeles

utmärkt med vanliga Ethernet-kort. Över huvud taget kom vi undan mycket billigt. Det blev mindre än halva priset för iSCSI mot FC, inklusive hårddiskarna. Det är SCSI-diskar i arrayerna, inte FC-diskar.

Qbranch vakar över Stokabfiber

Men RB har inga servrar i huset själva utan har utkontrakterat driften till Qbranch AB i Stockholm, som sköter om maskinvaran och lämnar support till RB:s 2 600 anställda.

– Qbranch är hjärtat i vår driftmiljö och håller tummen i ögat på Telenor som har våra WAN-länkar. De håller samma tumme i ögat på maskinvaruleverantörerna, och så är de ovärderliga för vår tekniska utveckling. Vi började ha våra länkar hos Utfors, som sedermera togs över av Telenor, som har skött sig mycket bra och idag levererar fasta anslutningar till alla 230 orter där vi är representerade. I Stockholm har vi egen redundant Stokab-fiber till huvudkontoret, men i övrigt kör vi Telenors nät.

Litet SAN jag koppla vill...

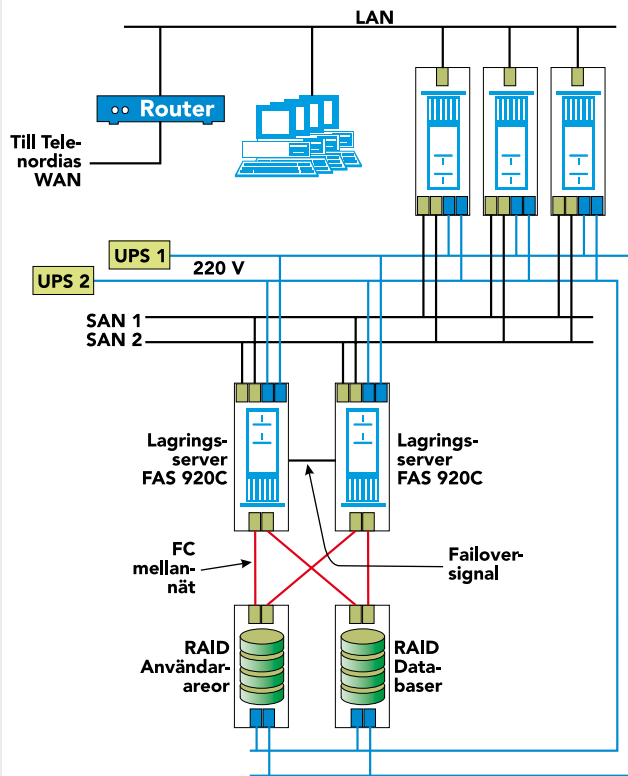
Namnfloran är yvig inom alla bran-

4

fakta om Riksbyggen och deras installation

- **Alla de 2 600 anställda** på de 230 lokal-kontoren i landet kommer nu åt sina data på ett billigare och snabbare sätt än tidigare.
- **Tillförlitligheten och driftsäkerheten** har ökat betydligt, fastän kostnaderna har minskat. Riksbyggen är numera ett 365 x 24-företag.
- **Raid-arrayen håller idag 2,5 terabyte**, men den är raskt på väg mot 5 terabyte – allting ligger på SCSI-diskar och hanteras av dubblerade lagringsservrar.
- **Storebror ser dig:** Qbranch övervakar med programmet Big Brother, som i mycket liknar till exempel Tivoli. Netapps lagringsserver råkade redan ha rutiner för att kunna rapportera till Big Brother, men vanliga servrar måste förses med en agent för att storebror ska kunna se.

Dubbelt upp av allt – för säkerhets skull



REDUNDANSERNAS HÖGBORG. Hos Riksbyggen har alla servrar dubbla nät-aggregat som försörjs från dubbla elnät, som kommer in i var sin ände av dator-hallen, från var sin UPS. SAN:et är dubblerat, varför alla servrar har dubbla Ethernet-kort. Anslutningen till hårddiskarrayerna är dessutom korskopplad så att en lagringsserver kan ta över om den andra skulle gå sönder. Så det finns lite Fibre Channel trots allt, mellan lagringsserverna och diskarna. Användar-nätet, LAN:et, är inte dubblerat, utan allt krut har lagts på att få säker datalagring.

scher, som vanligt, tills allt har lugnat ned sig. Det finns två kombattanter på lagringsarenan som hävdar sin särart när de säljer SAN (Storage Area Network) respektive NAS (Network Attached Storage), men det är egentligen bara hårklyverier. Båda systemen går ut på att man flyttar ut hårddiskarna ur serverna och har dem i ett centralt skåp

istället, oavsett kommunikations-protokoll. Lagringsskåpet styrs i båda fallen av en styrenhet – en lagringsserver med enda uppgift att flytta data mellan hårddiskarna och nätet.

... som aldrig får stå still

Hur kopplar man då ihop det? Servrarna som ska konsolideras behö-

Proact skingrar dimridåerna: Därför föredrar folk dyra Fibre Channel

Vi vänder oss till Proact, som byggde RB:s lösning, med de kinkiga frågorna. Om nu ISCSI är så billigt, varför köper man då FC? Låt oss börja med en pris/prestanda-översikt och jämföra kostnaderna för olika typer av nätverk. Vi antar att företaget har 48 servrar som står i ett och samma serverrum, inte mer än 50 meter från varandra och arbetar mot samma lagrings-system. I exemplet struntar vi i alla redundanser. Lagringssystemet är lika i båda fallen och ingår inte i priset.

Ledningslängden på 1 mil var tidigare till fördel för FC, men det är inte längre någon skillnad mellan FC och ISCSI på fiber. Längden på FC över koppar är betydligt kortare, men den verkar ha fallit ur mode för ingen tycks använda den.

Tiden för att flytta 10 gigabyte är intressant om man ska spegla data mellan två disksystem. Vi antar att du ska spegla en burst på 10 gigabyte som uppstått under en belastningstopp på en minut. Hinner man få undan alla data innan nästa topp kommer?

Priset för att flytta en gigabyte är räknat på en pay-off-tid på fem år per länk.

Valet är både lätt och svårt. Ser man bara till priset för dataöverföringen vinner ISCSI på UTP med hästlängder, men om du inte kan vänta 88 sekunder på att få över 10 gigabyte får du gå upp i pris och välja FC. FC är inte lika dyrt som det har varit, tack vare konkurrensen med Ethernet, och priserna kan förväntas sjunka ytterligare när 10 Gbps FC kommer. Då har å andra sidan snabbare varianter av Ethernet också blivit vanligare.

Så, snälla Proact: varför köper man FC när det finns ISCSI?

– Det vanligaste är att man redan har FC-utrustning, kompetens, verktyg, övervakning och så vidare, och ska komplettera. Nästa skäl är att man kräver absolut högsta överföringshastighet, eller vill uppgradera sitt 1 Gbps-nät till 4 Gbps. Men för nyinstallationer pekar allt på ISCSI. Vilken systemavdelning som helst kan administrera ett UTP-nät och behöver inte investera i ny övervakningsutrustning.

ver egentligen inget särskilt utöver två nätverksportar: en för användarnätet och en för lagringsnätet. SAN:et ska lämpligen inte förbrukas bandbredd på LAN:et. Vill man ha extra säkerhet, sätter man in två nätverkskort mot SAN:et. Dessutom krävs lämplig drivrutin i servern, så att den kan boota från SAN:et. Helt vanliga CAT5-kablar

dras sedan till en gigabit-Ethernet-switch från i stort sett valfri tillverkare. Ett par UTP-kablar (för redundansens skull) går vidare till lagringsservern, i RB:s fall alltså två FAS 920 C från Network Appliance. Servrarna ansluts i sin tur till ett skåp med hårddiskar, i RB fall ett FAS 3000 på 5 terabyte.

Fast visst vore det väl otur om allt skulle slockna på grund av att strömmen gick eller UPS:en stannade? Det har Qbranch redan insett och har därför dubblerad elförsörjning i alla hallar, med ledare som kommer in från två håll, om det skulle bli brand i ena änden av huset. De båda ledarna försörjs av var sin UPS. Nu kan det bli svårt att slå omkull systemet.

Men det skulle förstås kunna gå så illa att en lagringsserver går sönder. Av den anledningen är RB:s två servrar klustrade och kan ta över belastning från varandra. Fibrerna

Så billigt är ISCSI jämfört med Fibre Channel

| Länktyp | FC, fiber | FC, fiber | FC, fiber | ISCSI, fiber | ISCSI, UTP |
|--|--------------------------|--------------------------|---------------------------|-------------------------|----------------------|
| Länkkapacitet (Gb/s) | 1 | 2 | 4 | 1 | 1 |
| Max teoretisk datahastighet (MB/s) | 106 | 212 | 424 | 113 | 113 |
| Maximal ledningslängd | 1 mil | 1 mil | 1 mil | 1 mil | 100 m |
| Priser* i kronor | | | | | |
| 48 dubbelriktade ledare, 50 meter | 1 200 × 48 = 57 600 | 1 200 × 48 = 57 600 | 1 200 × 48 = 57 600 | 1 200 × 48 = 57 600 | 500 × 48 = 24 000 |
| Switch(ar) för totalt 48 ingångar, exempelvis 2 × 32 | 611 730 | 611 730 | 611 730 | 170 000 | 8 956 |
| 48 nätverkskort (HBA) | 13 300 × 48 = 638 400 | 13 300 × 48 = 638 400 | †19 950 × 48 = 957 600 | 5 475 × 48 = 262 800 | 385 × 48 = 18 480 |
| Totalt | 1 307 730 | 1 307 730 | 1 626 930 | 490 400 | 51 436 |
| Totalt per länk | 27 244 | 27 244 | 33 894 | 10 217 | 1 072 |
| Tid för att flytta 10 GB (sekunder) | 94,3 | 47,2 | 23,6 | 88,5 | 88,5 |
| En förflyttad GB/länk (milli-öre) [‡] | 162,9 | 81,5 | 50,7 | 57,3 | 6,0 |

* Proacts priser, utom UTP-priserna som kommer från Dustin. †) Hypotetiskt, eftersom produkterna är för nya. 50 procent har lagts på priset för 2 Gb/s-produkten. ‡) Payofftid: 5 år.

⇒ från lagringsserverna till diskarrayerna är dessutom korskopplade för att man ska kunna klara det sällsynta fall att en diskarray går sönder. Då går det att backa över data till den andra arrayen och köra vidare ändå.

Inget smör och bröd-program

Summarum är ett svenskt system för fastighetsförvaltning, som hanterar precis allt. För RB är det ett strategiskt system som måste vara igång jämnt. Förlora det, och verksamheten går i stå helt och hållet. Inga avlopp blir rensade, inga trappor skottade och inga hyror betalda. Systemet innehåller en fastighetsdatabas, det sköter hyresuppbörd, tar emot felanmälningar och skickar vidare till teknikernas mobila klienter. Det håller också reda på schemalagda aktiviteter (skruva, skotta, rensa ogräs), arbetsorder, medieförbrukning, nyckelhantering, ritningar – ja allt som kan behövas för att hålla ett hus i toppskick. Alla de 2 600 anställda måste kunna komma åt data hela tiden. Även trappstädare och snöröjare har mobila terminaler.

Summarums föregångare kördes på en IBM S/390 stordator fram till år 2003 och datakonverteringsjobbet för att få över det till Windows-miljö var allt annat än enkelt. Datalagret är på väg mot fem terabyte med en farlig fart eftersom RB har ett projekt igång med att digitalisera alla gamla husritningar, ända från 1940-talet. För dem som är inblandade i själva byggprocessen finns en enorm mängd dokument: ritningar, avtal och skisser.

RB är dessutom en av landets största redovisningsbyråer, i och med att man kör bokföringen åt 2 500 bostadsrättsföreningar och fastighetsägare.

RB har lyckats standardisera mycket på server- och klientsidan. Alla applikationer körs på applikationsservrar, genom terminalservrar. Alla anställda har bärbara datorer av samma sort, och de som är rörliga har handdatorer, en sorts förvuxna, oömma mobiltelefoner från Panasonic. Skrivarfloran har bringats ned från ett oändligt antal till bara fem.

– Vi är ett riktigt Microsoft-fäste,

Hjärta och hjärna i Riksbyggens SAN-slott

► Övervakningsprogrammet ger en snabb överblick av dataflödet. Medelflödet ut ligger på 19,4 megabyte per sekund, men toppar på 51.



Christer Wallin klappar om sin RAID-array. Där finns nu 2,5 terabyte data, men det sprutar in mer hela tiden och snart är man uppe i 5 terabyte.



1 gigabit per sekunds ISCSI (gröna UTP) kopplas som vanligt med CAT5-kabel. Fibrerna från lagringsserverna kommer in längst ned.

myser Christer. Vi har bara sådant. Maskinparken består av ett 70-tal servrar, inklusive terminalservrar. Visst, det märks ju att Microsoft är lite för stöddiga och det vore kul att gå en annan väg, men ...

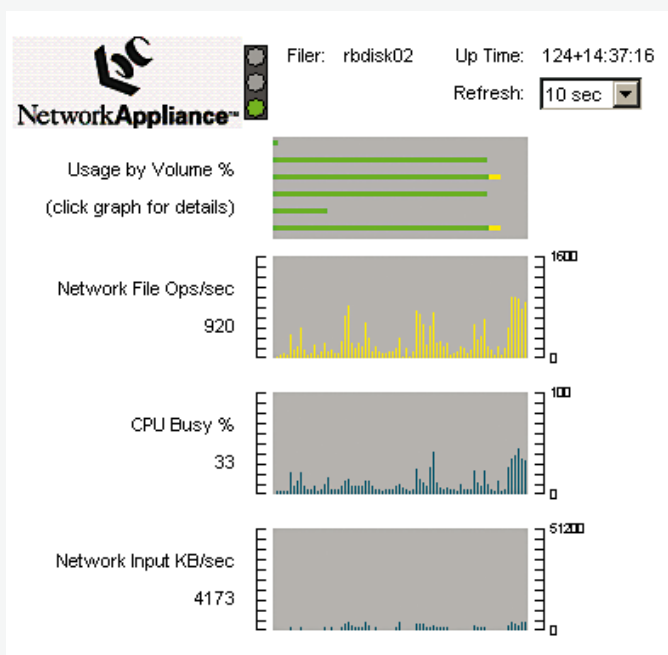
Inga Linuxplaner?

– Nej, inga planer, men vi håller naturligtvis ögonen på det.

Framtiden: "Vi sover lugnt"

Nästa steg för RB är att helt gå över till digital hantering av fakturor. De kommer också att bygga vidare på sin kommunikations- och fastighetsdriftscentral i Kalmar och utveckla ekonomisystemen.

– För en mindre peng än idag kommer vi att kunna lagra mera data, ha flera applikationer och ännu fler klienter. Allt detta har upp-



Britta, kan du ge mig lite 8GFC?

Det kryllar av konstiga ord. HBA (Host Bus Adapter) är namnet på det kort man sätter i servern för att ansluta den till ett SAN, oftast använt av FC-folket. TOE (TCP Off-load Engine) är ett Ethernet-kort med extra processor för att själv kunna göra all nätverkshantering (fast det kan väl snart alla Ethernet-kort idag ...). Säg du "nätverkskort" i båda fallen! Men sen blir det värre. ISCSI-HBA-kort är endast avsedda att ta hand om SAN-trafiken, medan ISCSI-TOE-kort är generellare och kan hantera både SAN och vanlig LAN-trafik samtidigt. Av den anledningen har HBA-korten färre CPU:er och blir billigare. Om rådet i största allmänhet är att köra LAN och SAN på separata nät, så gör det inget alls att HBA är billigare än TOE. Akronymikerna har också klämt ihop till exempel 8 Gigabit per second Fibre Channel till 8GFC. GbE betyder Gigabit Ethernet. Ack ja.

SÅ GÅR DU VIDARE

- Riksbyggen presenterar sig på www.riksbyggen.se, dock ingen teknik.
- Det var Proact som gjorde RB:s installation. De har också stått för prisuppgifterna, www.proact.se.
- Qbranch sköter driften, titta på www.qbranch.se.
- Lagringsservern Network Appliance FAS 920 finns på www.netapp.com/products/file/enterprise-filers.html medan ISCSI beskrivs på www.netapp.com/products/iscsi-technology.html. Själva hårddiskskåpen FAS3000 beskrivs på www.netapp.com/products/file/fas3000_ds.html.

Jörgen Städje är elektronikkonstruktör, nätverksspecialist, teknikhistoriker och datorgrafiker. Han nås på jorgen@qedata.se.