

Terugdringen beheerkosten in tegenspraak met groeiende vraag naar opslag

Bestandsvirtualisatie: uitweg uit dilemma?

Ondernemingen staan voor de uitdaging een oplossing te vinden voor het opslagprobleem van de almaar toenemende hoeveelheid informatie. Dit stelt ze voor een dilemma: aan de ene kant moeten de operationele uitgaven voor IT-beheer worden teruggedrongen, aan de andere kant nemen de complexiteit en operationele kosten toe bij elk nieuw toegevoegd opslagsysteem.

Bram Dons bespreekt de mogelijkheden van bestandsvirtualisatie en bekijkt of die vorm van virtualisatie een remedie biedt.

Bram Dons

De enige manier om het opslagprobleem op korte termijn op te lossen is het efficiënter benutten van de opslagsystemen; bij de meeste enterprise-ondernemingen

wordt slechts veertig tot vijftig procent van de opslagcapaciteit benut. Dit inefficiënte gebruik is niet het resultaat van onderbezetting, maar meer van *over-provisioning*. Er is geen eenvoudige, niet-verstorende manier om te beantwoorden aan de vraag naar meer opslagcapaciteit zonder daarbij het opslagsysteem te moeten verstoren. Een manier om dit dilemma op te lossen is de toepassing van intelligente filevirtualisatie. Zo heeft de firma Acopia een filestorage-infrastructuur ontwikkeld die een oplossing kan bieden voor dit capaciteitsprobleem. Daarmee rijst de vraag waarom onze huidige filestorage-infrastructuren (zie figuur 1) tekortschieten en met welke problemen, bestandssystemen worden geconfronteerd als gevolg van de informatiegroei.

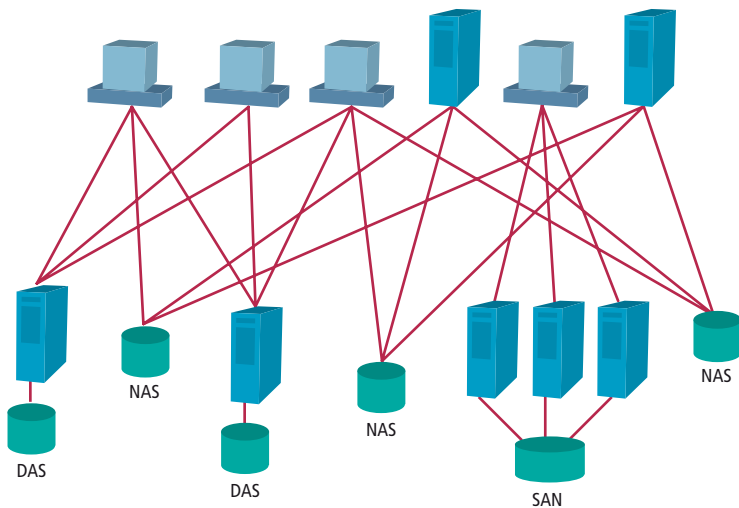
Tekortkomingen huidige bestandssystemen

Een doorsnee-enterprise-opslaginstructuur bestaat uit een complexe verzameling opslagplatforms, file- en besturingsystemen die vaak nog van verschillende leveranciers zijn ook. Helaas kunnen de meeste apparaten niet zonder de nodige aanpassingen goed met elkaar samenwerken en moet elk 'opslageiland' afzonderlijk beheerd worden. Gebruikers en applicaties zijn statisch verbonden

met fysieke bestandsopslagbronnen. Iedere keer dat er iets binnen de opslagomgeving voor de bestandssystemen gewijzigd moet worden, bijvoorbeeld de toevoeging van een nieuw opslagsysteem of de verhuizing van bestanden, worden deze statische *mappings* verstoord. Het updaten van de omgeving is een handmatig proces waarbij het systeem vaak tijdelijk moet worden uitgeschakeld.

Een oplossing hiervoor is de statische koppeling tussen client en opslagbronnen te elimineren en deze te vervangen door een systeem waarin de plaats van opslagbronnen dynamisch is aan te passen, met de mogelijkheid om data vrijelijk tussen opslagbronnen te kunnen verplaatsen zonder dat de toegang van gebruiker tot de data verstoord wordt. Oftewel: bestandsvirtualisatie.

Een ander probleem ligt in de beheerkosten van nas-systemen. De eenvoudige toepassing van een nas-systeem voor het bieden van file services is voor de IT-manager een aantrekkelijke optie. En hoewel het beheer van enkele nas-systemen nog betrekkelijk eenvoudig is, wordt het beheer van tientallen of meer behoorlijk arbeidsintensief en dus kostbaar. Ook voor de beheerder is uitbreiding een mooie ontsnapingsmogelijkheid.



Figuur 1 Huidige statische filestorage-infrastructuur

Een enkel bestandssysteem dat tegen de fysieke grenzen van de servercapaciteit aan zit, vraagt bij een uitbreiding naar een grotere server met meer capaciteit een grote beheerinspanning. Het is voor de beheerder dus veel eenvoudiger om een nas-server bij te plaatsen. Het nadeel hiervan is wel dat er weer een opslageland gecreëerd wordt. Elk toegevoegd nas-systeem moet afzonderlijk en handmatig beheerd worden, omdat elke nas-server zijn eigen afzonderlijke bestandssysteem heeft.

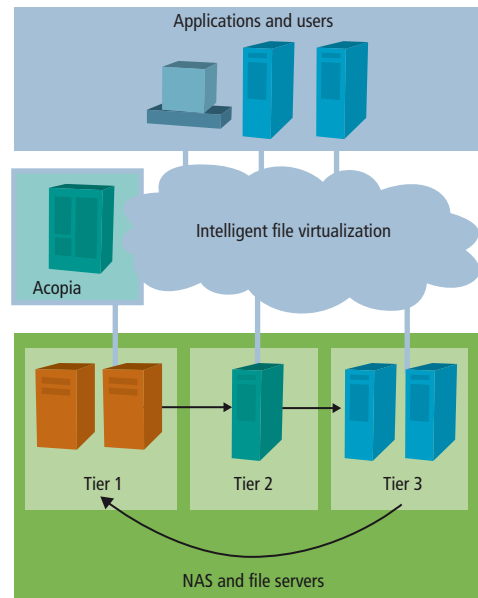
De meeste bestandssystemen ondersteunen slechts een enkel type opslag. Dat bemoeilijkt de implementatie van een gelaagd opslagsysteem, bijvoorbeeld als onderdeel van een toepassing van information lifecycle management (ilm). Tot op heden nemen veel ondernemingen hun toevlucht tot de implementatie van clusterbestandssystemen, om hun opslag- en prestatieproblemen het hoofd te bieden. Maar de kosten en complexiteit van dergelijke toepassingen zijn hoog. Ze vereisen de nodige expertise van de beheerder, en ondernemingen worden na de installatie doorgaans geconfronteerd met een *forklift upgrade* bij de migratie naar het nieuwe bestandssysteem en mogelijk nieuwe hardware. Een nieuwe toepassing van virtualisatie-technologie kan deze bezwaren onder-

vangen: intelligente bestandsvirtualisatie. Hiermee wordt het mogelijk om de complexiteit van meerdere nas-systemen voor de beheerder te maskeren, routinematige en administratieve taken te automatiseren, te stroomlijnen, en daarbij te voldoen aan de gestelde eisen van een hoog presterende bestandssysteem met lage tco.

Acopia's virtualisatie

Het door de firma Acopia geboden filevirtualisatiesysteem – dat hier bij wijze van voorbeeld gepresenteerd wordt – brengt een intelligente laag aan binnen het IP-netwerk, die de logische toegang tot bestanden ontkoppelt van de fysieke locatie ervan. Met behulp van deze laag zijn data vrijelijk te verplaatsen en kunnen opslagbronnen worden verhuisd. De door Acopia toegepaste technologie introduceert geen nieuw bestandssysteem, maar maakt gebruik van de bestaande systemen CIFS en/of NFS.

Deze intelligente filevirtualisatielaag biedt talrijke voordelen voor de IT-omgeving. Ten eerste, de talrijke clientnaar-opslagbronnen-mappings kunnen worden geconsolideerd tot een beter beheerbaar aantal. Bovendien hoeven deze mappings nooit meer veranderd te worden, ze zijn van blijvende aard. Opslagbeheertaken, zoals data provisi-

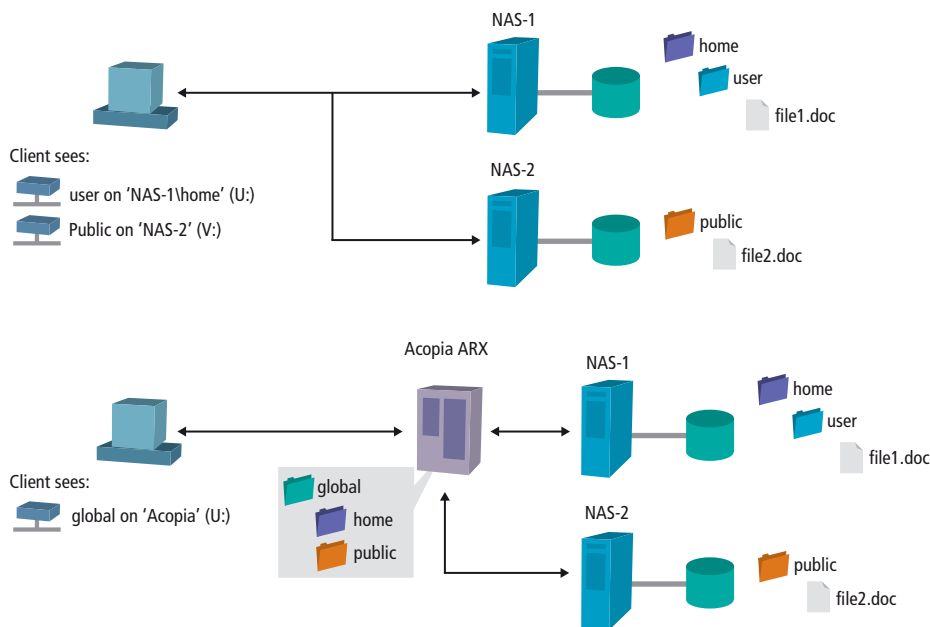


Figuur 2 Gelaagde opslagarchitectuur van Acopia

oning, consolidatie en migratie, kunnen direct worden uitgevoerd, zonder de noodzaak tot configuratie van de client. Omdat data niet langer gekoppeld zijn aan een fysieke opslagbron kunnen ze worden verplaatst zonder gevolgen voor de client die op dat moment toegang tot deze data heeft. Overtollige opslagcapaciteit kan transparant, zonder gevolgen voor de client, worden gedeeld door verschillende heterogene opslagsystemen. Op die manier is de beschikbare opslagcapaciteit optimaal te benutten. Ten slotte, filevirtualisatie biedt keuzevrijheid bij de aanschaf van opslagsystemen en voorkomt een *vendor lock-in*.

Intelligente filevirtualisatie voegt bestaande bestandssystemen bij elkaar, ongeacht of deze zich op meerdere of verschillende typen nas-apparaten, file-servers of nas gateways bevinden. Na de virtualisatieslag hebben clients en applicaties toegang tot bestanden via de global namespace. De client ziet nog slechts één 'global' drive mapping en hoeft er niet langer rekening mee te houden (en weet ook niet) op welke nas-systeem of share/export het bestand is geplaatst (zie figuur 3). De client behoudt dezelfde logische drive mapping zoals die bestond voor de virtualisatieslag, ongeacht de fysieke locatie van een bestand binnen

virtualisatie



Figuur 3 Bestandstoegang zonder en met virtualisatie

de gevirtualiseerde namespace. Dit komt doordat clients zich nu via de global namespace toegang verschaffen tot files in plaats van de fysieke apparaten zelf. Nadat de bestandssystemen op alle nas-systemen zijn gevirtualiseerd, kan de beheerder policy's instellen op het Acopia-apparaat voor de automatische verplaatsing en vervanging van files. Deze actie is volkomen transparant voor de client.

Typen gelaagdheid

Acopia's opslagbeheersysteem (zie figuur 2) ondersteunt een aantal beheertaken die voorheen handmatig moesten worden uitgevoerd door de beheerder. De uitgevoerde taken blijven transparant voor de gebruiker, wat voordelen biedt voor de opslagbeheerder. Zo kunnen verschillende taken geautomatiseerd worden, waaronder datamigratie tussen opslagplatforms van verschillende leveranciers (zonder de toegang van clients tot de data te hoeven verstoren), de implementatie van *storage tiering* (een gelaagd opslagmodel), of ilm-policy's en datareplicatie waarbij CIFS/NFS-bestanden tussen heterogene storageapparaten kunnen worden gespiegeld. Acopia ondersteunt geautomatiseerd een gelaagd opslagmodel dat in drie fasen wordt doorlopen: specificatie van (1) welke data, (2) waarheen en (3) wanneer verplaatst moet worden.

Werking

De eerste stap in een geautomatiseerd storagetieringmodel of ilm is te bepalen welke data op een bepaalde storage tier verhuist of verplaatst moeten worden. Dit staat bekend als *data-classificatie*. De classificatietoepassing van Acopia biedt een eenvoudig toepasbare dataclassificatiemethode, op basis van metadata, die veel flexibiliteit bij het classificeren van data biedt. In tegenstelling tot veel andere classificatietools, die een intensieve coördinatie tussen de verschillende bedrijfsafdelingen vereisen voor een diepgaande analyse van de relaties tussen bestanden en applicaties.

De meeste basisclassificatie is gebaseerd op de leeftijd van bestanden. Acopia's leeftijdgebaseerde policy's – die trouwens te combineren zijn met andere policy's – bieden de beheerder de mogelijkheid om bestanden in de delen op basis van de 'laatst gemodificeerde' of 'laatste toegang' bestandsattributen. Bestanden die voor een bepaalde tijd niet zijn benaderd of gemodificeerd, worden dan naar een andere (goedkopere) opslaglaag gemigreerd. Er kunnen ook policy's worden gedefinieerd om bestanden die naderhand toch zijn gemodificeerd of opgevraagd, weer naar een primaire opslaglaag terug te zetten.

Na de dataclassificatiefase geeft de beheerder aan op welk doelbestemming de geselecteerde bestanden worden opgeslagen. Het doelopslagsysteem kan een specifieke share of export naar een file-server of een *share farm* (een groep met shares of exports die gebundeld zijn in een storage pool) zijn. In de context van een gelaagd ilm-model kan de beheerder voor groepering zorgen binnen een bepaalde opslaglaag, op basis van kosten, betrouwbaarheid, snelheid of andere door de gebruiker gedefinieerde criteria.

Andere toepassingen

Ook bestaat de mogelijkheid van replicatie van bestanden voor de creatie van een disasterrecoverytoepassing. Replicatie gebeurt op gezette tijden over een IP-netwerk, waarbij alleen de verandering worden overgestuurd. De gereplieerde bestanden kunnen daarna in een *read-only mode* door gebruikers worden benaderd, maar ook via een omzetting naar een *read/write mode* voor een recoveryoperatie.

Met Acopia's *appliances* (apparaten met ingebouwde software, red.) kunnen beheerders load balancing policy's configureren, gebaseerd op applicatiebelasting en/of de verdeling van capaciteit over meerdere fysieke opslagapparaten. Dit alles gebeurt transparant voor de applicaties en gebruikers.

Voordelen

Acopia's toepassing biedt talrijke voordelen ten opzichte van soortgelijke producten:

- Ten eerste blijft de verhuizing van bestanden tussen de opslaglagen tijdens de gehele levensduur transparant voor gebruikers en applicaties. Beheerders hebben de vrijheid om een willekeurig aantal opslaglagen te construeren, terwijl de meeste andere toepassingen een beperkt aantal opslagplatforms ondersteunen.
- Het is niet nodig om software agents op servers of clients te installeren en te onderhouden. De apparaten van Acopia leiden niet tot de invoering



Producten en prijzen Acopia

Acopia biedt drie *appliances*: ARX500, ARX1000 en ARX6000. Qua softwarefunctionaliteit is er geen enkel verschil tussen de drie apparaten. De keuze voor een bepaald platform kan plaatsvinden op basis van schaalbaarheidskenmerken, waarbij niet het aantal terabytes maar ook het aantal files en de gewenste doorvoersnelheid een rol spelen.

De ARX500 (vaste configuratie) schaaft tot 128 miljoen files en heeft een maximale doorvoersnelheid van 100 Mb/sec, de ARX1000 schaaft tot 384 miljoen bestanden en biedt 433 Mb/sec doorvoer, en de ARX6000 gaat tot twee miljard bestanden en meer dan 2 Gb/sec.

De catalogusprijs van de ARX500 bedraagt \$ 29.995, die van de ARX1000 \$ 69.000. De prijsstelling van de ARX6000 is wat lastiger te specificeren, omdat dit een modulair systeem is en er met een zogenaamde *starter kit* wordt gewerkt. De prijs voor deze starter kit bedraagt \$ 215.000. Voor multiprotocol (dus zowel NFS en CIFS) betaalt men een extra licentieprijs: voor de ARX500 \$ 3.600, voor de ARX1000 \$ 3.600 en voor de ARX6000 \$ 12.000. Wat men er in euro's voor moet betalen, is natuurlijk afhankelijk van de euro-dollarkoers.

Er wordt geen ondoorzichtige licentiestructuur gehanteerd voor de softwarefuncties. Dit is een onderscheidende eigenschap in de storagemarkt, aangezien het gebruikelijk is om met terugkerende licentiekosten te werken, bijvoorbeeld voor replicatiefunctiefunctionaliteit.

www.acopia.com

van nieuwe bestandssystemen of opslaghardware.

- ILM-policy's worden op bestandsniveau en niet op het niveau van datablocks of complete bestandssystemen toegepast. Dit is een voordeel, want bestanden bieden een 'bedrijfscontext', datablocks uiteraard niet. Bovendien is een ilm-policy die wordt toegepast op een compleet bestandssysteem weinig inzichtelijk en van weinig waarde.
- Dataclassificatie vindt plaats door middel van eenvoudige, flexibele criteria op basis van metadata-attributen.
- Los van de toegepaste virtualisatietechnologie is het voordeel van deze virtualisatietoepassing de te bereiken kostenbesparing. Er wordt efficiënter gebruikgemaakt van de beschikbare opslagcapaciteit. Uiteraard zijn de te verwachten kostenbesparingen en de *return on investment* afhankelijk van

de IT-architectuur en liggen ze voor iedere onderneming weer anders.

Bram Dons is zelfstandig IT-analist.
E-mail: b.dons@IT-trendwatch.nl.

