

Compellent zet belangrijke stap

Maximaal profiteren van thin provisioning blijkt lastig

Thin provisioning (TP) lijkt op dit moment als technologie daadwerkelijk door te breken. Om te kijken of deze technologie in de praktijk alle beloftes waarmaakt, onderwerpt Bram Dons de storageoplossing van de firma Compellent aan een kritische test.

Thin provisioning is niet nieuw. Al lange tijd ondersteunen 3PAR, Compellent en andere kleinere leveranciers TP in hun storage arrays. Maar nu zijn ook de belangrijke storageleveranciers het afgelopen jaar begonnen met het introduceren van thin provisioning. De brede beschikbaarheid van TP heeft ook de discussie tussen storageleveranciers over welke gebruikte techniek het meest effectief is, aangewakkerd.

Niet alle implementaties van thin provisioning zijn hetzelfde, omdat er een groot verschil bestaat tussen de kleine storageleveranciers, zoals 3PAR, Compellent, EqualLogic en LeftHand Networks, en anderzijds grote leveranciers als Hitachi Data Systems. Zo stelt bijvoorbeeld de Dynamic Provisioning Software van HDS een storage pool samen uit 42 MB pagina's die het over devices of RAID-sets streept. Volgens sommige experts is dat een effectievere methode dan bijvoorbeeld 3PARs opdeling in kilobytes.

Ook de maximumgrootte van een storage pool verschilt per leverancier. Tevens bestaan er verschillen in prestaties van 'thin provisioned' LUNs, met name door de manier waarop data over de disks

wordt gedistribueerd. Zo kan een disk array van leverancier Compellent dynamisch de data bij minder gebruikte disks distribueren. Sommige systemen kunnen cache aan specifieke storage pools toewijzen.

Afgezien van al deze onderlinge verschillen is de meest belangrijke vraag hoe een applicatie omgaat met onverwacht gebrek aan diskruimte. Daarvoor moet een TP-product hecht zijn geïntegreerd met de bestaande management- en business continuity-processen; eenvoudig en snel diskcapaciteit aan een thin-provisioned pool kunnen toevoegen en de hoeveelheid door de host gevraagde opslagcapaciteit kunnen aanpassen wanneer een bepaalde capaciteitsdrempel wordt overschreden.

Eigenschappen

Tot op de dag van vandaag gaan OSs en applicaties ervan uit dat het nog steeds direct met een 'hard' drive met een vaste grootte is verbonden. Of dat nu een disk is die via DAS of een SAN, iSCSI of Fibre Channel is gekoppeld, maakt voor het OS verder niet uit. Sterker nog, deze ziet niet eens het verschil. Hoewel de laatste tien

jaar de meeste servers en applicaties hun bestanden in een array subsysteem opslaan, die op een flexibele wijze een opdeling in LUNs kunnen maken, worden ook LUNs nog steeds door het OS als een vaste hoeveelheid storage, de Physical Unit, gezien.

Er is geen enkele reden om aan te nemen dat het gedrag van operating systemen en applicaties in directe relatie tot storage systemen de komende jaren gaat veranderen. Ook niet in een virtuele omgeving, want het OS ziet geen verschil tussen een gewone LUN en een virtuele LUN. Voor het OS blijft een LUN een LUN.

Naast de kosten voor opslagsystemen, vormt de directe relatie tussen het OS en het disksysteem voor het IT-beheer, bij uitbreiding of wijziging van de benodigde hoeveelheid opslagruimte, een enorme belasting. TP is een technologie die de hoge kosten die daarbij tevoorschijn komen, kan terugdringen.

Thin provisioning is een storage management eigenschap die het OS als een virtuele drive presenteert met een grote capaciteit, maar in werkelijkheid is dit natuurlijk niet het geval. Sommige experts vinden dat TP een vorm van virtualisatie is. De redenering hiervoor is dat TP op basisniveau als een systeem werkt dat stukken gealloceerde storage in kleinere eenheden of blokken opdeelt en die vervolgens automatisch over de verschillende componenten binnen de storage pool verspreid. Daarmee verbreekt het de traditionele afhankelijkheid van een LUN of RAID-set, dat niets meer of minder is dan een stel gebundelde diskdrives die

een gemeenschappelijk schema gebruiken voor het lezen en schrijven van data. Door de introductie van TP als laag tussen een LUN of RAID-set, ontstaat een abstractie van de fysieke datalocatie en aldus een vorm van virtualisatie, waarvoor wat te zeggen valt. TP maakt de creatie van een pool met logische diskbronnen afkomstig van meerdere arrays mogelijk. De Hitachi Universal Storage Platform (USP V)-oplossing biedt zelfs TP met externe disk arrays van verschillende storageleveranciers!

Toepassingen

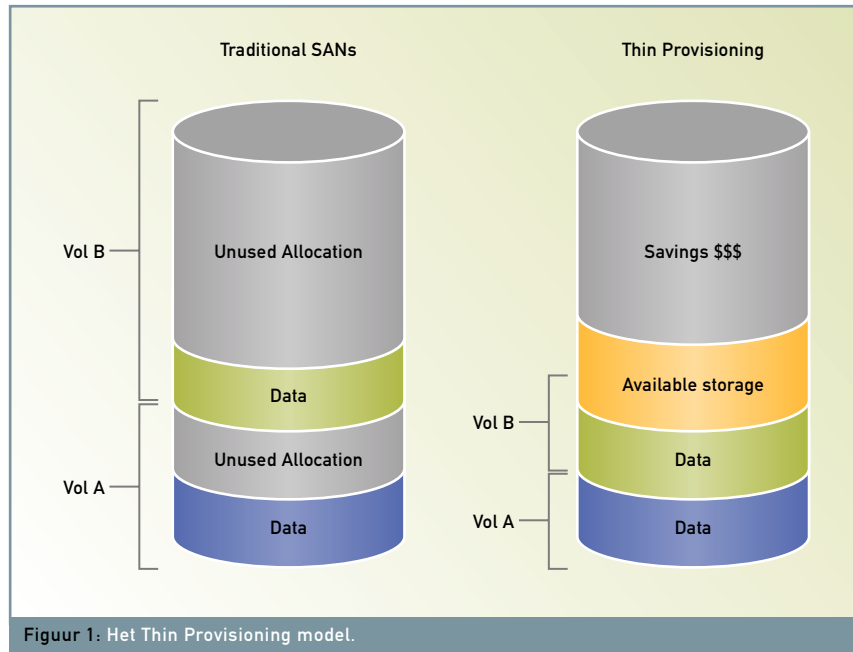
Een van de vanzelfsprekende voordelen van TP is dat de noodzaak voor de handmatige configuratie bij de toewijzing van storage komt te vervallen, omdat uitbreiding daarvan automatisch gebeurt. Doorgaans wordt de omvang van een volume iedere twee of drie jaar vergroot en, afhankelijk van de applicatie, het OS en de storage array, neemt deze operatie minstens een paar uur in beslag. In grote omgevingen met bijvoorbeeld meer dan 150 volumes betekent een volume-uitbreiding voor IT-beheer een gemiddelde aanpassing van één keer per week, een taak die met de invoering van TP komt te vervallen.

Een minder in het oog springend voordeel is dat bij deze methode een bijna 100 procent gebruiksgraad van storage kan worden bereikt. Daarnaast bespaart

NIET ALLE IMPLEMENTATIES ZIJN HETZELFDE

TP, in combinatie met Remote Mirroring, diskruimte op beide kanten van de mirror, naast een substantiële besparing in de hoeveelheid benodigde netwerkbandbreedte voor het opzetten van de mirror. Dat komt omdat bij creatie van een mirror bij block-gebaseerde remote mirroring, eerst het complete diskvolume over het WAN moet worden getransporteerd. Met TP hoeft echter alleen de werkelijk gebruikte diskruimte te worden getransporteerd, in plaats van het hele volume. Dit kan aanmerkelijk de benodigde tijd reduceren om een mirror op te zetten en biedt de mogelijkheid om een lagekosten-netwerkverbinding te gebruiken.

De meest voor de hand liggende vraag van een bedrijf bij toepassing van TP is hoeveel kosten ermee kunnen worden bespaard. Daarop vooraf een antwoord



Figuur 1: Het Thin Provisioning model.

geven is niet eenvoudig. Het aantonen van de feitelijke besparing is alleen achteraf aantoonbaar, ná de invoering van thin provisioning. Volgens storageleveranciers kan een doorsnee scenario voor thin provisioning de opslaggebruiksgraad van 60 naar 80 procent verhogen, uiteindelijk dus een besparing van 20 procent.

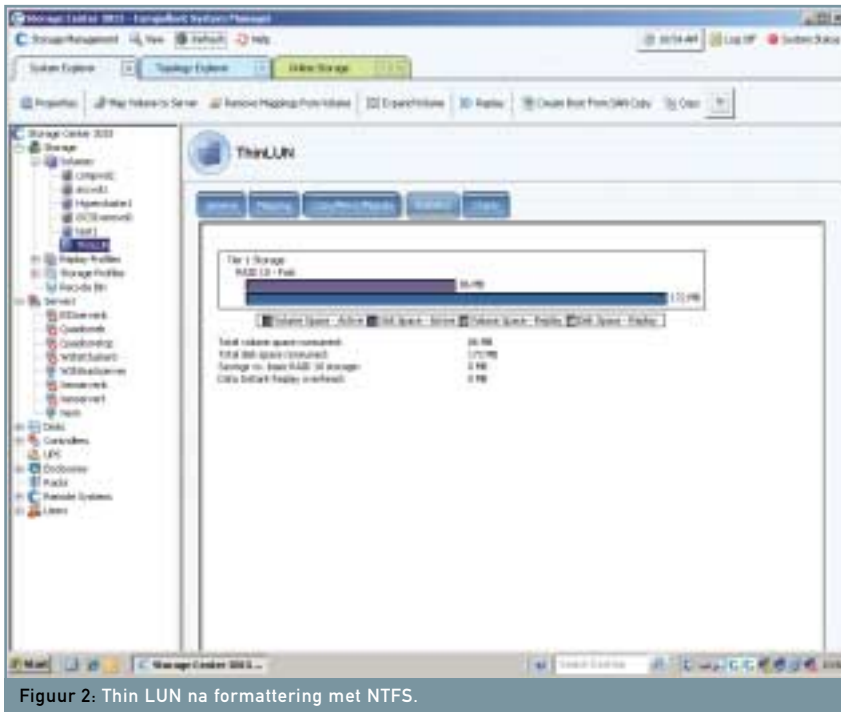
Nadelige effecten

Net zoals elke andere nieuwe technologie, brengt de implementatie van TP niet alleen nieuwe uitdagingen, maar ook problemen met zich mee. De IT-beheerder moet bij TP met een aantal zaken rekeninghouden. Onder andere wanneer de maximale fysieke opslagcapaciteit wordt bereikt, want op dat moment denkt de

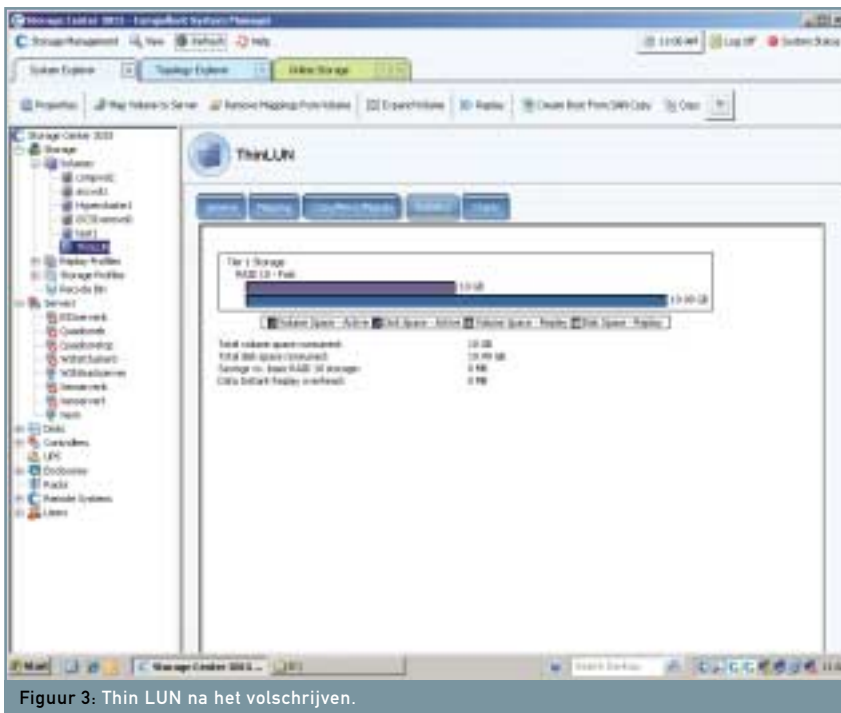
applicatie dat er nog genoeg lege fysieke opslagruimte beschikbaar is. In werkelijkheid is er echter minder ruimte beschikbaar. Een mogelijk scenario is dat wanneer de actuele maximale capaciteit wordt bereikt, de toegang tot het volume wordt geblokkeerd, totdat er meer capaciteit is toegevoegd. De beheerder krijgt dan, via bijvoorbeeld e-mail, een waarschuwing en vervolgens moet hij dan opslagruimte toevoegen. Dit verschuift de verantwoordelijkheid voor een 'einde opslagcapaciteit' situatie van de applicatieserver naar de storagebeheerder en wordt hiermee nu zijn verantwoordelijkheid om er voor te zorgen dat er altijd op tijd genoeg fysieke storage beschikbaar is. Een goed TP-systeem voorziet de beheerder van voldoende informatie om de

(Advertentie)

The advertisement features a background image of a server rack. The text reads: 'Backup/Restore Backup to disk net zo goedkoop als tape? Met deduplicatie is dit realistisch.' Below this is the website 'www.storageXperience.nl' and the ISIT logo with the tagline 'THE STORAGE COMPANY'. At the bottom, it says 'Reserveer nú een 1-op-1 sessie met een ISIT consultant'.



Figuur 2: Thin LUN na formattering met NTFS.



Figuur 3: Thin LUN na het volschrijven.

storagegroei te voorspellen en tijdig extra storagecapaciteit toe te voegen. Ook kunnen niet alle servers die de storage pool gebruiken tegelijkertijd extra opslagcapaciteit krijgen. Het is te vergelijken met een bank waarbij alle spaarders tegelijkertijd hun spaargeld in contanten willen opnemen, de meeste banken kunnen op dat moment daaraan niet voldoen. In de omgekeerde vorm, kunnen niet alle file systems en databases de eerder ver-

wijderde diskruimte opnieuw claimen. Sommige systemen hebben de voorkeur om voor eens en altijd nieuwe ruimte te alloceren, in plaats van hergebruik van vrij gekomen ruimte. Dit kan een probleem opleveren wanneer een applicatie constant files creëert en verwijderd. Dit komt omdat het TP-systeem continu zal groeien en om steeds meer fysieke diskruimte vraagt, terwijl de meeste van de vrijgekomen diskruimte niet door wordt hergebruikt. Afhankelijk van de imple-

mentatie, kan TP invloed hebben op de prestaties vanwege de noodzaak om grote tabellen en potentieel gefragmenteerde bestanden te moeten beheren. Ten slotte, ook het gebruik van TP in combinatie een file system kan tot problemen leiden, waaronder met het Windows NTFS file system.

Compellent thin provisioning

De disk array-architectuur van de firma Compellent biedt een TP-systeem waarin een ongelimiteerde hoeveelheid en grootte aan volumes kan worden aangeemaakt. Alle volumes worden in een enkele storage pool ondergebracht op basis van een automatische, 2 MB grote, page-gebaseerde allocatie. De allocatie gebeurt met behulp van de 'Dynamic Block Architecture' en gebruikt 'Dynamic RAID Level' conversie. Daarmee houdt het disksysteem specifieke informatie bij van elk diskblock, waaronder de schrijftijd, type disk en data, RAID-niveau. Op basis daarvan neemt de disk array beslissingen om data dynamisch te verplaatsen binnen het geautomatiseerde storage tiering model.

THIN PROVISIONING IS VORM VAN VIRTUALISATIE

Tevens ondersteunt de disk array replicatie, synchroon en asynchroon, en voorzieningen voor data recovery. Compellent's 'Automated Tiered Storage', 'Data Progression' genaamd, beheert inactieve data die zich op de verschillende storage tiers bevindt. Het drielaagse storage model is op de, binnenkort ondersteunde, SSD-, SATA II- en FC-diskens gebaseerd. Bij het aanmaken van een thin volume kunnen beheerders kiezen uit een van de drie disksystemen, of zij laten de keuze over aan de Compellent software waarbij in de loop van de tijd diskblokken, afhankelijk van het gebruik, automatisch van de ene naar de ander storage tier worden verplaatst. De 'Fastest Track'-voorziening verhuist automatisch veel gebruikte diskblokken naar de snelste buitenste disktracks. Een harddisk drive transporteert immers data sneller van de buitenste tracks dan de binnenste omdat deze meer data bevatten waardoor de invloed van de diskzoeksnelheid kan worden gemaskeerd. De transfersnelheid van binnenste tracks op een harddisk is ongeveer de helft van de buitenste!

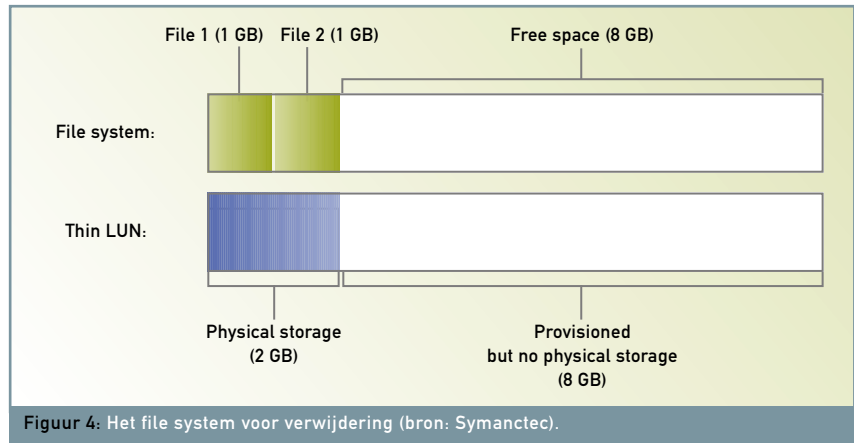
Test

Om te zien hoe TP in praktijk werkt, creëren we een 10 GB 'ThinLun' met behulp van de 'Compellent System Manager'. Wij kiezen voor een 'Tier 1', RAID 10, op FC-disks gebaseerde LUN. Op een Windows 2003 formatteren het LUN tot een NTFS-volume, 'TPLUN' genaamd. Wij constateren dat Compellent slechts 172 MB aan fysieke diskruimte heeft gebruikt van het onder Windows gecreëerde 10 GB volume (zie figuur 2). Vervolgens gaan we de 10 GB diskruimte onder Windows volschrijven om daarvan de gevolgen te zien voor het thin LUN. Bij het vollopen van het volume genereert het Windows file system een code 'I/O error 112' en stopt met het schrijven van data. Wij zien dat het TP-systeem automatisch tijdens het volschrijven van het volume de gedefinieerde 10 GB Thin LUN als fysieke ruimte voor Windows beschikbaar maakt (zie figuur 3).

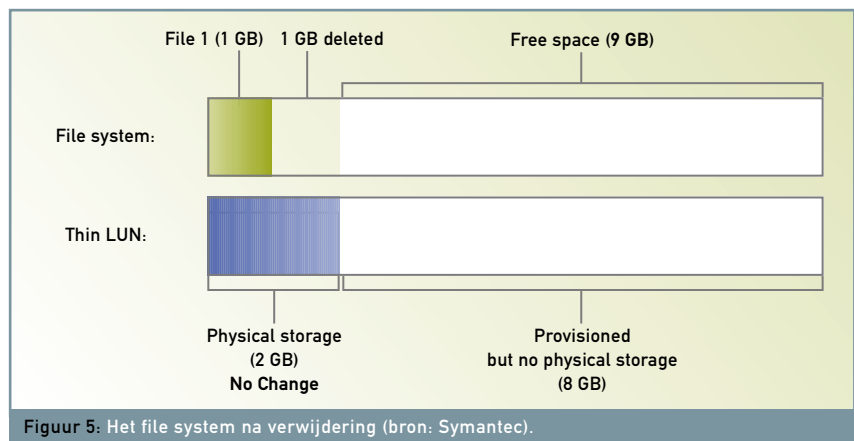
Verder biedt het TP-concept van Compellent de mogelijkheid om weer vrijgekomen fysieke diskruimte terug te geven aan de bestaande storage pool. Daartoe verwijderen we eerst weer alle data op het 10 GB NTFS volume 'TPLUN'. Daarna constateren we dat het Thin LUN nog steeds 19.99 GB in beslag neemt op de storage array. Hier zien we dus de discrepantie van het TP-mechanisme en het gebruikte file system. Maar gelukkig kent Compellent echter een special recover feature, 'Space Recovery', waarmee automatisch op regelmatige tijden, of onmiddellijk, vrijgekomen opslagruimte van het Windows NTFS-volume kan worden teruggewonnen en in de storage pool worden opgenomen.

Thin provisioning voor Windows

Uit de test blijkt dat op het moment dat gebruikers of applicaties files of records verwijderen, of dat nu op een file system of databaselaag is, er opslagruimte wordt vrijgemaakt. Dat geldt echter niet voor de fysieke storage in de thin LUN. Stel als voorbeeld een thin LUN van 10 GB, waarvan fysiek maar 2 GB aanwezig is. De 2 GB vertegenwoordigt voor het file system het hoogwatermerk voor gebruikte data. Stel dat dit systeem slechts twee bestanden heeft, elk 1 GB groot en de gebruiker verwijdert een file waardoor de actuele logische opslagruimte op de file systemlaag 1 GB wordt. De thin LUN blijft echter 2 GB! Eenzelfde situatie kan ontstaan als gevolg van een uitgevoerde defragmentatiebevoering. Kortom, zonder extra voorziening kan de vrijgekomen fysieke opslagruimte niet door TP zonder meer worden herbenut.



Figuur 4: Het file system voor verwijdering (bron: Symantec).



Figuur 5: Het file system na verwijdering (bron: Symantec).

De Free Space Recovery feature van Compellent biedt voor dit probleem een oplossing. De feature maakt deel uit van Compellent's Enterprise Manager en werkt in combinatie met het Windows file system om diskruimte dat niet langer in gebruik is te detecteren en deze weer aan de storage pool beschikbaar te stellen. De 'Server Agent for Windows' die op de betreffende server draait, vergaart en laat

informatie zien van het door het Windows OS gerapporteerde informatie. Free Space Recovery maakt het mogelijk om ongebruikte opslagruimte van het Windows NTFS file system te hergebruiken. Via de Enterprise Manager kan op regelmatige basis of onmiddellijk een recovery procedure worden gestart. In de praktijk wordt melding gemaakt van een 10 tot 20 procent teruggewonnen diskruimte door

(Advertentie)

The advertisement features a blue and white background with a perspective view of a server rack. The text is as follows:
File & E-mail Archiving
HSM is nog geen archiving! Kom kijken hoe het wel moet.
www.storageXperience.nl
Reserveer nú een 1-op-1 sessie met een ISIT consultant
ISIT THE STORAGE COMPANY
Part of Symantec

gebruik van Free Space Recovery. De Recovery feature werkt alleen op een Windows NTFS file system met 'Basic' disks en niet voor volumes die gestreept of gespiegeld zijn. Microsoft geeft echter aan dat er nog aanvullende beperkingen zijn bij TP.

Windows NTFS en thin provisioning

De media en documentatie van storageleveranciers berichten wel uitgebreid over de voordelen van TP, maar geen van hen gaat verder diep in over de toepassing binnen de Windows-omgeving, met name het NTFS file system. Dat er een probleem is, wordt geïllustreerd door een artikel dat valt te lezen op de support website van Microsoft. In het artikel staat te lezen dat: "het verbergen van de actuele storage niet goed werkt met het huidige ontwerp van NTFS en de betrouwbaarheid van het file system ongunstig kan worden beïnvloed door TP volumes". Het artikel gaat verder met: "We (aut: Microsoft) kijken naar een manier om beter thin-provisioned SANs te ondersteunen". In het bulletin wordt aanbevolen om geen TP te gebruiken voor een systeem volume, de opslag van een page file of gebruik van een disk fragmenter, die kan extra storage ruimte innemen. Windows Vista en Server 2008 hebben ingebouwde tools om een volume te laten toe- of afnemen. Na defragmentatie met de juiste tool, kan een NTFS-volume opleveren met ongebruikte, maar provisioned, block storage aan het einde van het volume. Daarna kan het volume worden ingekrompen. Of een TP-oplossing het ingekrompen volume herkent en de vrijgekomen block-ruimte weer inlijft in haar storage pool, valt onder de verantwoording van de TP-oplossing en niet van het Windows file system. Verder wordt in dit bulletin aanbevolen om bij een 'overcommitted' TP volume, deze eerst te 'unmounten'. Gebeurt dit niet, dan is de kans groot dat data die zich nog in het cache bevindt niet naar het volume wordt weggeschreven en verloren gaat. Een ander probleem is het verlies van data als er binnen TP een fout optreedt op het moment dat data wordt weggeschreven. Nog erger zijn de gevolgen als er zich een fout voordoet bij de toewijzing van een block storage cluster op het moment dat NTFS probeert metadata weg te schrijven. Een niet triviale oplossing is om de beheerder de NTFS disk quotas in te stellen op een zodanig niveau ,dat de limieten eerder worden bereikt voordat de block storage clusters op zijn. Dit vereist natuurlijk wel een handmatige aanpassing. NTFS

beheert block storage in eenheden, clusters. Het gebruik van TP bij Windows NTFS file system vereist een gedisciplineerde aanpak dat de toepassing van de juiste tools en technieken. Er ligt dan ook voor de leveranciers van TP een schone taak om een goede integratie van TP met NTFS-volumes tot stand te brengen.

Toekomst

Thin provisioning is een technologie die waarschijnlijk een standaardoptie gaat worden op alle nieuwe moderne storage-systemen. Naarmate TP meer en meer algemeen zal worden toegepast, zullen operating systems en applicaties rekening moeten gaan houden dat diskvolumes ook 'thin volumes' kunnen zijn. Elke oplossing voor TP zal dan ook moeten worden uitgerust met verbeterde geoptimaliseerde storage allocatiemechanismen.



Figuur 6: Compellent Space Recovery functie.

Ook de aan TP gerelateerde technologieën zullen de acceptatie bespoedigen. Een daarvan is servervirtualisatie. In een virtuele serveromgeving heeft elke VM zijn eigen boot- en datavolumes waarin de grootte van het boot volume wordt bepaald door de grootte van het OS en de bootsoftware, in overeenstemming met de grootte van het geheugen. Dit is nodig om de memory dump files te kunnen opslaan wanneer een applicatie faalt. In het algemeen reserveren gebruikers minimaal 4 tot 8 GB voor booting, terwijl in werkelijkheid slechts 1 GB nodig is. In een virtuele omgeving met honderden VMs betekent dit een verspilling die in de TBs kan oplopen. TP biedt daarvoor een oplossing door het operating system toch 8 GB te bieden, terwijl slechts 1 GB wordt gealloceerd. Bij uitval van de applicatie wordt alsnog de resterende GBs bij de creatie van de dump file toegewezen. Een andere technologie die zeker de toepassing van TP zal gaan versnellen is de

flashgebaseerde Solid State Disk (SSD). Vanwege de hoge prijs is het wenselijk dat de gebruiksgraad van SSDs tussen 80 en 90 procent ligt. Omdat SSDs geen last hebben van verminderde prestaties door fragmentatie, hebben sommige nadelen van TP geen invloed bij gebruik van SSDs. Verder biedt TP de mogelijkheid om het Storage Resource Management (SRM) van het applicatie- en file-niveau naar het storagestelsel en block-niveau te verplaatsen. Een van de redenen waarom SRM geen succes is geworden, komt door het gebrek om gebruikte diskruimte op applicatie- en server niveau te kunnen meten. Dit speelde vooral in omgevingen met honderden of duizenden servers met verschillende operating en file systems. Bij toepassing van TP wordt de gealloceerde diskruimte op block-niveau gemeten, in plaats van op file-niveau, met als voordeel dat dit op centraal niveau kan worden gemeten, onafhankelijk van de hoeveelheid of type applicatieservers. De nieuwe generatie SRM software zal, door de toepassing van thin provisioning, een enorme verbetering zijn.

Conclusie

Thin provisioning is een storage array gebaseerde technologie om efficiënt opslagruimte op aanvraag aan applicaties toe te wijzen, waarmee het gebruik van opslagruimte wordt geoptimaliseerd. Echter, het volledig profijt trekken uit een TP-oplossing is een uitdaging. Veel file systems kunnen niet efficiënt omgaan met gebruikte en niet langer gebruikte opslagcapaciteit. Bovendien kunnen bestaande applicaties niet intelligent de migratie van traditionele storage arrays naar thin provisioned arrays uitvoeren. Zonder een intelligent file system en integratie met een storagestelsel is het niet mogelijk om overbodige opslagcapaciteit naar de storage pool terug te sluisen. Het resultaat is een inefficiënt gebruik van opslagruimte. Tot de dag van vandaag heeft dit de toepassing van thin provisioned storage binnen de grootzakelijke markt vertraagd. ■

MEER INFORMATIE

Voor meer informatie over de problemen tussen thin provisioning en Windows NTFS: <http://support.microsoft.com/kb/959613>.