



# Volume snapshots

## Onmisbare aanvulling op de herstelstrategie

Voor het gebruik van volume snapshots komt steeds meer belangstelling. Er zijn legio redenen om er handig gebruik van te maken. NetOpus gaat in dit artikel dan ook dieper in op het gebruik van volume snapshots.

---

BRAM DONS

Een veelgebruikte toepassing is het maken van een backup door een kloon van het originele volume te maken. Het voordeel van deze methode is dat de actieve applicaties ongestoord van het originele volume gebruik kunnen blijven maken. Met snapshots wordt namelijk het zogenaamde 'open file' probleem bij backups opgelost. De snapshotmethode wordt ondermeer in Windows XP/2003 Volume Shadow copy Service (VSS) toegepast. Een andere toepassing van snapshots is om een kloon van 'live data' te nemen, voor bijvoorbeeld het gebruik in datamining-applicaties. Snapshots zijn ook handig om een nieuwe versie van applicaties te testen in een bestaande omgeving, maar ook als onderdeel van een complete disasterrecovery-oplossing neemt het gebruik van snapshots snel in populariteit toe. Er zijn echter tal van manieren te bedenken waarop snapshot-technologie een aanvulling kan zijn op een algemene herstelstrategie. Een belangrijke overweging om voor de toepassing van snapshots te kiezen, is de factor tijd. Want, de tijd die nodig is om (naast verkleining van de zogenaamde 'backup window') een restore uit te voeren kan met snapshots aanzienlijk worden bekort. Microsoft zelf, beschrijft in een white paper hoe snapshot-technologie kan worden ingezet om een snel herstel bij een crash van de Active Directory mogelijk te maken. We beschrijven hierna in het kort de toegepaste methode en we laten je tevens zien op welke wijze Microsoft de snapshot-technologie in Windows Server 2003 heeft geïmplementeerd. Volume Shadow copy Service en 'snapshots' zijn synoniem, maar om de intellectuele eigendomsrechten van het woord 'snapshot' te respecteren, hanteert Microsoft de naam 'Volume Shadows'.

### Verschillende snapshot-implementaties

Een snapshot is niets meer dan een consistente point-in-time kopie van een volume en maakt dus geen volledige kopie van de originele image-data. Het is een terstond aangemaakte set van pointers (of bitmap) die wijst naar de originele data op het moment van aanmaak. De vraag is hoe een volume snapshot wordt gecreëerd?

Daarvoor bestaan weliswaar verschillende methodes maar in feite maken ze allemaal van de duplicatie van een schrijfoperatie gebruik. Het enige verschil is de manier waarop dit gebeurt in soft- of hardware en de plaats waar dit gebeurt.

De eerste methode is gebaseerd op hardware waarbij een volume wordt gespiegeld en daarna afgesplitst (de zogenaamde 'split-mirror' technologie). Iedere schrijfoperatie wordt in twee duplicaten schrijfoperaties opgesplitst: een naar de source-, en ander naar de duplicaat disk-target. Voordeel van deze methode is dat het bijzonder snel en betrouwbaar is. Nadeel zijn de hoge kosten die het dubbel uitgevoerde opslagsysteem en de benodigde (proprietary) disk-controller hardware met zich meebrengen. De tweede manier is om een volume snapshot als file system filter driver in het NTFS- of FAT file system op te nemen en iedere IRP (I/O Request Packet) naar deze driver te sturen. Dit proces is echter vrij complex waarbij het niet eenvoudig is om van de door de hardware geboden snapshot-technologie gebruik te kunnen maken. Voorbeelden van applicaties die van filters 'boven' NTFS gebruik maken, zijn producten als St. Bernard Open File Manager, Vinca (nu LEGATO) Open File Manager en Cheyenne Open File Agent.

Een derde methode is om binnen het file system de voor snapshots benodigde driver-stack op te nemen. NetApps Write Anywhere File Layer (WAFL) en Linux SnapFS file systems zijn daarvan voorbeelden. Uiteraard draaien deze producten niet op het Windows NT file system. Nu is het schrijven van een file system op zich al complex genoeg, maar de toevoeging van snapshot-technologie maakt het nog aanzienlijk complexer. Op dit moment is er binnen het 'native' file system van Windows NT geen

systeem dat van deze methode gebruik maakt. Tenslotte is er de methode om 'onder' het file system volume snapshots te implementeren door een filter driver daarin in op te nemen. De methode is gebaseerd op de 'copy-on-write' techniek waarbij de door applicaties gewijzigde logische blokken eerst naar een ander opslagsysteem worden gekopieerd, voordat deze op het primaire opslagsysteem worden doorgevoerd. Deze techniek staat ook bekend als differential snapshot omdat alleen de verschillen worden opgeslagen (in plaats van een complete kopie bij mirroring).

## Win XP/2003 Volume Shadow Copy Service

Microsoft heeft op Windows XP/2003 de Volume Shadow copy Service (VSS) geïmplementeerd waarmee een framework wordt geboden voor het maken van een gecoördineerde en consistente point-in-time kopie van disk volumes. De service is als een filter driver geïmplementeerd (volsnap.sys) 'onder' het file systeem. De nieuwe VSS-service biedt een hoge mate van betrouwbaarheid en voor de uitvoering van complexe backup-up scenario's, die voorheen onder Windows niet mogelijk waren.

Microsoft heeft VSS op non-disclosure-basis een ontwikkelkit beschikbaar gesteld en ondersteunt daarmee een drie soorten ontwikkelaars, te weten:

1. ISV's die zogenaamde VSS 'writers' willen ontwikkelen voor bijvoorbeeld Exchange, SQL Server, Oracle, SAP, Sybase.
2. ISVs, die voor backup- en storage management (SRM) applicaties 'requestors' voor VSS willen ontwikkelen.
3. Onafhankelijke soft- en hardwareleveranciers (IHV's en ISV's) die hard- en software willen ontwikkelen voor backup, fault tolerance en data integrity; zoals VERITAS en EMC.

ISV's hoeven met VSS niet langer zelf een VSS-service te schrijven. De door Microsoft geboden VSS-service is namelijk te vergelijken met een print spooler, waarvan er slechts één binnen het systeem nodig is. Sommige ontwerpers (zoals de leveranciers van 'Providers') hoeven dan nog alleen maar het equivalent van een printer driver te schrijven en andere leveranciers de printer-applicatie.

## VSS architectuur

De met Windows XP/2003 meegeleverde VSS bestaat uit vier type modules: Writers, Requestors, VSS service en Providers. De snapshot writers zijn gewoon de

applicaties die data schrijven. Voorbeelden van snapshot writers zijn Exchange, SQL Server 2000, SAP en Oracle. De bedoeling is dat (niet alleen Microsoft zelf) met behulp van de SDK de ISVs op snapshot gebaseerde applicaties gaan ontwikkelen. Daarbij ontvangt de applicatie van de snapshot services waarschuwingen om de schrijffactiviteiten tijdelijk te beëindigen en na het maken van de snapshot weer voort te zetten (leesacties zijn uiteraard geen probleem).

Het gebruik van de VSS-service heeft een groot voordeel ten opzichte van de traditionele hardwaregebaseerde snapshot-systemen die binnen Windows 2000 en vorige Windows-versies werden gebruikt. Bij de klassieke versies voor de creatie van snapshots had het hardwaregebaseerde systeem geen enkele mogelijkheid om vast te stellen wat de staat van de applicatie en het OS was waarin deze verkeerden, met in bijzonder het Windows cache-systeem. Dat betekende in de praktijk dat veel snapshots inconsistent waren en alleen door het draaien van een speciale 'data consistency checker' kon achteraf worden vastgesteld of de genomen snapshots wel consistent waren. In tegenstelling tot de VSS-gebaseerde architectuur waar niet alleen het cache wordt leeggemaakt (flushed) en het schrijven wordt opgehouden maar ook binnen een minuut kan worden vastgesteld of de gecreëerde snapshot wel of niet consistent is. Snapshot writers leveren ook meta-data aan voor backup en restore, bijvoorbeeld welke bestanden gekopieerd moeten worden. Microsoft zal writers leveren voor zowel SQL Server 2000 als Exchange en werkt met ISV's samen om writers voor andere applicaties te ontwikkelen, waaronder Active Directory.

Requestors zijn doorgaans de backup-applicaties die aan de API van de VSS service vragen om een snapshot te creëren. Het voordeel hiervan voor de ontwerpers van backup applicaties is dat ze niet langer met allerlei moeilijke zaken rekening hoeven te houden, zoals de plaats waar de backup-data wordt opgeslagen, welk deel van de data bestaat uit de applicatie log-bestanden en welke speciale behandeling deze bestanden nodig hebben. De betreffende writer (bijvoorbeeld SQL Server) is verantwoordelijk voor de specificatie van de bestanden en directories die in de backup moeten worden opgenomen. Een restore wordt ook eenvoudiger omdat de applicatie niet de data hoeft te lokaliseren en welke bestanden het aan de API van de applicatie moet doorsturen. De restore-applicatie overhandigd gewoon de verzameling data aan de writer (applicatie) en laat die vervolgens

de restore uitvoeren.

Snapshot providers zijn de entiteiten die daadwerkelijk de snapshot of volume kopie creëren. De bedoeling is dat providers door ISV's en IHV's worden geschreven voor het creëren, verwijderen en manipuleren van snapshots. Met behulp van de snapshot-SDK moeten snapshot-providers als een 'COM out-of-proc' worden geschreven. De provider kan een kernel-component bezitten, bijvoorbeeld een filter-driver tussen het file system en de LVM, maar de kernel mode functionaliteit kan optioneel ook in hardware worden opgenomen. Een voorbeeld van een snapshot-driver is de door Microsoft default met XP en Windows Server 2003 meegeleverde volsnap.sys driver.

### **Bescherming van Active Directory informatie**

In Windows Server 2003 is Active Directory information service de verbeterde directory service voor het Windows OS. Het is de plaats om informatie over objecten op het netwerk, zoals gebruikers, computerbestanden, printers en applicaties op te slaan. Een middelgrote organisatie kan tienduizenden objecten omvatten. De meeste van die objecten zijn gebruikers. Enterprise-organisaties kunnen zelfs honderduizenden, of zelfs miljoenen objecten hebben. Het mag duidelijk zijn dat wanneer de AD-informatie niet meer beschikbaar is, gebruikers niet meer op het netwerk kunnen inloggen en dus ook geen toegang tot netwerkbronnen hebben. Het is aan te bevelen om de AD-servers op drie volumes te implementeren: een volume voor het OS, een AD database-volume en een voor de AD log-bestanden. De Active Directory maakt deel uit van de 'System State' backup, die de database bevat, log-bestanden, registry, system boot bestanden, COM+ registratie-database en Sysvol. Het is dan ook belangrijk dat een backup en restore met deze volumes als een set wordt uitgevoerd.

De AD informatie moet tegen een aantal potentiële gevaren beschermd worden, zoals disk- of andere hardware-fouten. Dit doe je door fault tolerant beschermde volumes (via mirroring of RAID-5) te gebruiken en door tenminste twee Domain Controllers (DC) te gebruiken in een Active Directory Domain. Hoewel mirroring en RAID-5 configuraties fault tolerance bieden voor potentiële fouten, bieden deze oplossingen geen bescherming tegen corruptie van data. Immers, de gespiegelde datakopie is, net zoals het origineel, ook beschadigd. Dit is te voorkomen door een enkele point-in-time backup kopie van alle volumes te maken, en deze fysiek gescheiden van het origineel te bewaren. Daarvoor

was tot kort in de meeste gevallen tape backup de aangewezen methode.

### **Beperkingen tape backup**

Een bekend nadeel van tape backup is dat het een tijdsintensief proces is en dat het op de server de nodige belasting vormt. Om deze reden wil men in een traditionele IT-omgeving meestal maar een keer in de week een full-backup maken en op de andere dagen een incremental backup. Vaak wordt voorbijgegaan aan het feit dat een restore ook een lange tijd vergt, met name in een omgeving met Active Directory (AD). Want, des te langer het restore proces vergt des te groter is de samensmelting van de shadow copies en de Active Directory servers die bezig zijn met de uitvoering van schrijftransacties naar disk.

Een beter alternatief voor de traditionele tape-gebaseerde bescherming is het maken van point-in-time shadows copies. Het gebruik daarvan bij Active Directory configuraties maakt een snel herstel mogelijk bij een aantal specifieke systeemproblemen. Denk aan problemen bij installatie van een service pack; een onstabiel systeem, als gevolg van een slecht functionerende device- of filter driver, corruptie van de systeem registry, of een virus dat op het systeem heeft huisgehouden.

Omdat het shadow proces snel verloopt en vrijwel geen invloed heeft op de systeemprestaties, kunnen shadow copies vaker en sneller worden uitgevoerd dan een tape backup. De shadow kopie kan binnen een SAN worden opgeslagen zodat voor een restore de backup-data weer snel toegang is. Met een daarvoor geschikte hardware-voorziening kan de kopie naar een backup server worden getransporteerd, op tape worden opgeslagen en daarna eventueel nog naar een offsite opslagplaats worden overgebracht.

### **Fast recovery van Active Directory**

Zoals we zagen coördineert de VSS-service drie software-componenten die nodig zijn voor het maken van een full-mirror shadow kopie. Ten eerste de utility die om creatie van een shadow kopie verzoekt (de Requestor). Ten tweede de applicatie-specifieke software die er voor zorgt dat de applicatie-data klaar staat voor het maken van een shadow kopie (in dit geval dus de specifieke Active Directory binnen Windows Server 2003). En als laatste de interface die daadwerkelijk de shadow kopie uitvoert: de Provider. De VSS kan shadow kopieën maken via (in-box) software-provider, of een 3th party hardware-provider. In het voorbeeld van Microsoft's white paper wordt van de HP StorageWorks VDS hardware provider gebruik

gemaakt, die een point-in-time kopie van een, of meerdere volumes kan maken. Een aantal leveranciers van storage arrays hebben al lange tijd een dergelijke hardware provider geïmplementeerd, waaronder HP, EMC, HDS en XIOTech Corporation. Deze proprietary hardware-gebaseerde techniek, is gebaseerd op het nemen van een volledige point-in-time snapshot kopie van de data set om deze daarna naar een andere disk array te transporteren (uiteraard weer van dezelfde diskleverancier).

In geval van data-corruptie van de Active Directory biedt Virtual Disk Service (VDS) de mogelijkheid om de backup shadow kopieën van ongecorrumpeerde volumes voor gebruik beschikbaar te maken. Doch bij een SAN is het nodig dat opslagruimten van servers van elkaar gescheiden blijven en is het nodig dat een shadow kopie van de originele data 'unmasked' getransporteerd kan worden voor gebruik op de andere server. Daarvoor is een oplossing nodig die door een hardware provider wordt geboden. VDS maakt het mogelijk door de op het SAN opgeslagen shadow kopieën te 'unmasken'. Dat wil zeggen beschikbaar te maken op de server met de gecorrumpeerde data. Daarbij wordt de read-only status veranderd in read/write en tenslotte worden de volumes voor gebruik 'gemount'. Het is eigenlijk meer een virtueel- dan een fysiek proces, omdat de data altijd op de storage array binnen het SAN blijft opgeslagen.

## VSS met AD in de praktijk

In het volgende voorbeeld is Active Directory geïnstalleerd op twee servers binnen een SAN. Het proces voor het maken van shadow kopieën begint op het moment dat de backup-applicatie (Requestor) de VSS-service vraagt om een kopie van de System State Volumes te maken. De VSS-service, die als een coördinator fungeert, informeert elk van de System State writers, waaronder de Active Directory, om zich voor te bereiden voor het maken van een shadow kopie bewerking. Zodra de data beschikbaar is voor het backup-proces informeert de writer de VSS service coördinator, die op zijn beurt de informatie doorstuurt naar de backup requestor. De requestor stopt tijdelijk de Active Directory I/O-schrijfoperaties naar disk, om de provider een paar seconden de tijd te geven een single point-in-time shadow kopie van de betreffende volumes te maken. Zodra de shadow kopie van een van de volumes klaar is, kan de VSS service worden geïnstrueerd om de verbinding tussen het origineel en de shadow kopie te verbreken. De shadow-kopie bevindt zich nu in een 'read-only' staat en de server kan met de originele data op de

normale wijze doorgaan met de productie. Vanaf dit moment heeft de shadow kopie geen enkele relatie meer met een server en blijft op de storage array beschikbaar, tot het moment dat het nodig is voor een recovery operatie.

## LUN masking en Unmasking

Alhoewel servers in een opslagnetwerk gebruik maken van een gemeenschappelijk opslagsysteem heeft elke server slechts beperkt toegang tot een of meer specifieke LUN's. Door van de hardware-provider op het SAN gebruik te maken, kunnen point-in-time shadow copies voor het gebruik daarvan virtueel 'getransporteerd' worden naar een andere server. Dit wordt mogelijk gemaakt door het proces van 'masking' en 'unmasking' met de VDS DISKRAID utility. De gecorrumpeerde LUN's worden via het DISKRAID mask commando 'offline' geplaatst. Vervolgens worden de shadows kopie LUN's 'unmasked' en geconverteerd van read-only- naar read/write status en tenslotte 'gemount' op de uitgevallen server (die daarna opnieuw geboot moet worden om het deze weer in het productieproces op te nemen).

We zien dus dat op deze manier een snel herstel van een uitgevallen Active Directory Server mogelijk is. De restore van tape die normaal uren (of zelfs dagen) zou kunnen duren, wordt nu aanmerkelijk bekort tot enkele minuten (uiteraard afhankelijk van de grootte van de AD-database).

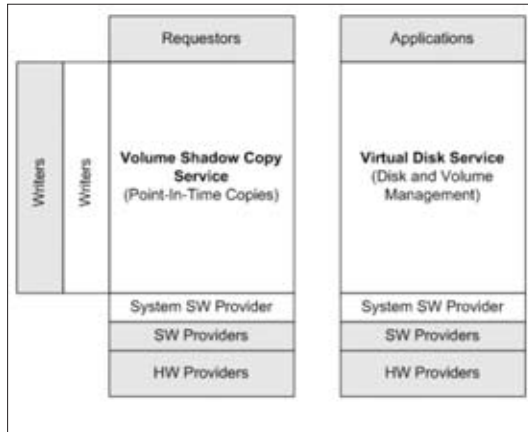
## Bescherming van Snapshots

VSS is een prima manier om bestanden te herstellen, verwijderde bestanden terug te halen of ongewilde veranderingen terug te draaien. Maar wat in het geval als de disk, of server niet langer beschikbaar is? En hoe staat met die bestanden en applicaties waarbij het verlies van data absoluut uit den boze is en die binnen een minuut een recovery vereisen?

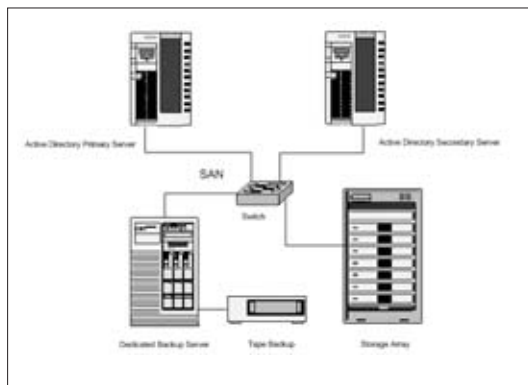
Dat is op basis van VSS alleen niet mogelijk. Native VSS is namelijk niet geschikt om een volledig volume te herbouwen, omdat snapshots geen complete data-image vertegenwoordigen: alleen de kopieën van de veranderde data. VSS kan wel snapshots creëren maar is niet geschikt voor het bieden van een Disaster Recovery (DR) oplossing.

Het bijhouden van kopieën van bestanden op de locale server biedt dus wel een uitstekende voorziening voor recovery maar beschermt snapshots niet in het geval er een disk of een server uitvalt. Snapshot kopieën zijn namelijk net zo kwetsbaar als de originele kopieën. Bij uitval is men nog altijd genooddaakt om alle data te restoren van de laatst beschikbare

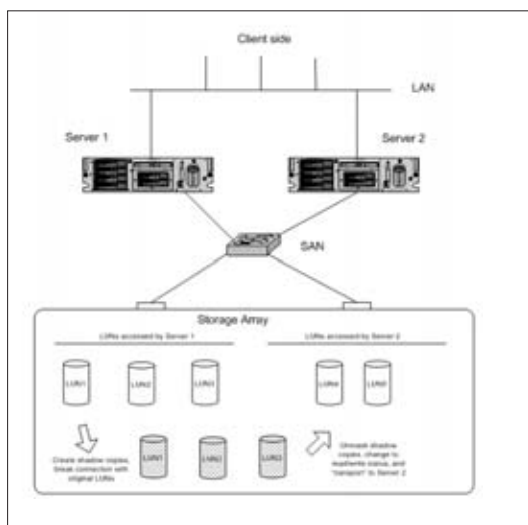
tape, wat een moeizaam en tijdrovend proces blijft (nog daargelaten of deze wel in 100% van alle gevallen lukt). Om deze reden is het aan te bevelen om op een remote locatie een backup van alle data en de snapshots te bewaren. ☒



☒ Afbeelding 1 » VSS en VDS architectuur (bron Microsoft)



☒ Afbeelding 2 » SAN configuratie met AD-servers (bron Microsoft)



☒ Afbeelding 3 » Mask en Unmask shadow copies (bron Microsoft)