

Is het oude back-up tape al een archief?

De titel suggereert het al: een oud back-up bandje is geen archief-opslag laat staan een volwaardig archief dat mogelijk ook nog aan wettelijke eisen moet voldoen. Niettemin horen



we van klanten die denken dat ze met hun tape drives of – erger nog – met een handvol externe USB-harddisks een archief kunnen opbouwen en beheren. Wij zullen in dit document uitleggen hoezo tape drives meestal de juiste keuze voor back-ups maar nagenoeg nooit geschikt voor archival storage zijn en hoezo Silent Cubes juist wél ideaal voor archieven maar (op één uitzondering na) ongeschikt voor back-ups zijn.

Om de verschillen tussen een digitaal archief en een data back-up goed in kaart te brengen is het handig eerst te kijken wat de doelen van een archief en van een back-up zijn. Met andere woorden: hoezo back-uppen wij en hoezo archiveren wij?

De intentie van een back-up is in het geval van dataverlies de verloren informatie weer toegankelijk te maken en dit liefst zo snel mogelijk om de bedrijfscontinuïteit zo min mogelijk te storen. Een back-up is dus altijd een kopie van originele gegevens waarbij deze meestal met de tijd wijzigen.

De intentie van een archief is een verzameling van onveranderlijke gegevens die langdurig bewaard moeten worden te creëren. Een archief is dus geen kopie van gegevens maar bevat de originele gegevens. De snelheid waarmee uit het archief gelezen kan worden is minder bedrijfskritisch, wél is het belangrijk dat het archief eenvoudig en snel doorzocht kan worden.

De voornaamste verschillen tussen een back-up en een archief zijn dus dat een back-up gemaakt wordt voor snelle data recovery en een archief gemaakt wordt voor het langdurige bewaren van originele gegevens die meestal niet dagelijks gebruikt worden.

Dit heeft invloed op de geschikte hardware en software, het is bijvoorbeeld geen probleem dat back-up hardware en software vaak in propriëitair formaat op speciale hardware schrijft – het gaat immers alleen om een kopie van originele gegevens. Bij een technologische verandering kunnen wij gewoon een nieuwe kopie maken.

Bij een archief is het echter fataal als wij onze originele gegevens van een aantal jaren geleden niet meer kunnen benaderen omdat de hiervoor nodige hardware of software niet meer beschikbaar is. Archiven moeten daarom zogenaamde vendor neutral archives – d.w.z. onafhankelijk van fabrikanten – zijn.

Nu is het niet zo dat archieven en back-ups helemaal niets gemeen hebben, er zijn natuurlijk wel enkele gemeenschappelijke eisen: aangezien de hoeveelheid gegevens permanent toeneemt moeten de capaciteiten van zowel ons archief als van ons back-up systeem uit te breiden zijn. Uit kostenoverwegingen moeten beide systemen met weinig administratief werk te beheren zijn en ze moeten zo energiezuinig als mogelijk zijn. Dit laatste geldt in bijzondere mate voor ons archief: aangezien een archief voor lange tijd, vaak voor decennia en soms voor eeuwig is aangelegd heeft het energieverbruik een bijzonder zwaarwegend aandeel in de TCO (total cost of ownership).

De onderstaande tabel geeft een overzicht over de eigenschappen van back-up en archief:

	Archief	Back-up
<i>Snel terugzetten van grote hoeveelheden data</i>	Nee, gegevens worden naar behoefte rechtstreeks benaderd	Ja, om bedrijfscontinuïteit in geval van calamiteit te waarborgen
<i>Snel doorzoekbaar</i>	Ja, willekeurige gegevens moeten in willekeurige volgorde op te vragen zijn	Nee, het is bekend welk back-up medium te gegevens bevat die terug gezet moeten worden
<i>Bevat originele gegevens</i>	Ja, altijd	Nee, altijd een kopie
<i>Herschrijfbaar</i>	Nee, moet WORM zijn om gegevensintegriteit te waarborgen	Ja, de originele gegevens waar het back-up de kopie van is kunnen veranderen
<i>Langdurig</i>	Ja, dit is kenmerkend voor een archief	Meestal niet, de kopie moet alleen zolang beschikbaar zijn als het origineel geldig is
<i>Bevat onveranderlijke gegevens</i>	Ja, dit is kenmerkend voor een archief	Meestal niet
<i>Capaciteit uitbreidbaar</i>	Ja, de hoeveelheid gegevens zal naar verwachting verder toenemen	Ja, de hoeveelheid gegevens zal naar verwachting verder toenemen
<i>Weinig administratie</i>	Ja, omdat archieven over lange periodes beheert moeten worden is het belangrijk dat de kosten hiervoor laag zijn	Ja, omdat back-ups regelmatig (dagelijks) gemaakt worden is het belangrijk dat het beheer eenvoudig is
<i>Fysiek gescheiden van het datacenter / bedrijfsnetwerk</i>	Nee, integendeel moet een archief meestal rechtstreeks te benaderen zijn	Ja, om bij een lokale destructie (brand, sabotage, aardbeving, ...) nog over een kopie te beschikken
<i>Energiezuinig</i>	Ja, omdat archieven over lange periodes actief zijn is het bijzonder belangrijk het energieverbruik laag te houden	Niet meer dan bij andere apparatuur die dagelijks gebruikt wordt
<i>Regelmatische gegevenscontrole</i>	Ja, om dataintegriteit over een lange periode te waarborgen moeten de gegevens regelmatig gecontroleerd worden	Nee, het moet alleen zeker gesteld zijn dat back-up en restore altijd werken
<i>Niet propriëtair, dus opgebouwd uit standaard componenten met standaard software protocollen</i>	Ja, omdat de gegevens ook na lange tijd nog benaderbaar moeten zijn moet een archief vendor neutral zijn, dus onafhankelijk van bepaalde hardware, software of besturingssystemen die mogelijk ophouden te bestaan	Nee, omdat een back-up altijd een momentele kopie van actuele gegevens is kan eenvoudig overgegaan worden op andere technologie

Om te bepalen welke storage technologie voor welke toepassing geschikt is kunnen wij deze tabel gebruiken om de verschillende media te toetsen. Als wij dit doen zullen wij – niet echt verbazend – zien dat een technologie, die speciaal voor een bepaalde toepassing ontwikkelt is de eisen voor deze toepassing ook het beste vervuld.

Zo zijn tape en RDX systemen goed geschikt voor back-up, want grote hoeveelheden data kunnen snel "bulk" terug gezet worden, de media zijn herschrijfbaar en daardoor meerdere keren te gebruiken, tape en RDX systemen zijn door aanschaf van extra cartridges goedkoop uit te breiden, en ze zijn relatief energiezuinig omdat ze alleen voor het schrijven en lezen van gegevens energie verbruiken, het bewaren van de gegevens zelf (bandje ligt in de kast) kost geen energie.

Tape en RDX systemen zijn echter ongeschikt voor archiefdoeleinden want ze voldoen aan veel van de eisen voor archieven niet, bijv. het WORM criterium, de regelmatige gegevenscontrole, de snelle doorzoekbaarheid, de duurzaamheid van de opslag.

De Silent Cubes zijn speciaal ontwikkeld met het oogmerk op archival storage, ze zijn dan ook niet geschikt voor back-up, alleen al omdat ze niet herschrijfbaar zijn. Maar ze zijn bij uitstek geschikt voor dat waar ze voor gemaakt zijn: archief opslag. Silent Cubes zijn hardwarematig beveiligd tegen het overschrijven van de gegevens, ze zijn dus een écht WORM systeem, gemaakt voor het langdurig bewaren voor gegevens, die regelmatig en automatisch op integriteit gecontroleerd worden. Daarbij is een Silent Cube uiterst energiezuinig en opgebouwd uit standaard componenten met standaard netwerkprotocollen.

Optische media zoals CDs, DVDs en BDs zijn ontwikkelt voor de digitale distributie van (mediale) gegevens. Hiervoor zijn ze ook zeer geschikt omdat ze goed transportabel zijn en een grote oppervlakte voor bedrukkingen hebben. Optische media zijn echter ongeschikt voor back-up omdat ze (op uitzonderingen na) niet herschrijfbaar zijn en vaak niet voldoende capaciteit bieden. Ze zijn beperkt geschikt voor archief-opslag: het zijn WORM media, dus gegevens kunnen niet overschreven worden, maar de capaciteit is beperkt en een regelmatige gegevenscontrole is omslachtig.

RAID systemen zijn ongeschikt voor zowel back-up en archief. RAID systemen zijn ontwikkelt om snelle, relatief betrouwbare toegang tot permanent veranderlijke gegevens (bijv. databases) te geven, voor back-up en archief zijn ze te duur, te slecht uit te breiden en te energie-intensief. Voor archief doeleinden zijn ze tevens veel te onbetrouwbaar.

De volgende tabellen geven een overzicht over de verschillende systemen en hun geschiktheid voor back-up en archief.

Eisen aan een back-up systeem:

	RAID	Optisch	Tape / RDX	Silent Cube
<i>Snel grote hoeveelheden data terugzetten</i>	Ja	Nee	Ja	Ja
<i>Herschrijfbaar</i>	Ja	Nee	Ja	Nee
<i>Capaciteit uitbreidbaar</i>	Ja, maar zeer duur	Ja	Ja	Ja
<i>Weinig administratie</i>	Ja	Nee, jukeboxen en drivers moeten beheert worden	Nee, tape libraries en drivers moeten beheert worden	Ja
<i>Energiezuinig</i>	Nee	Ja	Ja	Ja
<i>Mogelijkheid tot fysieke scheiding</i>	Ja, met een tweede installatie	Ja	Ja	Ja, met een tweede installatie

Eisen aan een archief systeem:

	RAID	Optisch	Tape / RDX	Silent Cube
<i>Snel doorzoekbaar</i>	Ja	Gemiddeld	Nee	Ja
<i>Niet herschrijfbaar (WORM)</i>	Nee	Ja	Nee	Ja
<i>Langdurig veilig</i>	Nee, alleen korttijdig.	Ja, maar geen redundantie	Nee	Ja, viervoudig redundant
<i>Capaciteit uitbreidbaar</i>	Ja, maar zeer duur	Ja	Ja	Ja
<i>Weinig administratie</i>	Ja	Nee, jukeboxen en drivers moeten beheert worden	Nee, tape libraries en drivers moeten beheert worden	Ja
<i>Energiezuinig</i>	Nee	Ja	Ja	Ja
<i>Regelmatische gegevenscontrole</i>	Nee, alleen met additionele software	Nee, alleen met additionele software	Nee, alleen met additionele software	Ja, door ingebouwde digitale audit
<i>Niet propriëtair</i>	Nee, RAID systemen zijn altijd vendor afhankelijk	Nee, voor jukeboxen zijn vendor afhankelijke drivers voor elk besturingssysteem nodig	Nee, voor tape libraries zijn vendor afhankelijke drivers voor elk besturingssysteem nodig. Tape formaten zijn verschillend en het opslagformaat is afhankelijk van de gebruikte software	Ja, met standaard componenten worden de Silent Cubes via standaard netwerkprotocollen (TCP/IP, Ethernet) als standaard share (SMB, CIFS, NFS) vanuit elk besturingssysteem aangesproken. Geheel vendor neutral dus.

Het is overduidelijk dat het Silent Cube systeem het meest geschikte opslag systeem voor archieven is. Dit is ook niet verbazingwekkend aangezien dit systeem speciaal voor archieven ontwikkeld is.

Voor back-up zijn in de meeste gevallen tape of RDX systemen het meest geschikt alhoewel hier wel uitzonderingen zijn: bij een zeer bedrijfskritisch omgeving kan het zinvol zijn een tweede RAID systeem als back-up-eenheid in te zetten en de hoge energie- en aanschafkosten voor lief te nemen.

Uiteraard moet ook van een archief een fysiek gescheiden back-up in de vorm van een kopie gemaakt worden, anders kunnen zulke rampen gebeuren als bij de brand van de bibliotheek van Alexandrië of onlangs bij het instorten van het stadsarchief van Keulen door de bouw van de metro. In beide gevallen zijn grote hoeveelheden erfgoed voor altijd verloren gegaan.

Aan het back-up van een archief zijn bijna dezelfde eisen gesteld als aan alle andere back-ups, maar met één uitzondering: omdat de originele gegevens (het archief) nooit wijzigen hoeft het back-up van een archief niet herschrijfbaar te zijn. Dit maakt wederom de Silent Cube het meest geschikte systeem voor een archival back-up, vooral omdat hierdoor direct een compleet zelfstandig werkend

archief ontstaat dat in geval van een calamiteit onmiddellijk ter beschikking staat, zonder dat eerst gegevens van een back-up medium terug gezet moeten worden. Maximale redundantie dus.

De conclusie is dat voor back-up systemen meestal de oude tapes of de moderne RDX systemen gebruikt kunnen worden en voor archief systemen altijd de Silent Cubes de juiste oplossing zowel voor de originele data als voor de back-up kopie zijn.

©2013, Comex sales bv, www.comex.eu

Dit document mag zonder uitdrukkelijke toestemming vrij gepubliceerd en/of verdeeld worden, echter uitsluitend in zijn geheel, onder benoeming van de bron en onder verwijzing naar de website www.comex.eu. Voor publicatie van delen uit dit document kunt u toestemming vragen onder info@comex.eu