

Whitepaper

Duurzame IT – gegevens langdurig en energie-efficiënt archiveren



Voor de archivering van data gelden speciale eisen. In vergelijking met de gebruikelijke investeringsperioden van drie jaar hebben energieverbruik en onderhoud van het systeem voor archiefopslag een hogere prioriteit.

Ook voor de hardwarecomponenten en leveranciers moeten strenge selectiecriteria worden gehanteerd. In deze whitepaper nemen we al deze aspecten onder de loep en vergelijken we technologieën voor de lange-termijnopslag van data.

Gegevensopslag voor de lange termijn

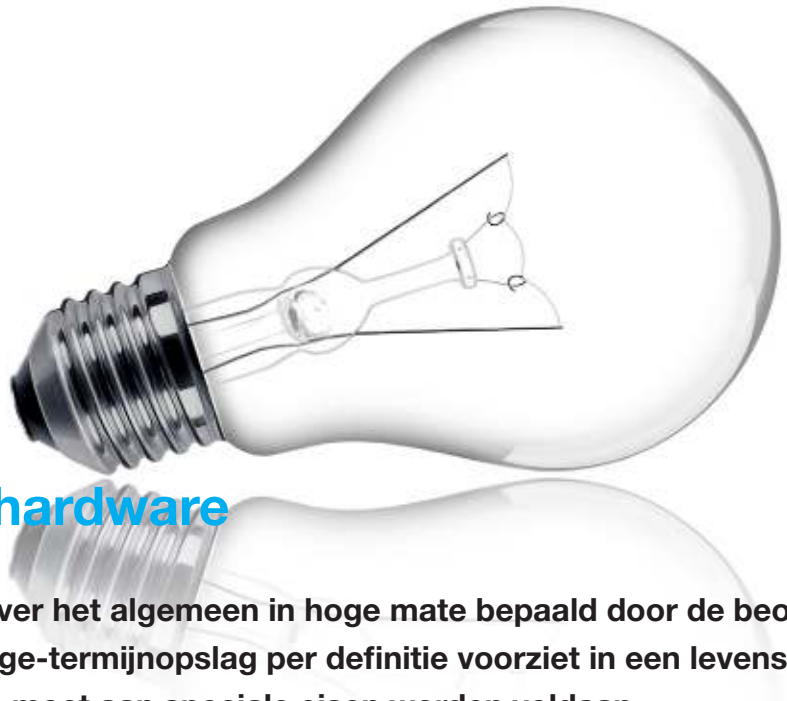
Het archiveren van data omvat het verplaatsen van gegevens die in de productieomgeving niet langer of nog maar zelden benodigd zijn. Hoewel de trend zich in de richting beweegt van het 'actieve archief' dat in tegenstelling tot klassieke, tapegebaseerde offline archieven snelle toegang tot gegevens kan bieden op basis van random access, blijft het voornaamste doel om data jarenlang of zelfs decennialang op veilige en betrouwbare wijze op te slaan. Het kostenaspect speelt eveneens een belangrijke rol. Omdat systemen voor het archiveren van data meestal een langere gebruiksduur kennen als de in de IT-sector algemeen geldende drie tot vijf jaar, spelen factoren als energieverbruik, onderhoudsoverhead en service level agreements een belangrijkere rol. Investeringskosten hebben in dit geval echter een incrementeel karakter: de capaciteit wordt stap voor stap uitgebreid. De vervanging van opslagsystemen wordt zoveel mogelijk vermeden met het oog op de almaar groeiende datavolumes en de noodzaak om data te migreren.

Bij langdurige gegevensopslag komen dus de volgende aspecten kijken:

Hardware: Duurzaamheid en beschikbaarheid, CO₂-voetafdruk, energie-efficiëntie, uitbreidbaarheid, bescherming tegen downtime

Software: Betrouwbaarheid, certificering, compatibiliteit met standaarden, datamigratie, beveiliging tegen cyberaanvallen

Service: Onderhoudsoverhead, voorspelbaarheid van kosten, beschikbaarheid van vervangende onderdelen



Langdurig inzetbare hardware

De keuze van de hardware wordt over het algemeen in hoge mate bepaald door de beoogde archiveringstoepassing. Omdat lange-termijnopslag per definitie voorziet in een levensduur van ettelijke jaren of zelfs decennia moet aan speciale eisen worden voldaan.

Veilige opslag

Bij back-ups worden gegevens gekopieerd om na een bepaalde periode weer te worden gewist. Bij de archivering van gegevens daarentegen worden gegevens verplaatst, zodat ze niet langer in het bronstelsel aanwezig zijn. Als de gegevens in het archief worden gewist, gaan ze dus permanent verloren. Data moeten daarom op zodanige wijze worden opgeslagen dat gegevensverlies tijdens de beoogde bewaartermijn effectief is uitgesloten.

Er kunnen diverse maatregelen worden getroffen tegen gegevensverlies, zoals bescherming tegen de uitval van afzonderlijke gegevensdragers (kopieën in het geval van tape en redundantie in het geval van harde schijven), bescherming tegen downtime van complete systemen (georedundantie, waar mogelijk opslag op een veilige externe locatie van gegevensdragers) en bescherming tegen het per ongeluk verwijderen of overschrijven van data (WORM). Meestal is het niet mogelijk om met standaard IT-oplossingen aan deze criteria te voldoen. Dit vraagt om speciale archiefopslag.

Een belangrijk aspect in dit verband, zeker als er gebruik wordt gemaakt van harde schijven, is de preventie van gecorreleerde uitval ('epidemic failure'). Voor de meeste storage arrays wordt gebruikgemaakt van een groot aantal schijven van hetzelfde type, vaak zelfs met leveranciersspecifieke firmware. Aangezien schijven die deel uitmaken van dezelfde productiebatch met dezelfde snelheid verouderen, bestaat bij de uitval van de eerste exemplaren het gevaar van 'massale sterfte'. Dit vraagt om een tijdige vervanging van het complete opslagsysteem.

Schaalbaarheid

Omdat gegevens in het archief normaliter niet worden verwijderd, zal de benodigde capaciteit gestaag groeien. Een uitzondering hierop zijn archieven die louter dienen om te voldoen aan de eisen van de wet- en regelgeving, zonder dat organisaties enig belang hebben bij de opslag van de gegevens in kwestie. In dat geval moeten gegevens na het verlopen van de wettelijke bewaartermijn tijdig worden verwijderd of vernietigd. Om deze reden wordt in dit verband

nog altijd vaak gebruikgemaakt van opslag op optische media of zelfs papieren documenten die uiterst vatbaar voor vernietiging zijn.

Alle andere archieven dienen voor het bewaren van informatie. Data heet tegenwoordig het nieuwe goud te zijn en wordt daarom niet gewist. Dankzij de opkomst van nieuwe oplossingen op het gebied van artificial intelligence en big data-analyse bestaat namelijk de kans dat grote datasets in een later stadium van nut kunnen zijn.

Systemen voor archiefopslag moeten dus optimaal schaalbaar zijn. Omdat de parallele gegevensdoorvoer in de regel niet hoeft mee te groeien met de opslagcapaciteit spreekt men wel van scale up storage: de onderliggende infrastructuur moet zo gestroomlijnd mogelijk zijn, maar wel ondersteuning bieden voor een maximale uitbreiding van de opslagcapaciteit. Scenario's waarbij ettelijke petabytes zijn gemoeid zijn zonder meer realistisch.

Energie- en milieubeheer

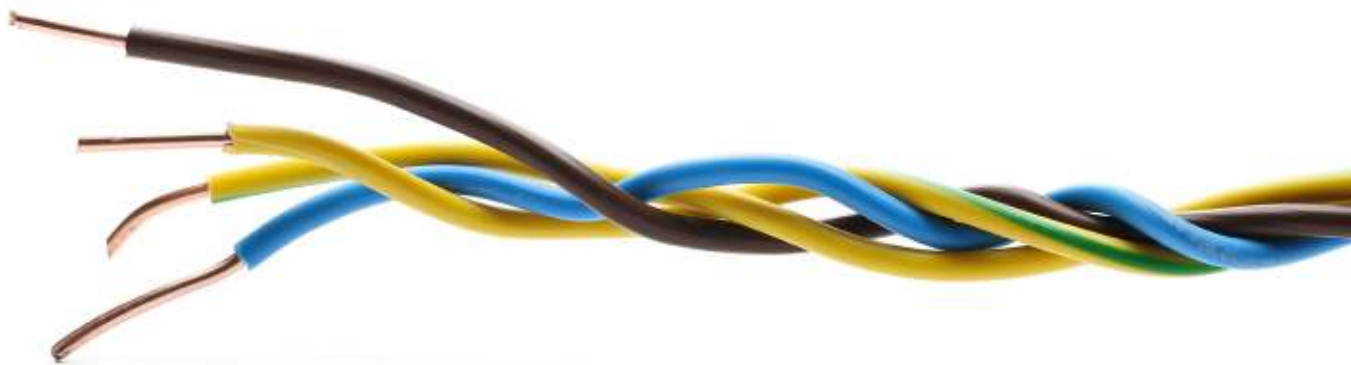
Hoe langer een systeem in gebruik is, hoe groter het aandeel van de energiekosten in de totale kosten. Omdat gearchiveerde gegevens gewoonlijk steeds minder vaak worden opgevraagd naarmate de opslagduur toeneemt, zou de opslag van zeer oude gegevens gepaard moeten gaan met zo min mogelijk of liefst helemaal geen energieverbruik. Dit kan worden gerealiseerd met behulp van

tiering, oftewel de opslag van zeer oude data op louter offline opslagmedia. Dit vraagt echter om een extra beheerinspanning en aanvullende opslagsystemen, waarmee de complexiteit en overhead verder toenemen. Een betere oplossing is om gebruik te maken van systemen met een modulaire opzet en efficiënt energiebeheer.

Een ander aspect is de CO₂-voetafdruk van de gebruikte componenten. De materiaalselectie en compensatie voor de CO₂-uitstoot die gepaard gaat met het produceren, leveren en afdanken daarvan dragen stuk voor stuk bij aan de CO₂-voetafdruk.

Duurzaamheid en beschikbaarheid

Hoewel oude systemen minder energie-efficiënt zijn dan de laatste generatie, is het aanhouden daarvan vaak beter voor het milieu dan het vroegtijdig afdanken daarvan. De duurzaamheid van onderdelen en de beschikbaarheid van eventueel benodigde vervangende onderdelen voorkomen het ontstaan van schadelijk elektronisch afval en dragen daarmee direct bij aan een beter milieu. Het is vooral de keuze van de gegevensdragers die het duurzaamheidsgehalte van de archiefoplossing bepaalt. Als versies elkaar in hoog tempo opvolgen en er sprake is van beperkte terugwaartse compatibiliteit, gaat dat gepaard met de noodzaak van kostbare en foutgevoelige migraties van data en meer schadelijk afval.



Software die aan duurzaamheid bijdraagt

Het gebruik van de juiste software kan het verschil maken tussen bruikbare hardware en waardeloos elektronisch schroot. Als de software die voor een hardwarevoorziening wordt gebruikt niet langer wordt doorontwikkeld of de ondersteuning voor de kerninfrastructuur wordt stopgezet, verliest het complete systeem daarmee zijn nut.

Standaarden, certificeringen en datamigratie

Een belangrijk criterium voor een duurzaam softwareconcept is de ondersteuning voor standaarden. Dit is de enige manier om ervoor te zorgen dat de vervanging van een opslagsysteem relatief probleemloos verloopt. Compatibiliteit met branchestandaarden waarborgt de uitwisselbaarheid en het behoud van gegevens, met inbegrip van alle data van de betrokken software. De eventueel benodigde migratie van data is alleen een haalbare kaart als aan de toepasselijke standaarden wordt voldaan.

Veel archiveringstoepassingen vragen om allerlei certificeringen. Deze zijn nodig om aan de wettelijke eisen te voldoen. Zonder certificering zal zelfs het technisch meest vernuftige archiefsysteem niet auditproof zijn. Omdat dergelijke certificeringen aanzienlijke financiële en organisatorische overhead voor fabrikanten met zich meebrengen, hebben zij belang bij een langdurige inzetbaarheid van hun oplossing en daarmee de jarenlange doorontwikkeling van de gebruikte software.

Betrouwbaarheid en veiligheid

Er bestaan geen algemeen geldende criteria voor de betrouwbaarheid van software. Omdat moderne opslagsystemen geen volledig autonome en in eigen beheer ontwikkelde systemen zijn, maar uit talloze softwarecomponenten zijn opgebouwd, blijft de regelmatige installatie van updates en beveiligingspatches zelfs na ettelijke decennia absoluut noodzakelijk. De ondersteuning voor branchestandaarden vraagt om de integratie van open source libraries. En hoe groter de schaal waarop die worden ingezet, hoe groter de kans dat cybercriminelen misbruik maken van kwetsbaarheden in de beveiliging. Een snelle reactie van de fabrikant is even wenselijk als transparante communicatie.

De functionaliteit van de beheersoftware kan eveneens bijdragen aan het duurzaamheidsgehalte van de archiefoplossing. Als de beheerinterface de mogelijkheid biedt om uitgebreide beveiligingsmaatregelen te configureren maakt dat een einde aan de kans op gegevensverlies als er bijvoorbeeld een gegevensdrager uitvalt. Organisaties kunnen zich daarnaast tegen de gevolgen van ransomware-aanvallen indekken met behulp van speciale geautomatiseerde procedures zoals continuus snapshots en hun gegevens daarmee ook buiten het WORM-archief beschermen.



Langdurige service

Hoe complexer een IT-omgeving wordt, hoe belangrijker de service level agreements. Het gebruik van SLA's is voor de meeste bedrijfssystemen verplicht. De mogelijkheid om de onderhoudskosten op de lange termijn te plannen biedt duidelijkheid over het totale kostenplaatje.

Onderhoudskosten

Opslagsystemen met weinig mechanische onderdelen zijn over het algemeen minder vatbaar voor storingen. Flashgebaseerde systemen bevatten helemaal geen mechanische onderdelen, maar zijn in de meeste gevallen te duur voor het archiveren van gegevens. Daarom is het belangrijk om de voordelen en nadelen tegen elkaar af te wegen. Vaste schijven bevatten draaiende magnetische schijven die gevoelig zijn voor schokken. Snelle serverschijven moeten continu in bedrijf blijven om te voorkomen dat de schrijf-/leeskop verharst.

Voor tape-opslag wordt gebruikgemaakt van flinterdunne magneetbanden die in mechanisch complexe tape drives worden ingezet. Zodra een bepaalde opslaglimiet is bereikt wordt gebruikgemaakt van tape libraries die gebruikmaken van robotica voor de wissel van opslagmedia.

Kosten en vervangende onderdelen

Onderhoudsbezoeken kosten geld, gaan gepaard met reistijd en vragen vaak ook om een gedeeltelijke of volledige uitschakeling

van het archiefsysteem. Dit is niet alleen vanuit milieuoogpunt onwenselijk. Het is eenvoudiger om gebruik te maken van systemen die organisaties zelf kunnen onderhouden, bijvoorbeeld door het vervangen van individuele gegevensdragers. De beschikbaarheid van componenten blijft wat dat betreft ook na tientallen jaren een belangrijke rol spelen. Als systemen zodanig zijn ontworpen dat daarvoor ook latere modellen gegevensdragers kunnen worden gebruikt, zal de onderhoudsoverhead geheel een stuk lager uitvallen.

Na het verstrijken van de gebruikelijke levensduur van drie tot vijf jaar hanteren veel fabrikanten vaak hogere tarieven, zeker als het om onderhoud gaat. En daar is een goede reden voor. Want hoe langer systemen worden gebruikt, hoe groter de kans dat ze onderhoud vereisen. Bovendien zijn fabrikanten meer geïnteresseerd in de verkoop van nieuwe oplossingen dan een langdurig gebruik van bestaande systemen. Om deze reden worden systemen vaak uit de rolatie genomen en afgedankt terwijl ze in feite prima functioneren. Dat komt omdat de onderhoudskosten buitenproportioneel hoog zijn geworden.

Een vergelijkend overzicht van technologieën

Al naargelang de eisen zijn er diverse technologieën voor archivering geschikt. Tape, schijfopslag en cloudoplossingen wedijveren met elkaar en beloven elk weer andere voordelen.

Tape

In het verleden werd voor de archivering van data vaak gebruikgemaakt van tape vanwege de lage kosten van dit opslagmedium. Hoewel tape met zijn lage opslagkosten per TB inderdaad alle concurrentie achter zich laat, verliest dit medium steeds meer aan relevantie voor opslagdoeleinden. De reden hiervan is dat data niet of slechts in beperkte mate opvraagbaar is vanwege het strikt lineaire karakter van tape. Daarom is het voortdurend nodig om data over te zetten naar toegankelijke opslagmedia zoals harde schijven of flashgeheugen. Daarmee stijgen de overhead, kosten en complexiteit. In sommige gevallen resulteert het gebruik van betaalbare tape zelfs in hogere kosten omdat er aanvullende systemen en beheerssoftware nodig zijn. Vanwege het hoge mechanische gehalte van tapegebaseerde opslagsystemen en alle componenten die daarbij komen kijken is er sprake van een relatief hoge onderhoudsoverhead. En die kan gedurende de levensduur van de oplossing flink toenemen. Daarom voorzien onderhoudscontracten voor de lange termijn vaak na een paar jaar in fors hogere tarieven. Dit maakt het moeilijk om de kosten vooraf te berekenen.

Vaak wordt tape aangeprezen als een duurzame oplossing vanwege de terugwaartse compatibiliteit die dit opslagmedium biedt. Maar sinds LTO-8 is die alleen nog maar van toepassing op één generatie, terwijl voorheen het dubbele gold. Wie dus wil profiteren van de hogere opslagcapaciteit van moderne LTO-media loopt daarmee het risico dat er regelmatig data moet worden gemigreerd, om nog maar niet te spreken van de kostbare en arbeidsintensieve vervanging van tape drives.

Wat de opslagbeveiliging betreft zijn tapes redelijk geschikt als individuele gegevensdragers. Er zijn echter altijd kopieën nodig om bescherming te bieden tegen storingen. Dit gaat gepaard met een verveelvoudiging van de benodigde capaciteit en een behoorlijk tijdrovend proces.

Oudere generatie tapes zijn echter meestal nog jarenlang beschikbaar en uiterst betaalbaar. Tapes zijn compact, en veel IT-managers hebben het beheer van hun tapegebaseerde opslagsystemen in de loop der jaren geautomatiseerd. Als het niet te verwachten valt dat de gearchiveerde gegevens na de beoogde bewaartermijn nog worden gebruikt, of slechts in speciale gevallen, dan is het mogelijk voor de archiefopslag gebruik te blijven maken van tape.



Cloud-archieven

Er komen steeds meer cloudoplossingen voor de archivering van data op de markt. Sommige namen verraden de beoogde toepassing. Zo noemt Amazon zijn oplossing voor langdurige opslag 'Glacier' (gletsjer). De focus ligt dus op langdurige opslag. Het opvragen van de data kan al naargelang de ingestelde beleidsregels met lange wachttijden gepaard gaan. Er zijn aanvullende lokale opslagsystemen nodig voor het waarborgen van de beschikbaarheid van minstens een deel van de data.

De kostenstructuur van cloudarchieven ziet er ook anders uit dan die voor opslagsystemen op locatie. Organisaties zijn relatief weinig geld kwijt aan de opslag van data, maar betalen relatief veel voor het opvragen daarvan. Als afzonderlijke bestanden moeten worden hersteld blijven de kosten beheersbaar. Deze storage-oplossingen zijn echter minder geschikt voor gebruik als actief archief.

Wat men niet moet vergeten, is dat ook cloudarchieven gebruikmaken van opslaghardware. Al naargelang het gewenste beschikbaarheidsniveau worden gegevens opgeslagen in

omvangrijke schijfgebaseerde storage arrays of in tape libraries. Gebruikers van cloudoplossingen wordt vaak geen inzicht geboden in de CO₂-voetafdruk en compensatie daarvoor of de energie-efficiëntie en levensduur van de gebruikte hardware, en al helemaal niet als ze oplossingen van internationaal actieve leveranciers afnemen.

Vragen die vaak nog moeilijker te beantwoorden vallen, zijn hoe de beveiliging is geregeld en of er bepaalde certificeringen nodig zijn. Vaak worden gegevens opgeslagen buiten de Europese Unie. In de meeste gevallen moet men afgaan op de beloftes van de aanbieder van de cloudoplossing. Beloftes van extreem hoge beschikbaarheid moeten klanten een gevoel van zekerheid bieden, maar zodra er problemen rond de toegankelijkheid van data optreden zullen er vaak meerdere diensten tegelijkertijd uitvallen. Ook als de opslagdienst zelf niet onder die storing te lijden heeft.

De belangrijkste voordelen van cloudopslag zijn eenvoudige schaalbaarheid en een lage onderhoudsoverhead.





Silent Cubes en Silent Bricks: unieke schijfgebaseerde opslag

FAST LTA biedt met Silent Cubes en Silent Bricks speciale systemen voor de archivering van data aan. Silent Cubes en Silent Brick WORM zijn opslagvoorzieningen die gebruikmaken van hardwarematige beveiliging volgens het write once read many (WORM)-principe. Dit draagt bij aan auditproof archivering en overeenstemming met de eisen van de AVG. Silent Bricks kunnen functioneren als omvangrijk virtual tape library (VTL)-archief met een opslagcapaciteit van ettelijke petabytes en bieden daarmee een alternatief voor tape libraries.

De software die aan de basis van deze opslagsystemen staat heeft zich decennialang in de praktijk bewezen en wordt voortdurend verder ontwikkeld. De allereerste Silent Cubes werden in 2008 in gebruik genomen. Sindsdien doen ze bij duizenden klanten dienst. En in de meeste gevallen doen ze de naam 'silent' eer aan door onopgemerkt en zonder de noodzaak van noemenswaardig onderhoud op de achtergrond

te functioneren. Aangezien FAST LTA deze producten aan strikte kwaliteitscontroles onderwerpt, hebben ze niet te lijden onder kwetsbaarheden zoals Log4j.

Schijfgebaseerde systemen bieden random access leesbewerkingen en bieden razendsnelle opslagprestaties. Dankzij beveiliging op basis van erasure coding kunnen van de 12 gegevensdragers in de Silent Cube of de Silent Brick vier gegevensdragers uitvallen zonder dat er kans op gegevensverlies ontstaat. Bovendien wordt er binnen elke storagemodule gebruikgemaakt van gegevensdragers die van drie verschillende productiebatches afkomstig zijn. Dit voorkomt gecorreleerde uitval. Replicatie naar een tweede locatie zet de beveiliging kracht bij met georedundantie.

Silent Cubes en Silent Bricks zijn ontworpen als modulaire opslagsystemen waarvan de capaciteit tot ettelijke petabytes kan worden

opgeschaald. Afzonderlijke opslagmodules die niet worden gebruikt worden uitgeschakeld. De Silent Bricks kunnen zelfs fysiek uit het systeem worden verwijderd voor veilige opslag op een externe locatie (air gapping). Dit draagt bij aan een forse energiebesparing, omdat zo alleen de actieve opslagmodules energie verbruiken. De modules wekken zichzelf regelmatig om zichzelf te controleren en waar nodig zelfstandig herstelbewerkingen uit te voeren.

Ook qua CO₂-voetafdruk scoren deze systemen goed. Met de introductie van de Silent Cube DS is FAST LTA begonnen om de CO₂-uitstoot die gepaard gaat met het produceren, leveren en afdanken van deze systemen volledig te compenseren. Voor toekomstige versies is het gebruik van gerecyclede materialen gepland. Dat geldt bijvoorbeeld voor de behuizing. FAST LTA moedigt zijn klanten actief aan om gebruik te maken van klimaatneutrale, hernieuwbare energie.

De opslagsystemen van FAST LTA zijn ontwikkeld met het oog op een gebruiksduur van minimaal tien jaar. Het gebruik van standaardschijven (SATA) draagt aan optimale compatibiliteit gedurende meerdere generaties bij. Als een schijf uitvalt kan de gegevensdrager in kwestie in een handomdraai worden vervangen. Er is in dat geval geen specifiek model nodig. De in eigen beheer ontwikkelde controller past de beschikbare capaciteit automatisch aan de aanwezige configuratie aan. Daarmee blijft de beschikbaarheid van vervangende onderdelen nog jarenlang gewaarborgd.

Ook de onderhoudscontracten zijn op deze lange levensduur afgestemd. Ze kunnen bovendien onder gelijkblijvende voorwaarden worden verlengd. Het goed geoliede servicenetwerk van FAST LTA staat 24/7/365 klaar om gebruikers van ondersteuning te voorzien.



Samenvatting

Het nut van een opslagsysteem varieert al naargelang de beoogde toepassing. Steeds meer organisaties vervangen volledig 'koude' archieven door actieve opslagsystemen met een hoog beschikbaarheidsniveau. Dit vraagt om een nieuwe duurzaamheidsaanpak. Schijfgebaseerde opslagsystemen moeten in dit kader aan diverse eisen voldoen. En alleen modulaire opslagsystemen die speciaal voor archiveringsdoeleinden zijn ontwikkeld zijn daartoe in staat.

Slim energiebeheer, een gebalanceerde CO2-voerafdruk, hardware en software met een lange levensduur en de langdurige beschikbaarheid van ondersteuning tegen planbare kosten zorgen ervoor dat zelfs 'online' opslagsystemen als duurzaam kunnen worden aangemerkt.

FAST LTA biedt met Silent Cubes en Silent Bricks twee oplossingen aan die perfect aansluiten op de eisen die bij langdurige gegevensopslag komen kijken. Ze bieden bescherming tegen onbevoegde wijzigingen van data, voldoen aan de eisen van de AVG en kunnen worden ingezet als omvangrijk virtual tape library (VTL)-archief.

