

FAST LTA
Wir sichern Terabytes.

COLD STORAGE

Tape zu langsam, Disk Libraries zu teuer?

NEUE TECHNOLOGIEN FÜR BACKUP UND ARCHIVIERUNG



Hallo,

wir sind die **FAST LTA AG**.

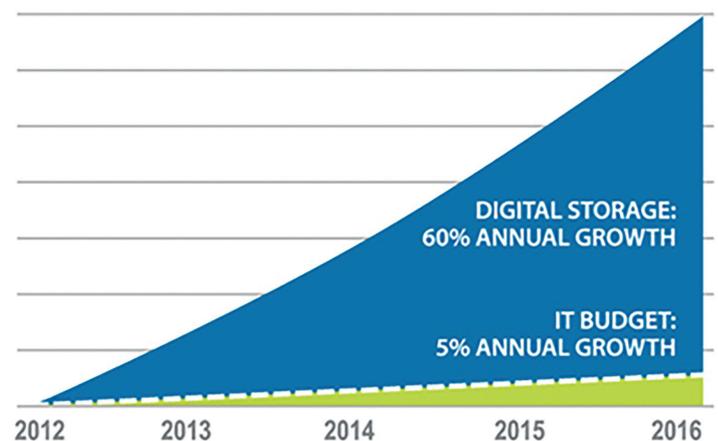
Unser Firmensitz ist in der Rüdeshheimer Str. 11 in 80686 München, Deutschland. Sie erreichen uns telefonisch unter +49 (89) 890 47 - 0, per Fax unter +49 (89) 890 47 - 890 und via E-Mail über info@fast-lta.de. Beim Amtsgericht München sind wir unter der HRB 127 484 eingetragen, unsere USt-ID ist DE204232266. Die Kontodaten bei der Kreissparkasse München Starnberg sind IBAN DE76 7025 0150 0022 2363 19 und BIC BYLA DE M1 KMS. Im Vorstand sitzen Matthias Zahn, Jörg Adelstein und Reiner Bielmeier, der Vorsitzende des AR ist Jens Gloede. Im Internet finden Sie uns unter www.fast-lta.de.

DISCLAIMER

Das Dokument erhebt keinen Anspruch auf Vollständigkeit. Änderungen jederzeit vorbehalten. Alle Bilder und Marken sind Eigentum ihrer jeweiligen Rechteinhaber. Veröffentlichung - auch auszugsweise - nur nach ausdrücklicher Genehmigung von FAST LTA.

Tape ist tot. Oder doch nicht?

Das weltweite Datenvolumen wächst bis 2020 um das Zehnfache, schätzen Experten¹. Von derzeit 4,4 Billionen Gigabyte auf 44 Billionen Gigabyte. Laut dem Branchenverband BITKOM wuchs die Datenmenge in deutschen Unternehmen allein im Jahr 2013 im Schnitt um 22 Prozent². Eine Erweiterung der Speicherkapazitäten plant daher die Hälfte aller befragten IT-Entscheider. Gleichzeitig stagnieren jedoch die IT-Budgets oder sinken sogar. Laut einer Analyse von Capgemini haben rund 40 Prozent der IT-Abteilungen dieselben finanziellen Mittel zur Verfügung, wie im vergangenen Jahr, fast 20 Prozent steht deutlich weniger Budget zur Verfügung³. Langfristig stehen IT-Verantwortliche also vor einem Dilemma: Bei geringerem Budget müssen sie immer größere Datenmengen sicher speichern.



Datenwachstum gegen
IT-Budget.
Quelle: sagecloud

Dabei ist eine Abwägung zwischen Schnelligkeit und Kosten erforderlich. Prinzipiell gibt es dazu im Moment vier Möglichkeiten:

1. In-Memory (direkt im Hauptspeicher)
2. Flash Drive oder Solid State Disk (SSD)
3. Disk Storage (Festplatten)
4. Tape (Bandlaufwerke).

Die ersten beiden Kategorien kommen in der Regel bei aktuellen Daten zum Einsatz, auf die häufig und schnell zugegriffen werden soll, etwa für Datenbanken oder virtuelle Maschi-

¹ „Digitales Universum explodiert durch Sensordaten“, Studie von EMC, April 2014 (<http://bit.ly/1CBkcfz>)

² „Datability“, Vortrag von Prof. Dieter Kempf, 9. März 2014, S. 6. (<http://bit.ly/1Cod8SZ>)

³ „IT-Trends Studie: Budget-Prognosen für 2015 sind vorsichtig optimistisch“, Capgemini, Dezember 2014 (<http://bit.ly/1ABbqL9>)

nen. Die relativ hohen Kosten sind dabei gerechtfertigt. SSDs werden sich für I/O-intensive Speicheranforderungen mittelfristig durchsetzen, da sie einen deutlichen Geschwindigkeitszuwachs, niedrigeren Energieverbrauch und keine Probleme bei Fragmentierung vorweisen.

Ist allerdings eine längerfristige Speicherung von Daten gefordert, auf die nur selten zugegriffen werden soll – wie bei Backup-Daten, Archiven oder Videodaten –, gewinnen die Kosten eine höhere Bedeutung. Die Alternativen lauten dann Disk oder Tape. Dabei bieten Festplatten eine deutlich höhere Zugriffsgeschwindigkeit, während das langsamere Bandlaufwerk deutlich billiger ist.

Sowohl Disk Storage als auch Tape haben jedoch Nachteile. Bei traditioneller Speicherung auf Festplatten sind das in erster Linie der hohe Energieverbrauch und die Fragmentierung der Daten, bei Tape die langen Zugriffszeiten.

Disk Storage

Disk Storage organisiert die Daten in kleinen Einheiten (Blocks). Die logische Zuordnung von Blöcken übernimmt der sogenannte Disk Drive Controller für alle Blocks, Sektoren und Tracks auf der Platte⁴. Dies führt dazu, dass Daten auf unterschiedlichen Sektoren der Platte abgelegt werden. Bei logischen Blöcken ist es für ein Betriebssystem oder ein Subsystem dadurch unmöglich festzustellen, auf welchem Sektor der Festplatte sich eine Datei befindet. Die Folge ist langfristig eine starke Fragmentierung der Daten. Da es produktionsbedingt auf jeder Hard Disk sogenannte Bad Sectors gibt, bedeutet das eine Gefahrenquelle für die Datenintegrität und ein hohes Fehlerrisiko.

Um diesem Risiko vorzubeugen und zur Steigerung von Kapazität und Geschwindigkeit wurde das Konzept des Block Storage bereits 1987 zum „Redundant Array of Independent Disks“ (RAID) ausgebaut⁵. RAID fasst eine bestimmte Anzahl von Festplatten zu einem

4 „Data Structures on Disk Drives“, Storage Networking Industry Association (<http://bit.ly/1JI8ozP>)

5 „A Case for Redundant Arrays of Inexpensive Disks (RAID)“, David A. Patterson, Garth

Verbund zusammen und steigert so Geschwindigkeit und Kapazität. Durch das zusätzliche Speichern so genannter Parity-Informationen können einzelne Festplatten bei Ausfall ersetzt und wieder hergestellt werden, ohne dass Daten verloren gehen. Je nach RAID-Level können eine

(RAID-5) oder zwei (RAID-6) Festplatten ausfallen und wieder hergestellt werden. Damit ist eine Absicherung durch Redundanz ohne vollständige Kopie möglich.

RAID löst jedoch nicht das Problem der Fragmentierung. Im Gegenteil: Statt auf nur einer Festplatte sind die Daten nun über die Disks im gesamten Verbund verteilt.



Auf herkömmlichen Festplattenspeichern liegen die Daten fragmentiert und über den gesamten RAID-Verbund verteilt vor.

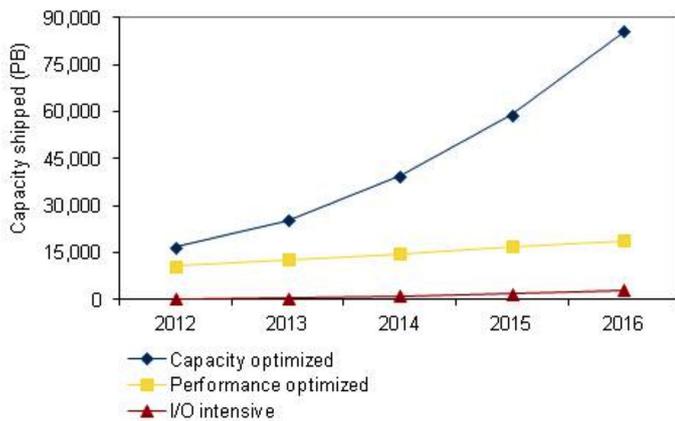
Tape Storage

Die strukturelle Ablage von Daten ist der elementare Unterschied zwischen Disk und Tape Storage. Auf Magnetband sind sämtliche Informationen strikt hintereinander gespeichert. Da ein Datensatz so zusammenhängend gespeichert wird, eignen sich Tapes als Transportmedien. Allerdings ist der Zugriff auf einzelne Daten langwierig, da nicht nur das entsprechende Tape gefunden und eingelegt, sondern auch zu der entsprechenden Stelle gespult werden muss. Für moderne „Search & Browse“-Anwendungen ist dies nicht praktikabel.

Tapes gelten heute zwar generell als zuverlässig, jedoch ist Datenverlust durch Alterung und mechanische Beanspruchung nicht ausgeschlossen. Tapes sollten deshalb regelmäßig gelesen und geprüft, im Idealfall auch umkopiert werden, um die Datenintegrität sicherzustellen. Absicherung durch Redundanz ist nur über vollständige Kopien möglich.

Linear Disk Storage – Die Zukunft der Disk Library

Worldwide Capacity-Optimized, Performance-Optimized, and I/O-Intensive Storage Systems Shipments, 2012–2016



Quelle: IDC

Da geschwindigkeitsrelevante und I/O-intensive Speicheranwendungen zukünftig vermehrt auf SSDs setzen werden, bleibt für Disk Storage der Bereich des „capacity optimized storage“, dem die IDC überproportionales Wachstum vorhersagt⁶.

Linear Disk Storage versteht sich dabei als Lösung zwischen den beiden genannten Optionen Disk und Tape, die zwar alle Vorteile vereint, die Nachteile jedoch außen vor lässt. So werden hier die Daten nicht fragmentiert wie auf klassischen Disks abgelegt, sondern linear. Gleichzeitig

kann aber wie bei Disk Storage üblich mit hoher Geschwindigkeit auf beliebige Daten zugegriffen werden (Random Access). So wird die Brücke zwischen den Welten Disk und Tape geschlagen.

Lineares Schreiben, Random Access



Der Schreibvorgang erfolgt ähnlich wie bei einer linearen Tape Library. Die Daten werden also in genau der Reihenfolge abgelegt, in der sie eingehen. So gibt es einen

Auf Silent Bricks werden die Daten linear gespeichert und sind somit genau lokalisierbar.

inhaltlichen und zeitlichen Zusammenhalt der Daten: Das System und damit auch der User wissen genau, wo sich die Daten befinden – und wo nicht. Daher sind zum einen deutlich weniger Bewegungen der Festplattenköpfe nötig, um die Dateien zu finden, zum anderen

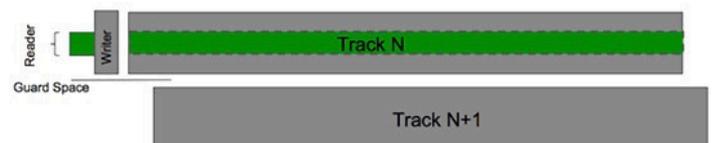
⁶ „Cold Storage is Hot Again“, Studie von IDC (<http://www.storiant.com/resources/Cold-Storage-Is-Hot-Again.pdf>)

lassen sich Daten leichter lokalisieren und bei Bedarf vernichten. Auch energetisch hat diese Technologie Vorteile. Einzelne Festplatten, die aktuell in der Library nicht benötigt werden, können einfach abgeschaltet werden. Das senkt den Energieverbrauch entscheidend und erhöht gleichzeitig die Lebensdauer der Hard Disks. Zugriffe erfolgen immer nur auf exakt den Platten, die benötigt werden. So entsteht eine granulare Energieverwaltung, die sehr effizient ist. Da keine Defragmentierung erforderlich ist, reduziert sich die mechanische Belastung zudem erheblich, was Wartungsaufwand und Administration vermindert.

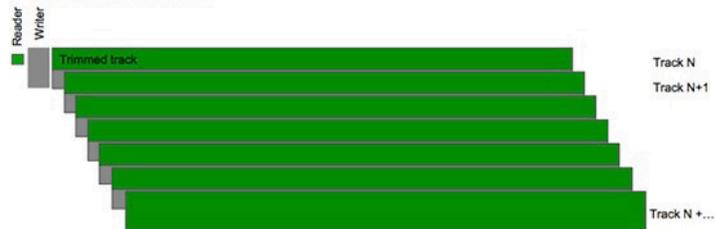
Neue Festplattentechnologien

Unterstützt wird diese effiziente Festplattennutzung durch neueste Festplatten-Technologien. Da die magnetische Auflösung an ihre Grenzen stößt, wird nach neuen Wegen gesucht, noch mehr Kapazität pro Disk zu speichern. Das „shingled magnetic recording“ (SMR)⁷ nutzt dabei aus, dass beim Lesen schmalere Tracks benötigt werden als beim Schreiben und überlappt die Datentracks deshalb. Dies ist ideal für lineare Speicherung, da einmal überlappe Tracks nicht einfach überschrieben werden können – der benachbarte Track würde dabei mit überschrieben. In herkömmlichen, fragmentierten Umgebungen resultiert dies in erheblich reduzierter Schreibgeschwindigkeit. Bei rein linearer Anwendung ist dies jedoch kein Nachteil.

Conventional Writes



SMR Writes



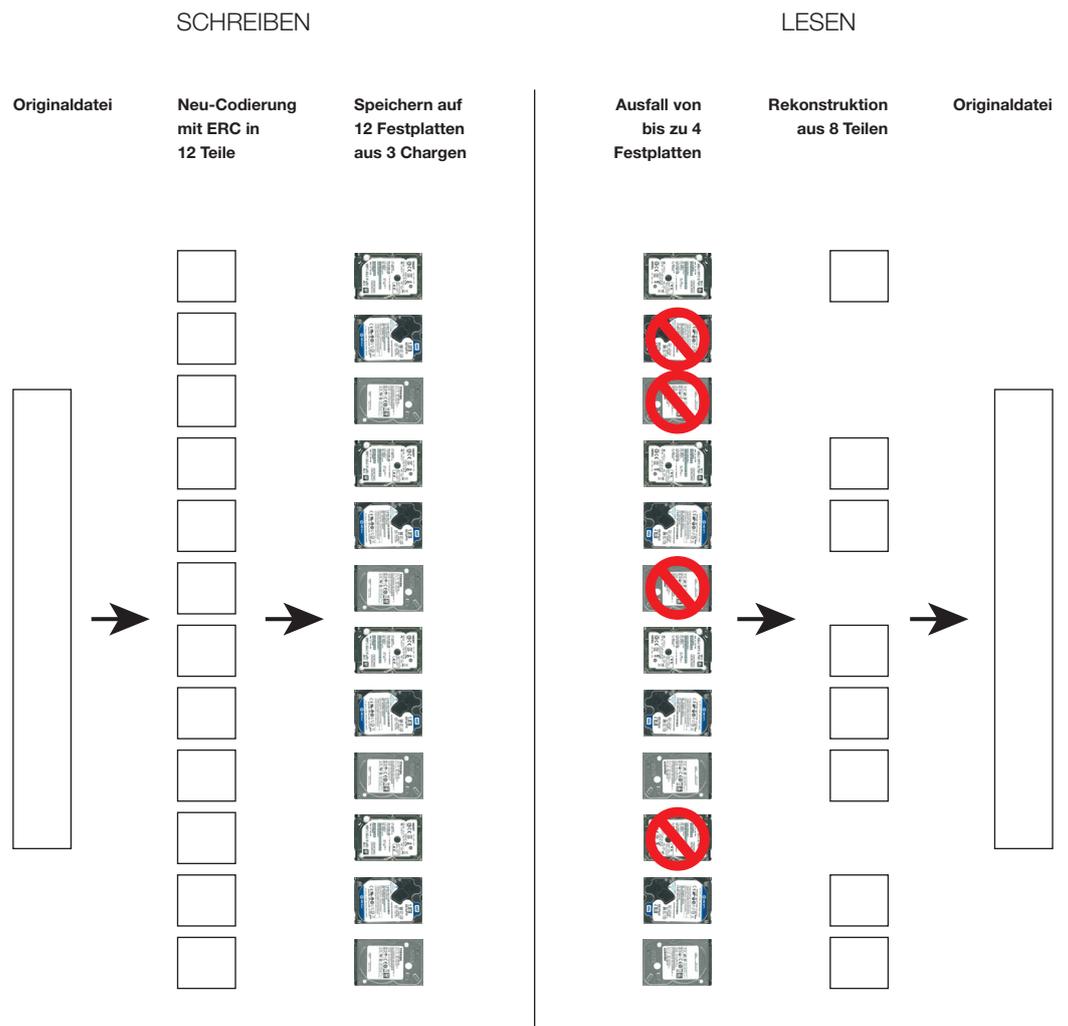
Shingled Magnetic Recording (SMR)
Quelle: Seagate

⁷ „Shingled Magnetic Recording and Two-Dimensional Magnetic Recording“, Roger Wood, 19.10.2010 (<http://bit.ly/1z3f9Vu>)

Cost Optimized Linear Disk (COLD) Storage

COLD Storage bezeichnet die von FAST LTA eingesetzte Technologie der linearen Datenspeicherung auf Standard-Festplatten bei gleichzeitiger Optimierung der Speicher-Grenzkosten, des Wartungsaufwands und der Energiekosten: Cost Optimized Linear Disk Storage. Dabei sind Daten so sicher gespeichert, dass kein weiteres Backup notwendig ist.

Erasure Coding: Der RAID-Nachfolger



Beim Erasure Coding im ERC 8/12-Schema können bis zu 4 beliebige Festplatten ausfallen, ohne dass Daten verloren gehen

Durch den Einsatz des sogenannten Erasure Coding wird eine sehr hohe Datensicherheit erzielt. Erasure Coding teilt Daten in Fragmente auf, erweitert sie mit redundanten Daten-Stücken und kodiert sie erneut. Danach werden diese über verschiedene Medien hinweg verteilt. FAST LTA setzt Arrays von 12 Festplatten ein, von denen beliebige 8 ausreichen, um die gespeicherten Daten wiederherzustellen (ERC 8/12). Selbst beim gleichzeitigen Ausfall von 4 Festplatten pro Silent Brick sind Funktion und Datenverfügbarkeit uneingeschränkt gewährleistet. Damit ist diese Konfiguration sicherer als eine gespiegelte RAID-6 Konfiguration.

Den gegenüber RAID gestiegenen Rechenaufwand lagert FAST LTA dabei auf eine leistungsfähige GPU (Grafikprozessor) aus, der für solche mathematischen Spezialoperationen herkömmlichen CPUs weit überlegen ist.

Mischbestückung: Schutz vor Chargenfehler

Dennoch gibt es Fälle, in denen auf Festplatten gespeicherte Daten ohne äußere Einflüsse verloren gehen können. Das eine Problem sind so genannte Chargenfehler (englisch: Epidemic Failure), bei denen innerhalb statistisch ungewöhnlich kurzer Zeit gleiche Festplattenmodelle Defekte aufweisen. Dies kann z.B. durch fehlerhafte Chargen in der Produktion entstehen (bedingt durch veränderte Produktionsbedingungen aber auch Fehler in Teilen der Firmware), oder einfach durch den Alterungsprozess von mechanischen Teilen auftreten. In Festplattensystemen mit Festplatten, die alle aus der gleichen Charge kommen – was durchaus üblich und z.T. sogar durch den Festplatten-Controller notwendig ist – bedeutet ein Festplattenausfall deswegen oft Gefahr für das Gesamtsystem.

FAST LTA setzt deswegen in allen seinen Festplattenspeichern immer Festplatten aus mindestens 3 unterschiedlichen Chargen von möglichst vielen unterschiedlichen Herstellern ein. Beim oben beschriebenen ERC 8/12 können so alle 4 Festplatten einer Charge ausfallen, ohne dass Datenverlust droht.

Digital Audit: Regelmäßige Selbstüberprüfung

Und noch ein weiteres Problem befällt die Festplatten: Der so genannte Unrecoverable Read Error, kurz: URE. Festplattenhersteller geben an, dass etwa eines von 10^{14} Bits einfach und ohne besonderen Grund nicht richtig gelesen werden kann. Das klingt nach einem zu vernachlässigbarem Problem, schließlich ist das eine enorme Zahl mit 14 Nullen. Allerdings sind die Datenmengen heutzutage dermaßen hoch, dass dieses Problem sehr real wird: Alle 12,5 Terabytes gibt es statistisch einen Lesefehler. Bei codierten Daten bedeutet das, dass die Datei schlicht nicht wiederhergestellt werden kann. Mit den heutigen Festplattengrößen (Stand 2015: bis zu 6 TB pro Festplatte) ist Datenverlust also fast schon sicher. Die Tücke dabei ist, dass der Fehler erst beim Lesen der Daten erkannt werden kann. Eine Speicherung, die auf Parity-Vergleich beruht (RAID), erkennt diesen Fehler also nicht – bis genau die eine betroffene Datei wiederhergestellt werden soll.

Die eigens entwickelten Festplatten-Controller von FAST LTA weisen deshalb eine wichtige Funktion auf: Den Digital Audit. Im Controller verankert und vom Host-System unabhängig liest diese Funktion eigenständig und regelmäßig jedes einzelne Bit der gespeicherten Daten und vergleicht die Hashes, die sich aus den Inhalten der Dateien ergeben und somit von jedem Bit abhängen, mit den bei der ursprünglichen Speicherung gesicherten Hashes. Tritt nun so ein Bitfehler auf, wird er von Digital Audit zuverlässig erkannt. Durch die hohen Redundanzreserven kann eine solche kompromittierte Festplatte automatisch im Hintergrund repariert oder abgeschaltet werden, ohne dass Datenverlust droht.

Die Silent Brick Library: So sicher, dass kein weiteres Backup notwendig ist

Fehlt einem Festplattenspeicher eines der beschriebenen Sicherheitsmerkmale, müssen die Daten zusätzlich, z.B. durch ein regelmäßiges Backup, abgesichert werden. Daten, die auf Systemen von FAST LTA gespeichert sind, benötigen dagegen kein zusätzliches Backup, vorausgesetzt, der Speicherstandort ist physikalisch gegen Totalverlust (Brand, Wasser, Diebstahl) gesichert.

Die Silent Brick Library von FAST LTA ist die logische Weiterentwicklung des Konzepts von Linear Disk Storage. Die Kombination aus Erasure Coding und Linear Disk Storage sorgt für eine unvergleichliche Flexibilität durch exakte Lokalisierbarkeit von Daten bei niedrigen Speicher- und Betriebskosten und enormer Sicherheit.

Erfahren Sie mehr zu COLD Storage und zur Silent Brick Library von FAST LTA. Besuchen Sie uns auf der CeBIT, Halle 2, Stand A37 oder im Internet unter www.silentbricks.de.



Silent Brick Library:
Disk Library für Cold
Storage Anwendungen

FAST LTA
Wir sichern Terabytes.

COLD STORAGE

Hallo,

wir sind die **FAST LTA AG**.

Unser Firmensitz ist in der Rüdeshheimer Str. 11 in 80686 München, Deutschland. Sie erreichen uns telefonisch unter +49 (89) 890 47 - 0, per Fax unter +49 (89) 890 47 - 890 und via E-Mail über info@fast-lta.de. Beim Amtsgericht München sind wir unter der HRB 127 484 eingetragen, unsere USt-ID ist DE204232266. Die Kontodaten bei der Kreissparkasse München Starnberg sind IBAN DE76 7025 0150 0022 2363 19 und BIC BYLA DE M1 KMS. Im Vorstand sitzen Matthias Zahn, Jörg Adelstein und Reiner Bielmeier, der Vorsitzende des AR ist Jens Gloede. Im Internet finden Sie uns unter www.fast-lta.de.