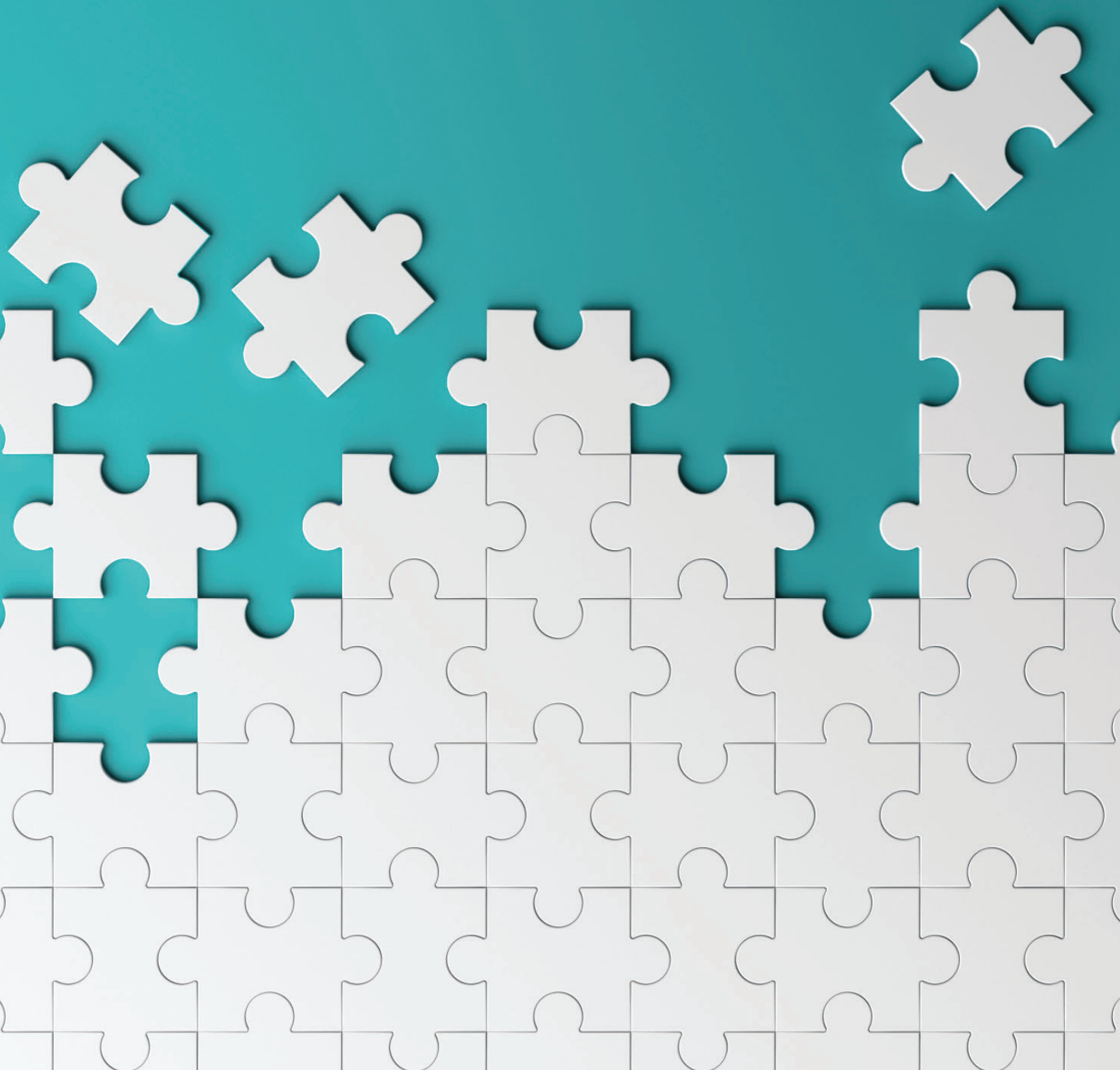


WHITE PAPER

# Speicherkonsolidierung für erhöhte Verfügbarkeit und niedrigere Kosten



# Inhaltsverzeichnis

Zusammenfassung .....	3
Einführung .....	4
Konsolidierungsarbeiten .....	5
Mögliche Vorteile der Speicherkonsolidierung .....	6
Downtime Causes .....	6
Nachbesserungshäufigkeit .....	7
Leistung bei Hardwareausfällen .....	8
Einsatz von RTOs und RPOs .....	8
Speicher-Array Verfügbarkeitsgrundlagen .....	9
Reales vs. Erkanntes Risiko .....	9
Verfügbarkeit .....	9
Richtlinien für das Design von Systemen mit hoher Verfügbarkeit .....	9
Durchschnittliche Dauer zwischen Datenverlusten .....	11
Anschaffungs- und Betriebskosten .....	11
Best Practices für die Speicherkonsolidierung .....	12
Fazit .....	12

# Zusammenfassung

Dieses Whitepaper zeigt, dass in den meisten Umgebungen die Bedenken von CIOs, Speicherarchitekten und IT-Direktoren hinsichtlich der Frage, wie Speicherkonsolidierung den Umfang von Fehlerdomänen erhöht, unbegründet sind, da moderne, auf marktvalidierte Unternehmensarrays mit mehreren Controllern eine wirksame 100%ige Verfügbarkeit bieten können<sup>1</sup>. Sie verfügen über eine höhere Nutzbarkeit, eine größere Fehlertoleranz und zuverlässigere, unterbrechungsfreie Softwareaufrüstung als Arrays mit zwei Controllern<sup>2</sup>. Die Speicherkonsolidierung vereinfacht die Infrastruktur-Topologien, indem die Anzahl der zu verwaltenden Arrays und die Möglichkeiten für Fehlkonfigurationen reduziert werden, was zum Budget für Ausfallzeiten beiträgt. Die Speicherkonsolidierung erzeugt einen größeren Fundus an freiem Speicherplatz, wodurch die nutzbare Verfügbarkeit und die operative Leistungsfähigkeit verglichen mit einem Verbund kleinerer Speichersysteme mit zwei Controllern verbessert wird. Durch die verringerte Anzahl von Controllern und Ports in der SAN-Infrastruktur wird ferner die Nachbesserungshäufigkeit sowie der Strom- und Kühlungsbedarf minimiert.

<sup>1</sup> „Marktvalidierte“ Arrays sind Arrays, die im Allgemeinen mindestens 9 bis 12 Monate verfügbar sind und 10s von Tausenden von Maschinenjahren Produktionserfahrung haben.

<sup>2</sup> „Nutzbare Verfügbarkeit“ bezeichnet die Fähigkeit, die Service-Level-Ziele bei Hardware-Ausfällen zu erreichen.

# Einführung

Bestehende Speicherinfrastrukturen haben häufig mit dem Manko von vermeidbaren Ineffizienzen zu kämpfen, wie Komplexität der Verwaltung, mangelnde Beweglichkeit, Verfügbarkeits- und Leistungsprobleme, Kapazitätsengpässe, Qualifikationsdefizite, Backup-/Wiederherstellungsprobleme und Geldverschwendung.

Speicherkonsolidierungsprojekte lösen viele dieser Probleme, indem sie Infrastrukturarchitekten die Möglichkeit bieten, frühere Entscheidungen neu zu bewerten und eine auf aktuelle und zukünftige Arbeitslasten zugeschnittene Speicherinfrastruktur zu entwerfen. Mit der Schaffung selbstverwaltender Speicher-Arrays, die intelligente Entscheidungen über die Datenplatzierung und den Zugriff auf eine Vielzahl von Migrationstools treffen, sind technische Risiken und Bedenken im Zusammenhang mit der Speicherkonsolidierung hinsichtlich der Datenverfügbarkeit, der Größe von Fehlerdomänen und Qualifikationsdefizite für eine Speicherkonsolidierung unbegründet.

Dieses Whitepaper zeigt, dass Konsolidierung eine erfolgreiche Strategie zur Verbesserung der Verfügbarkeit und der operativen Effizienz ist. Außerdem werden die Hauptursachen für Ausfallzeiten und Nachbesserungen untersucht; es bietet eine hochkarätige Anleitung zur Fehlermathematik, beschreibt die Vorteile der Speicherkonsolidierung und legt die besten Verfahren zur Durchführung eines Speicherkonsolidierungsprojekts fest.

## Konsolidierungsarbeiten

Skeptikern der „Volles Risiko“-Strategie hat die Vergangenheit gezeigt, dass die Risikostabilisierung ein funktionierendes Konzept ist, da nur ein Gefahrenmomentum zu überwachen ist. Drei gute Beispiele, die den Erfolg dieser Strategie unterstreichen, sind:

- ▶ Luftverkehr – Von Anfang an haben sich die Flugzeughersteller darauf konzentriert, größere und sicherere Flugzeuge zu bauen, die weiter, schneller und höher fliegen können. Das Ergebnis ist, dass Flugreisen pro Sitzplatz/Meile sicherer sind als Autofahren und Flugzeugabstürze eine Seltenheit sind. Warum? Größere Flugzeuge haben Größenvorteile, wodurch die Ergänzung von Systemredundanz und verbesserten Sicherheitsfunktionen verglichen mit kleineren Flugzeugen erschwinglicher wird, und die Fehlleitung von Flugzeugen durch die Flugsicherung verringert wird, da es weniger Flugverkehr gibt.

- ▶ Rohöltransporte – Das Wachstum der weltweiten Rohöltransporte führte zur gleichen „Größer ist besser“-Strategie, die auch von den Flugzeugherstellern angewandt wurde: weniger größere, dafür sicherere, mit GPS-Navigation und viel Automatisierung ausgestattete Tanker mit verstärktem Schiffsschutz anstelle vieler kleinerer Tanker. Da die Kosten für den Bau eines Schiffes eng an ihr Gewicht gebunden sind (mehr Low-Tech- als High-Tech-Schiffe) und das Volumen schneller zunimmt als die Fläche (d. h. das Gewicht), produziert der Bau großer Tanker mit verstärktem Schiffsschutz geringere Kosten, als der Bau kleinerer Tanker mit verstärktem Schiffsschutz. Das Ergebnis: schnellere, zuverlässigere und unfallfreie Lieferung von Öl weltweit.

- ▶ Speicherkonsolidierung – Marktstatistiken von IDC und Gartner zeigen, dass die Betreiber beim Versuch kostenwirksam die Verfügbarkeit der Speicherinfrastruktur und die operative Effizienz zu verbessern eine ähnliche Konsolidierungsstrategie verfolgen. Kurz gesagt zeigen die Statistiken einen Anstieg der jährlichen PB-Sendungen, auch wenn der jährliche Versand von Speicher-Arrays rückläufig ist. Mit anderen Worten: Die durchschnittlichen Kapazitätskonfigurationen von Speicher-Arrays nehmen zu.



## Mögliche Vorteile der Speicherkonsolidierung

Die Vorteile, die beim Supportaufbau für die Speicherkonsolidierung am effektivsten sind, sind jene, die den größten Einfluss auf betriebliche Achillesfersen haben. Das sind die bevorzugten Schwachstellen: mangelnde Beweglichkeit, Verfügbarkeitsprobleme, Leistungsprobleme, Kapazitätsengpässe, Qualifikationsdefizite, Backup-/Wiederherstellungsprobleme und Budgetbeschränkungen.

Nachfolgend finden Sie beispielhaft Vorteile für den Betreiber, die diese Schwachstellen thematisieren:

- ▶ Die Speicherkonsolidierung macht die Erweiterung der Klasse von Dual-Controller-Mittelklasse- zu Multi-Controller-High-End-Arrays finanziell erschwinglich.
- ▶ Moderne Multi-Controller-Arrays sind skalierbar und bieten eine konsistentere Leistung bei Hardware-Ausfällen und/oder Software-Aktualisierungen.
- ▶ Die Preisreduktion für Kosten \$/TB/Monat ermöglicht den Kauf von mehr Speicher.
- ▶ Bei einem selbstverwaltenden Speicher, der das Management vereinfacht, bleibt die Mitarbeiterzahl auch bei steigender Kapazität unverändert; dies gilt auch für Verbesserungen bei Leistung und D/R.
- ▶ Moderne Multi-Controller-Hybrid-Arrays, die RAID 6 oder die Löschcodierung implementieren, erhöhen, verglichen mit RAID 1- und RAID 5-Konfigurationen, die durchschnittliche Betriebsdauer zwischen Datenverlusten (MTBDLs) um das Zehnfache.
- ▶ Der Einsatz speichereffizienter Funktionen, wie Datenkomprimierung und Deduplizierung, senkt die Teileanzahl, die Nachbesserungsfrequenz und die Kosten \$/PB.

## Ursachen für Ausfallzeiten

Tabelle 1 zeigt in absteigender Reihenfolge der Häufigkeit die Hauptursachen für Ausfallzeiten von Speichersystemen. Während man über die konkrete Reihenfolge dieser Ursachen diskutieren könnte, gehen wir aber konform, dass Speicher- und SAN-bezogene Hardwareausfälle keine primären Quellen für Ausfallzeiten sind. Die Erfahrung und die Analyse der Grundursache haben gezeigt, dass, wenn Speicher-Hardwareausfälle Ausfallzeiten verursachen, es fast immer ein Softwarefehler ist, was durch den ausfallverursachenden Hardwarefehler aufgezeigt wurde. Wäre dies nicht der Fall, hätte die Cloud-basierte Überwachung und Analyse keine Auswirkungen auf die Verfügbarkeit des Speichersystems.

**TABELLE 1**

Ursache	Bemerkung
Menschliche Fehler	Proportional zur Komplexität und Häufigkeit der Interaktionen zwischen Speicher-Admins und FE's mit den Speicher-Arrays
Softwarefehler	Umgekehrt proportional zur Code-Reife und proportional zur Softwaregröße und Komplexität
Mangelhafte Steuerung der Softwareänderung	Verbessert durch Cloud-basierte Analyse
Infrastruktur-Fehlkonfigurationen	Proportional zur Infrastrukturkomplexität und zur End-to-End-Konfigurationsvalidierung, die häufig fehlt
Fehlerhafte D/R-Tests	Failover/Failback funktionieren nicht

## Nachbesserungshäufigkeit

Die Nachbesserungshäufigkeit verhält sich proportional zur Anzahl der Komponenten in einem Speichersystem und umgekehrt proportional zu den MTBFs der Komponenten. Mehr Teile bedeuten mehr Ausfälle, was mehr Nachbesserungen erfordert. Gleichung 1 legt die MTBF als Kehrwert der Ausfallrate fest, was bedeutet, dass sich eine höhere MTBF in einer geringeren Häufigkeit von Reparaturaktivitäten niederschlägt.

### Gleichung 1 – MTBF

$$\text{MTBF} = 1/\text{Ausfallquote}$$

Da großflächige PB-Rechenzentren über Hunderte bis Tausende rund um die Uhr in Betrieb stehende HDDs und SSDs verfügen, entfallen auf sie fast alle Hardwarenachbesserungen. Abbildung 1 zeigt, dass, geht man von der Theorie in die Praxis über, die Nearline-Festplatten von Seagate mit einer MTBF von 1,2 Mio. auf Jahresbasis eine Ausfallrate von 0,73 % aufweisen. Abbildung 1 zeigt, dass ein aus 480 dieser Seagate Nearline-Festplatten bestehendes Array, d. h. etwa 4 PB doppelt paritätsgesicherter Kapazität mit 12 TB-Festplatten, nicht mehr als 3,5 Festplattenreparaturen pro Jahr oder nicht mehr als 1 Festplattenreparaturen/PB/Quartal aufweisen sollte.

#### ABBILDUNG 1

MTBF (Stunden).....	1.200.000
Stunden/Jahr (24 x 365).....	8.760
Jährliche Ausfallrate (AFR) (Stunden pro Jahr / MTBF) .....	0,73%
# von Arrays HDDs.....	480
Jährliche Häufigkeit von Festplattenreparaturen (AFR x # Array-HDDs).....	3.504

Diese relativ häufigen Reparaturen im Verbund mit Plattenkapazitäten, die schneller als die Datenübertragungsraten wachsen, haben zum Einsatz von robusteren Datensicherungsschemata und niedrigeren Wiederherstellungszeiten zu kritischen Designzielen in modernen Arrays geführt, da schnelle Wiederherstellungszeiten „das Zeitfenster die Verwundbarkeit“ verringern, das sich immer öffnet, wenn eine HDD oder SSD ausfällt<sup>3</sup>. Die Verringerung der Verwundbarkeit erhöht die Datenhaltbarkeit bzw. die durchschnittliche Dauer zwischen den Datenverlusten.

Zwei Methoden, die sich bei verringerten Wiederherstellungszeiten bewährt haben, sind:

- ▶ Ersatz des Reserveplattenkonzepts durch freie Kapazität, die den Datenwiederaufbau durch Parallelisierung des Wiederherstellungsprozesses beschleunigt.
- ▶ Nur die Datenwiederherstellung anstelle des gesamten Geräts verkürzt die Datenwiederherstellungszeiten weiter.

Technologien zur Datenreduzierung (Komprimierung und Deduplizierung) verringern zwar nicht die Wiederherstellungszeiten, aber sie verkleinern die Anzahl der Geräte, die für eine bestimmte Datenmenge benötigt werden und verringern so die Ausfallraten: intelligente Wiederherstellungen. High-End-Multi-Controller-Architekturen verfügen über die Rechenleistung und Bandbreite, die erforderlich sind, um performante Implementierungen dieser Techniken zu erstellen, selbst wenn Hardwarefehler vorliegen. Die Mathematik ergibt eine erwartete jährliche Häufigkeit von Reparaturen für nicht medienbezogene Hardware-Ausfälle von 0,08 pro Jahr.

<sup>3</sup> Allgemein bekannte Beispiele für widerstandsfähigere Datenschutzsysteme: doppelte Parität, Löschkodierung und Reed Solomon

Selbst wenn wir die AFR auf 10 % erhöhen würden oder bei einer MTBF mit nur 87.600 Stunden, wäre die erwartete AFRA immer noch weniger als einmal pro Jahr. Mit anderen Worten: Ausfälle von Array-Controllern sind bei einer Entscheidung für eine Speicherkonsolidierung im Wesentlichen „irrelevant“.

## Leistung bei Hardwareausfällen

Da Controller-Ausfälle, ausbleibende Softwarefehler oder misshandelte Reparaturen unkritische Ereignisse sind, zeigen sich die sichtbaren Folgen eines Controller-Ausfalls in einem Dual- oder Multi-Controller-Array im Allgemeinen in Form von verringerter Leistung (IOPS), verringertem Durchsatz (GB/s) und/oder der Latenz (Millisekunden) und nicht als Verlust der Datenzugänglichkeit oder Datenintegrität. Die Ausfallmathematik und das Konzept der nutzbaren Verfügbarkeit begünstigen von Natur aus Multi-Controller-Arrays, da sich die Folgen von Controllerausfällen auf die Leistung umgekehrt proportional zur Anzahl von aktiven Controllern in einem Array auswirken. Der Ausfall eines Controllers in einem aktiven/aktiven Array mit zwei Controllern könnte die Leistung um bis zu 50 % oder mehr verringern, wenn er den überlebenden Controller zwingt, vom Schreib- in den Durchschreib-Cache-Modus zu wechseln; ein aktiv/aktives Array mit drei Controllern um bis zu 1/3; ein Array mit vier Controllern um bis zu 25 %, et al. Die Verwendung der Phrase „bis zu“ ist keine subtile „Sesam öffne dich“-Karte, sondern eine Bestätigung, dass die Aufteilung der erhöhten Arbeitsbelastung auf mehrere überlebende, nicht mit maximaler Kapazität arbeitende Controller die Auswirkungen von Controller-Ausfällen weiter minimiert.

Betreiber, die sich selbst einen Freiraum von 25 bis 30 % der überschüssigen Leistungsreserven überlassen, werden nur selten einen Ausfall bemerken, abgesehen von proaktiv gesendeten Warnmeldungen des Speicherarray-Anbieters. Der Freiraum ist wertvoll, da er den Arrays eine Hardwareausfall-Toleranz erlaubt. Er verbessert auch die Nutzverfügbarkeit, indem er Softwarefehler verbirgt, die nur dann gezeigt werden, wenn ein Array unter extremer Belastung steht. Durch die Beibehaltung eines Freiraums von 25 bis 30 % haben Betreiber, die keine COD- oder verbrauchsabhängigen Preismodelle verwenden, auch genügend Zeit, um kosteneffiziente Upgrades auszuhandeln.

Unternehmen mit Gewährleistungen, die nicht genügend Spielraum für unvorhergesehene Leistungsprobleme oder organisches Wachstum reservieren können, können aufgrund der Ineffizienz dieser Topologie von einer Strategie der Nichtkonsolidierung profitieren, allerdings auf Kosten einer geringeren betrieblichen und finanziellen Effizienz.

## Einsatz von RTOs und RPOs

Der Einsatz von RTOs und RPOs hängt sowohl von der Häufigkeit von Speicherauszügen als auch von der Menge der zu schützenden Daten ab. Eine Erhöhung der Häufigkeit von Speicherauszügen verkürzt die RPOs, indem die Zeit zwischen dem letzten Speicherauszug und dem Beginn einer Wiederherstellung verkürzt wird. Außerdem werden RTOs verkürzt, indem die Anzahl der Transaktionen, die verlängert werden müssen, oder die wiederherzustellende Datenmenge verringert wird. Diese Erkenntnisse führen unweigerlich zu dem Konzept der kontinuierlichen Datensicherung oder dass jede Datenänderung von einem Speicherauszug begleitet wird. Obwohl das Konzept verlockend ist, leidet es an zwei systemeigenen Problemen: Die Erstellung von Speicherauszügen führt bei jedem Schreibvorgang zu einem zusätzlichen Softwareaufwand (d. h. Aktualisierung von Metadaten) und Speicherauszüge verbrauchen bei jedem Schreibvorgang Kapazität.

Die Speicherauszugfrequenz wird daher durch Leistungserwägungen bestimmt und durch die Wirtschaftlichkeit begrenzt. Noch einmal: Ein Upgrade auf Multi-Controller-Architekturen ermöglicht im Hinblick auf die zuvor besprochene Leistungsfähigkeit häufige Speicherauszüge. Ferner erfolgt die Wiederherstellung nach einer Datenbeschädigung schneller, als die Wiederherstellung auf einem Array



mit zwei Controllern, das durch CPU-Zyklen oder Bandbreite eingeschränkt ist. Die Konsolidierung vieler Speicher-Arrays in weniger Multi-Controller-Arrays mit weniger und größeren Speicherpools können die Möglichkeiten von Multi-Controller-Array \$/TB mit Preisauflägen durch Verringerung der gestrandeten Kapazität und Komplexität, erhöhter Mitarbeiterproduktivität und kostenverminderten Ausfallzeiten zunichte machen.

## Speicher-Array Verfügbarkeitsgrundlagen

### REALES VS. ERKANNTES RISIKO

Gleichung 2 setzt Risiko, Fehlerbereiche und Verfügbarkeit in einer Weise in Beziehung, die dem gesunden Menschenverstand entspricht. Eine Vergrößerung des Wirkungsbereichs oder der Fehlerdomäne erhöht das Risiko und erhöhte Verfügbarkeit senkt das Risiko. Da alles unter 100 % Verfügbarkeit ein echtes Ausfallrisiko darstellt, dürfen Speicherarrays keine Single Points of Failure (SPOFs) haben, müssen fehlertolerant sein und alles unterbrechungsfrei bereitstellen: Software-Updates, Reparaturmaßnahmen und Kapazitätserweiterungen.

#### **Gleichung 2 – Risiko**

**Risiko = Wirkungsbereich x (1- Verfügbarkeit)**

Bleiben wir beim Beförderungswesen: Wenn ein Supertanker auseinander bricht, ist es eine Katastrophe; wenn ein Tanklaster verunglückt, ist es schlimm, aber selten eine Katastrophe. Dadurch wird es einfacher mehr in die Sicherheitsausstattung eines Supertankers zu investieren, als in einen Tanklaster, und der Schutz eines Tanklasters ist von Natur aus schwieriger als der Schutz des Supertankers.

### VERFÜGBARKEIT

Gleichung 3 definiert das Verhältnis zwischen Verfügbarkeit, MTBF und MTTR. Das unterstreicht, dass das Erreichen einer 100%igen Datenverfügbarkeit Fehlertoleranz erfordert und dass alle Reparaturmaßnahmen im Hintergrund laufen. Sie zeigt auch, dass niedrige MTBFs zu häufigen Reparaturen und nicht zu Ausfallzeiten führen.

#### **Gleichung 3 – Verfügbarkeit**

**Verfügbarkeit = MTBF/(MTBF + MTTR)**

Für diejenigen, die technische Risiken der Speicherkonsolidierung evaluieren, aber nicht an einer detaillierteren Analyse der Fehlermathematik interessiert sind, erhöht die Konsolidierung in weniger High-End-Speicher-Arrays mit mehreren Controllern die MTBDLs, indem sie die Wiederherstellungszeiten verkürzt, die Nutzverfügbarkeit, indem sie die Auswirkungen von Hardware-Ausfällen prozentual verringert, und die Häufigkeit von Reparaturen, indem sie die Teileanzahl Ihrer Speicherinfrastruktur reduziert.

## Richtlinien für das Design von Systemen mit hoher Verfügbarkeit

Die Verfügbarkeit der Speicher-Array-Hardware wird durch die in Tabelle 2 aufgeführten Faktoren bestimmt. Es ist allerdings wichtig zu beachten, dass eine 100%ige Verfügbarkeit des Speicher-Arrays keine Garantie für verlustsichere Daten sind. Der Schutz vor Datenverlust wird durch RAID- und Löschkodierungsschemata gewährleistet, wobei der Schutzgrad durch Leistungs- und Kostenerwägungen begrenzt ist, die später noch untersucht werden.

TABELLE 2

Faktor	Bemerkung
Komponentenanzahl im System	Falls nicht vorhanden, ist kein Zusammenbruch möglich. Dies und Kostenerwägungen sind die wichtigsten Triebfedern für die Vereinfachung und Verringerung der Systemteileanzahl.
Die durchschnittliche Dauer zwischen den Ausfällen jeder Komponente	Hochwertige Komponenten fallen seltener aus als Verbraucherkomponenten <sup>4</sup>
Die Anzahl der kritischen Ausfallmodi im Array <sup>5</sup>	Keine SPOFs ist das angestrebte Ziel, denn Software wird immer ein SPOF bleiben, auch wenn kritische Hardware-Ausfälle beseitigt sind.
Durchschnittliche Reparaturdauer	Eine 100%ige Verfügbarkeit erfordert unterbrechungsfreie Reparaturen.

Es gibt keine systemeigenen Verfügbarkeitsvorteile von Scale-Up- gegenüber Scale-Out-Architekturen, da die in Infinidat und vielen anderen Speicher-Arrays integrierten Funktionen für Fehlertoleranz, Wiederherstellung und unterbrechungsfreie Reparaturen die Verfügbarkeit praktisch von Hardware-Ausfällen und Software-Updates entkoppelt haben. Dadurch verlagert sich der Schwerpunkt von der Verbesserung der Speicherverfügbarkeit auf die Verbesserung der Softwarequalität, der Wiederherstellungsfunktionen und die Reduzierung der Anzahl der für den Aufbau des Arrays erforderlichen Teile.

Die Reduzierung der Teilanzahl verringert nicht nur die Kosten, sondern auch die Häufigkeit von Hardware-Reparaturen, wodurch die Fehlerquellen der Außendiensttechniker, die das Speichersystem zum Stillstand bringen, verringert werden. Die Analyse der Softwarequalität und der Wiederherstellungsfähigkeiten ist eine von Natur aus subjektive Analyse, da sie viele Unbekannte enthält. Daher wird die Häufigkeit von Software-Updates, ohne funktionale Verbesserungen, als Analogie zur Codequalität verwendet.

Da die MTBF der Controller-Elektronik nicht durch eingebrachte IOPS beeinflusst wird, ist es offensichtlich, dass die Teileverringern durch vermehrte Einbringung von IOPS durch jeden Controller ersetzt werden sollte. Die Anzahl der IOPS, die ein Controller unterstützen kann, wird von der CPU-Leistung und der Software-Effizienz bestimmt.

Ein Vergleich der jährlichen Verbesserungen der CPU-Leistung mit den Verbesserungen der HDD-Datenübertragungsrate zeigt, dass die Verbesserungen bei den Mikroprozessoren die Verbesserungen bei den Medien übertreffen: ca. 40 % pro Jahr bei Mikroprozessoren gegenüber 10 % bis 15 % pro Jahr bei HDDs. Dieser komparative Vorteil verringert das Verhältnis von CPU zu Kapazität, das zur Vermeidung von Leistungsengpässen erforderlich ist. Der Unterschied zwischen der CPU-Leistungsverbesserungen und Verbesserungen der Datenübertragungsrate begünstigt den weiteren Aufbau von Scale-Up-Arrays, die auch Scale-Out-Fähigkeiten besitzen, falls sie letztendlich benötigt werden.

Scale-Up-Arrays sind durch die Reparaturfrequenz und Kostenvorteile begünstigt, da sie im Allgemeinen weniger Controller und unterstützende Elektronik – HBAs, NICs, Netzteile, Lüfter usw. – haben, als Scale-Out-Arrays mit gleicher Kapazität. Dieser Vorteil nimmt im Allgemeinen mit der Kapazität zu, da Scale-Out-Arrays die Kapazität häufig durch das Hinzufügen von Knoten mit Controllern und deren unterstützender Elektronik erhöhen. Die Nutzung von weniger Komponenten mit niedrigerem Warenaufwand und die Umweltvorteile von Scale-Up- gegenüber Scale-Out-Arrays erklären den anhaltenden Erfolg von Scale-Up-Arrays auf dem Markt.

<sup>4</sup> Herstellungsfehler reduzieren die MTBF der Komponenten und scheinen daher nicht in Tabelle 1 auf

<sup>5</sup> Kritische Ausfälle sind Ausfälle, die das Array zum Absturz bringen oder störende Reparaturen erfordern.

## Durchschnittliche Dauer zwischen Datenverlusten

MTBDLs sind für einen zuverlässigen Rechenzentrumsbetrieb mindestens ebenso wichtig wie die Systemverfügbarkeit, da sie zu länger andauernden Vorfällen führen können als Softwarefehler oder Hardwareausfälle. Die meisten Anbieter beantworten Fragen zur Zuverlässigkeit oder Verfügbarkeit mit Angaben zwischen 99,99 % und 100,00 % Verfügbarkeit, halten sich aber im Hinblick auf Wiederherstellungszeiten und voraussichtliche Reparaturmaßnahmen bei HDD und SSD bedeckt. Die Wiederherstellungszeiten und MTBDLs sind untrennbar miteinander verbunden, und die Wiederherstellungszeiten können sehr lang werden, Stunden bis Tage in stark ausgelasteten Arrays mit traditioneller Architektur.

Nachfolgend die Faktoren, die MTBDLs beeinflussen.

- ▶ Die MTBF der SSDs oder HDDs, auf denen Daten gespeichert sind – eine höhere MTBF (d. h. die Qualität der Komponenten) verringert die Häufigkeit von Reparaturen und die Wiederherstellungszeiten der Daten.
- ▶ Die Gesamtzahl der SSDs und HDDs im Speicherarray – weil sie die Wahrscheinlichkeit des Auftretens mehrerer SSD- oder HDD-Ausfälle innerhalb eines RAID oder einer Löschruppe beeinflusst.
- ▶ Die Anzahl der SSDs oder HDDs in einem RAID oder einer Löschruppe – mehr Teile bedeuten mehr Ausfälle, mehr Reparaturen und mehr Zeit für die Datenwiederherstellung, da größere RAID- oder Löschruppen mehr Daten enthalten als kürzere.
- ▶ Die Anzahl der Ausfälle, die in einer RAID-Gruppe toleriert werden können – RAID 1, 10 und 5 garantieren die Datenintegrität bei Ausfällen einer einzelnen Festplatte oder SSD; RAID 6 garantiert die Datenintegrität bei zwei HDD- oder SSD-Ausfällen. Löschruppen können, da sie Daten als Gleichungssysteme speichern, vor einer beliebigen Anzahl von Ausfällen schützen, wenn ihre Auswirkungen auf Leistung und Kosten vertretbar sind.
- ▶ Wiederherstellungszeiten – intuitiv schätzen wir, dass kürzere Wiederherstellungszeiten zu einer höheren MTBDL führen, aber ihr tatsächlicher Einfluss auf die MTBDL wird oft stark unterschätzt, da es sich um eine nicht einprägsame Variable handelt.

## Anschaffungs- und Betriebskosten

Der Wettbewerb zwischen lokalen Speicheranbietern und Cloud-Anbietern hat das Vermögen der Anbieter, unterschiedliche Preise für HDD- und SSDs in High-End-Arrays und Mid-Range-Arrays zu verlangen, erheblich eingeschränkt. Daher werden die Unterschiede bei den Anschaffungs- und Betriebskosten zwischen Arrays mit zwei und mehreren Controllern zunehmend durch Unterschiede bei den Controller-Kosten, der Hardwarewartung und den einmaligen/jährlichen Software-Lizenzgebühren beeinflusst. Daher schrumpft das \$/TB-Kostendelta zwischen Dual- und Multi-Controller-Arrays mit wachsenden Konfigurationen, insbesondere bei Multi-Petabyte-Skalen.

Angesichts der Tatsache, dass Speicher-Arrays in der Vergangenheit hohe Listenpreise, komplexe Preismodelle und aggressiv ausgehandelte Einmal-Rabatte hatten, gibt es keine einfachen Faustregeln (ROT), mit denen sich die tatsächlichen \$/TB/Monat-Preise mit hoher Genauigkeit abschätzen lassen. InfiniBox mit seinen Reaktionszeiten im Sub-Millisekundenbereich, der Multi-Petabyte-Größe, dem Software-Komplettpreismodell und den umwälzenden \$/TB-Preisen erschwert die Erstellung einfacher ROTs für \$/TB-Preise ebenso wie die starke Nutzung von Replikationssoftware, Management-Tools und Skripten, die starke Abhängigkeiten schaffen, selbst wenn sie die architektonische Hässlichkeit von Arrays verbergen, deren Hinterlassenschaften sich über Jahrzehnte erstrecken könnten.

# Best Practices für die Speicherkonsolidierung

Speicherkonsolidierungsprojekte scheitern häufig, weil sie den Status quo in Frage stellen, und Veränderungen schaffen Gewinner, Verlierer und Risiken. Im Folgenden werden einige der wichtigsten bewährten Verfahren aufgeführt, bei der die Wahrscheinlichkeit am höchsten ist, dass Konsolidierungsprojekte erfolgreich abgeschlossen werden.

- ▶ Holen Sie sich die Unterstützung der Unternehmensleitung ein, bevor Sie mit einem Speicherkonsolidierungsprojekt fortfahren.
- ▶ Stellen Sie ein Team zusammen, das wichtige Interessengruppen umfasst: Speicherarchitekten, Betriebsabläufe, Entwickler, Finanzen und Rechtsabteilung.
- ▶ Beziehen Sie Kostenänderungen für die Anschaffung und den Besitz der Infrastruktur, Kosten für Ausfallzeiten, Migrationskosten und Produktivitätsverbesserungen in Ihre Risiko-/Ertrags- und ROI-Analyse ein.
- ▶ Die Anbieter sollten ein Eigeninteresse am Erfolg Ihrer Konsolidierungsprojekte haben und daher sollten die Kosten und Risiken der Datenmigration aufgeteilt werden.
- ▶ Betonen Sie die Vorteile der Speicherkonsolidierung, um den Support unter den Betreibern aufzubauen.

## Fazit

Im Multi-Petabyte-Bereich macht die Konsolidierung von Speicher auf Multi-Controller-Arrays, insbesondere Arrays mit einer Software-Komplettpreisgestaltung, Capacity on Demand (Kapazität bei Bedarf -COD) und verbrauchsabhängigen Preismodellen, die Speicherkonsolidierung zur bestmöglichen Geschäftsentscheidung. Bei dieser Entscheidung sollten die Auswirkungen auf Verfügbarkeit, Leistung, Mitarbeiterproduktivität und Gesamtbetriebskosten berücksichtigt werden.

Anbieter, die im Preiswettbewerb mithalten und bereit sind, die Array-Konfiguration und Datenmigration zu übernehmen, erhöhen zusätzlich die Wahrscheinlichkeit eines erfolgreichen Konsolidierungsprojekts. Infinidat ist ein solcher Anbieter von Speicherlösungen.



**STANLEY ZAFFOS** Sr. VP, Product Marketing, Infinidat

Bevor er zu Infinidat wechselte, war Stanley Zaffos Research VP bei Gartner mit dem Schwerpunkt Infrastruktur- und Betriebsmanagement. Seine Fachgebiete sind Speichersysteme, neue Speichertechnologien, softwaredefinierte Speicherung, hyperkonvergente Infrastruktur und hybride Cloud-Infrastruktur. Er arbeitete mit zahlreichen Kunden, entwickelte Messaging und Begleitmaterial, das auf eine Maximierung der Produktankündigungen und Vertriebsschulungen abzielte, war beratend bei der Ausarbeitung von Fahrplänen, die einen anhaltenden Wettbewerbsvorteil gewährleisten, tätig.