QNAP

QuTS hero
ZFS

Whitepaper

INDEX

Seite	Titel
01	**Übersicht**
02	**ZFS Einleitung und Konfiguration**
02	Datenintegrität und Selbstheilung
02	Empfohlene Blockgröße für Freigabeordner/ LUN Synchrone
03	E/A für ZIL (Auto / Alle / Keine)
04	Pool Overprovisioning
06	**Datenreduktion (Inline-Komprimierung / Inline-Deduplizierung)**
07	Inline-Komprimierung
08	Inline-Deduplizierung
09	**SnapSync** Günstige Replikationslösung für Sicherung, Datenschutz und Notfallwiederherstellung.
09	Einleitung
09	Funktionsweise von SnapSync
11	Was zu tun ist, wenn ein Notfall eintritt
12	Bewährte Praktiken für die Konfiguration von Realtime SnapSync
14	Vorteile
15	**QSAL für All-Flash Arrays und SSD-RAID (QNAP-patentierte Technologie)**
16	**Schreib-Koaleszenz** Der Schreib-Koaleszenz-Algorithmus steuert die zufällige Schreibleistung für All-Flash Arrays
17	**Empfohlene Lese-Cache-Konfiguration (L2ARC)**
19	**Leistungsabhängigkeit zwischen Arbeitsspeichergröße und Speicherpools**

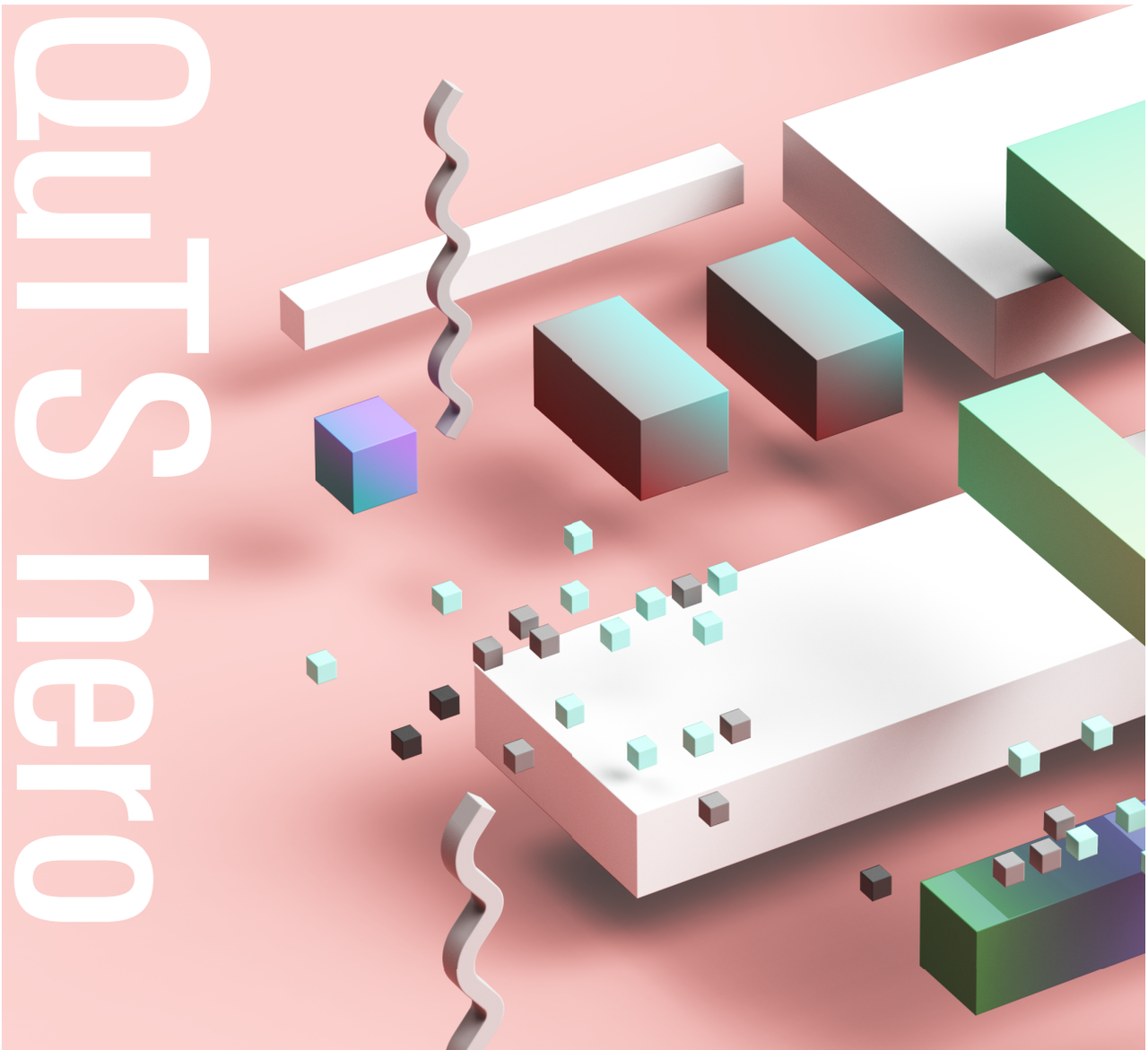
QuTS hero ZFS

Whitepaper

Übersicht

In der heutigen IT-intensiven Welt können sich schnell ändernde Anforderungen über Erfolg oder Misserfolg von Unternehmensorganisationen und Rechenzentrumsbetrieben entscheiden. In diesem Umfeld stehen die Budgets auf dem Prüfstand und jede potenzielle Ausfallzeit - ob Minuten oder Sekunden - ist potenziell verheerend. Jede potenzielle Steigerung der Effizienz und Produktivität sind wichtige Eckpfeiler für den Erfolg.

Das QuTS hero Betriebssystem verwendet ZFS, ein Dateisystem, das die Datenintegrität priorisiert, um die Anforderungen an eine leistungsstarke und hochstabile Datensicherung zu erfüllen. QuTS hero integriert hochleistungsfähige Datensicherungs-, Datenreduzierungs- und Virtualisierungsanwendungen und unterstützt Cloud-Umgebungen und QoS (Quality of Service)-Verwaltungsfunktionen, um günstigere Speicherlösungen für kleine Unternehmen, Rechenzentren auf Unternehmensebene und virtualisierte Arbeitsabläufe zu bieten.

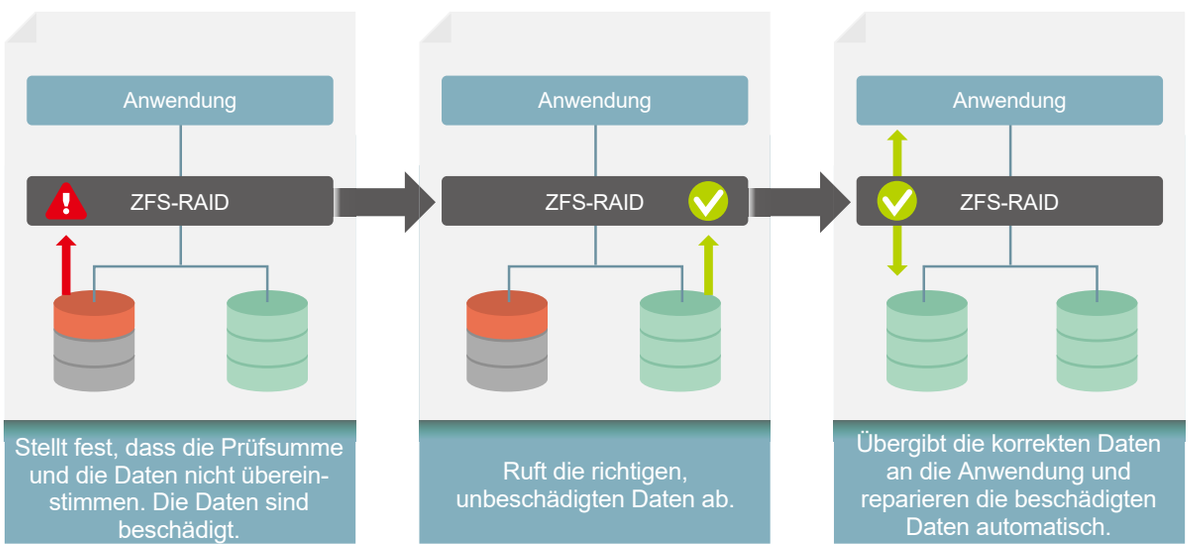


ZFS Einleitung und Konfiguration

Datenintegrität und Selbstheilung

Die Integrität und Vertrauenswürdigkeit von Daten ist der Schlüssel zur Gewährleistung des Betriebs von Anwendungen und Datenbanken. ZFS hat die Fähigkeit, schleichende Datenbeschädigung (Silent Data Corruption) zu verhindern, einen potenziell katastrophalen und schwer zu erkennenden Zustand, der durch Hardwareprobleme, defekte Kabel oder Metadatenfehler verursacht wird. ZFS verfügt über die Fähigkeit zur Selbstheilung dieser Daten, einschließlich der Überprüfung aller Datenblöcke und der automatischen Reparatur fehlerhafter Blöcke.

Es wird dringend empfohlen, regelmäßig ein "Pool Scrubbing" durchzuführen, um die Integrität der gespeicherten Daten sicherzustellen und potenzielle Probleme zu beheben, bevor sie auftreten. Sie können Pool Scrubbing so konfigurieren, dass es nach einem Zeitplan ausgeführt wird, z.B. wöchentlich oder monatlich.



Verhindert, dass schleichende Datenbeschädigung in ein laufendes System eindringt.

Empfohlene Blockgröße für Freigabeordner/ LUN:

128K - Videobearbeitung / Große Dateien / Sicherung
64K - Generisch / Hyper-V
32K - VMware
8k - OLTP
4K - VDI / Datenbank / Kleine Dateien

Synchroner E/A für ZIL (Standard / Always / None):

Das ZIL (ZFS Intent Log) ist der Schreibjournal-Mechanismus, der Dateiänderungen verfolgt, die noch nicht in das Dateisystem übertragen wurden. Das ZIL zeichnet die Absichten dieser Änderungen in einem Protokoll auf, das als Journal bezeichnet wird. Im Falle eines Systemabsturzes oder Stromausfalls überprüft QuTS hero die Journalprotokolle und wendet die geplanten Änderungen an, wodurch sichergestellt wird, dass das Dateisystem schnell und mit einer geringen Wahrscheinlichkeit von Datenbeschädigungen wiederhergestellt wird.

Wenn ein Freigabeordner/LUN erstellt wird, muss der ZIL-Schreibprotokoll-E/A-Modus konfiguriert werden, um die Datenintegrität oder die Leistung zu priorisieren.

Standard (Default): E/A-Transaktionen sind abhängig von der Anwendung und der Art der E/A-Anforderung synchron oder asynchron.

Always: Alle E/A-Transaktionen werden als synchron behandelt, was die Datenintegrität verbessert, aber die Leistung verringert.

None: Alle E/A-Transaktionen werden als asynchron behandelt. Diese Option bietet die höchste Leistung, birgt aber ein höheres Risiko von Datenverlusten bei einem Stromausfall.

Bitte beachten Sie, dass es empfohlen wird, ein individuelles ZIL als SLOG auf einem dedizierten High-Speed-Speichergerät (z.B. einer NVMe SSD) zu konfigurieren. Für das ZIL wird der Einsatz von Laufwerken mit extrem niedriger Latenz und hoher Lebensdauer (z.B. MLC-Laufwerke) empfohlen. Für den Lesecache sollte der Durchsatz priorisiert werden. Es wird empfohlen, das ZIL und den Lesecache aufzuteilen, um einen vorzeitigen Ausfall der gesamten SSD-Gruppe aufgrund der hohen ZIL-Schreibfrequenz zu vermeiden.

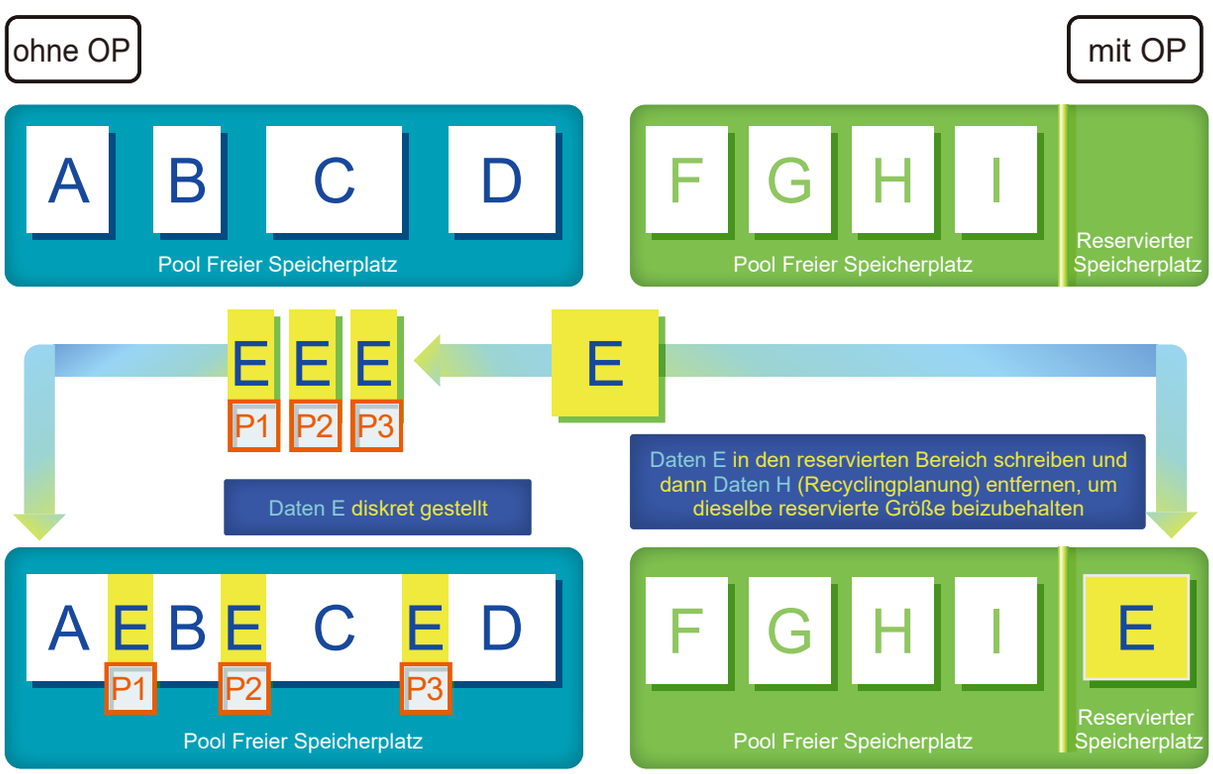
Pool Overprovisioning

Die ZFS-Leistung kann beeinträchtigt werden, wenn ein Speicherpool fast voll oder fragmentiert ist. Dies ist auf das Copy-on-Write Verhalten zurückzuführen: ZFS schreibt neue Daten immer in einen anderen Block, aktualisiert die Metadaten, um auf den neuen Speicherort zu verweisen, und gibt dann den alten Block frei, anstatt die alten Daten direkt zu überschreiben. Nach vielen Schreibvorgängen wird die Verteilung der Daten auf dem Speicherpool fragmentiert, wodurch das Schreiben neuer Daten schwieriger und zeitaufwändiger wird.

QuTS hero unterstützt Pool Overprovisioning, mit dem Sie einen bestimmten Prozentsatz des Speicherplatzes im Speicherpool festlegen und reservieren können, um eine konsistente zufällige Schreibleistung zu gewährleisten, wenn der Pool fast voll ist. Die folgende Abbildung zeigt, wie Daten in einen Speicherpool mit und ohne Pool Overprovisioning geschrieben werden. Wenn der Speicherpool fast voll und fragmentiert ist, werden die neu geschriebenen Datenblöcke Ohne Pool Overprovisioning eine Weile getrennt und die Leistung beeinträchtigt. Mit Pool Overprovisioning wird es für ZFS möglich, einen kontinuierlichen großen Block zu finden, um neue Daten zu speichern.

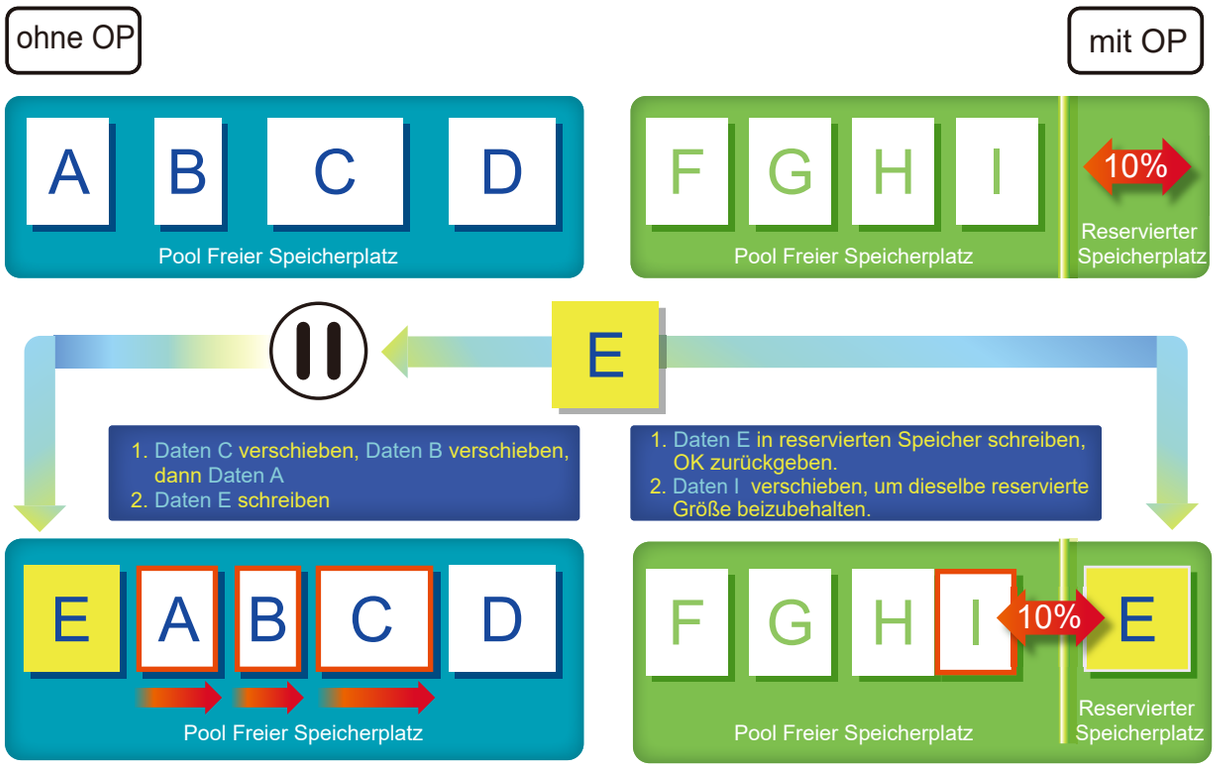
Overprovisioning

Verbessert die Leistung von fragmentierten Pools (z.B.: beim Schreiben eines Blocks auf eine **HDD**)



Overprovisioning

Verbessert die Leistung von fragmentierten Pools (z.B.: beim Schreiben eines Blocks auf eine **SSD**)



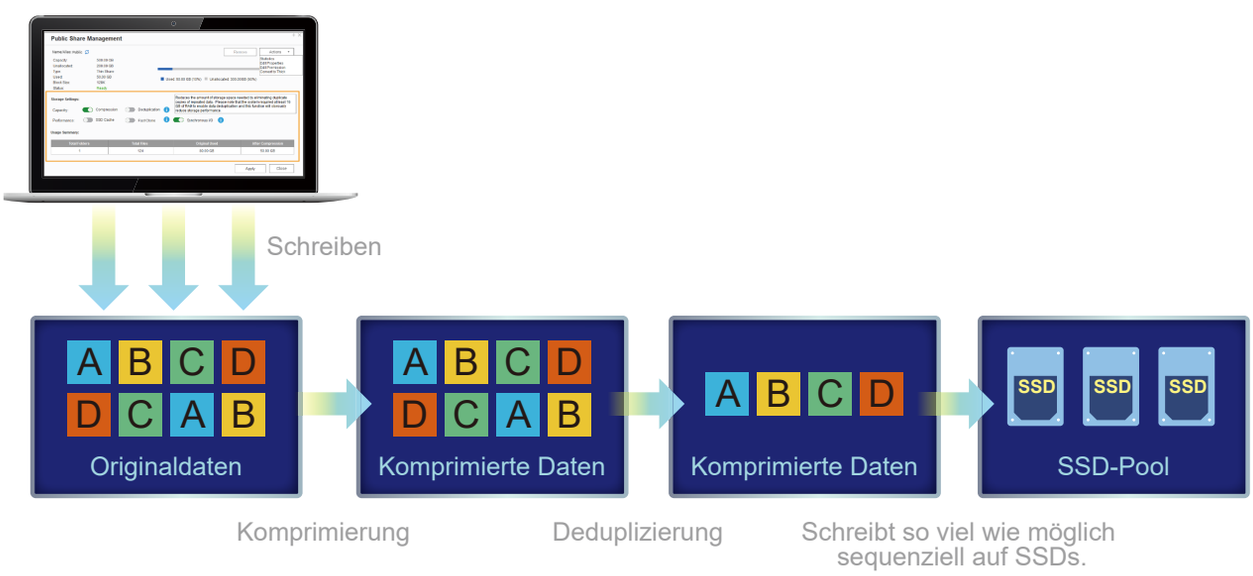
Das Aktivieren von Pool Overprovisioning sorgt für eine zufällige Schreibleistung auf Kosten der Speicherkapazität. Je mehr Speicherplatz reserviert wird, desto höher ist die Leistung. Der Schlüssel zur Bestimmung, ob ein solcher Kompromiss notwendig ist, liegt in der Betrachtung des Lese-/Schreibverhältnisses und der Anwendungsszenarien. Je mehr Schreibvorgänge anfallen, desto höher ist der Leistungsschub, den Pool Overprovisioning bietet.

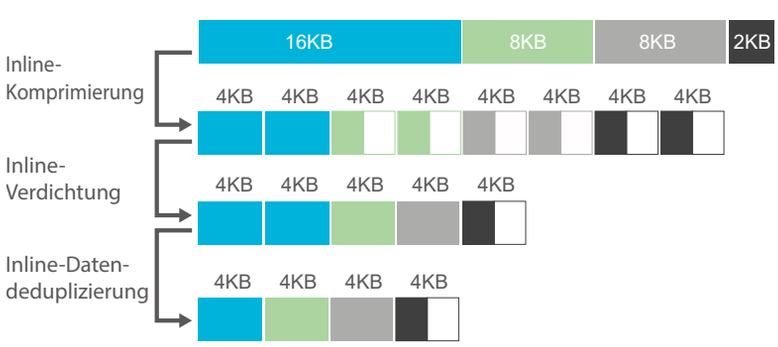
Es wird im Allgemeinen empfohlen, den Wert für Pool Overprovisioning auf 10-20% einzustellen. Für schreibintensive und leistungsfordernde Anwendungen (wie SQL-Protokolle, Datensicherung, Überwachung und medizinische Bildgebung) empfiehlt es sich, einen höheren Wert einzustellen. Für leseintensive Anwendungen (wie Webserver, OLAP und Medienstreaming) ist die Leistungssteigerung durch Pool Overprovisioning den Verlust an Speicherplatz möglicherweise nicht wert und es kann ein niedrigerer Wert verwendet werden.

Datenreduktion (Inline-Komprimierung / Inline-Deduplizierung)

Mit der Inline-Datendeduplizierung können sich wiederholende Daten vor dem Schreiben gelöscht werden, wodurch viel Speicherplatz gespart werden kann. Die Inline-Komprimierung wird ebenfalls unterstützt, um beim Speichern großer Dateien Speicherplatz zu sparen. Diese Funktionen bieten zusammen mit der Inline-Verdichtungstechnologie die Möglichkeit, die Speichernutzung erheblich zu optimieren. Beim Einsatz in All-Flash-Konfigurationen, die stark wiederkehrende Daten oder große Mengen kleiner Dateien speichern, ist die Möglichkeit die Speicheranforderungen zu reduzieren, von großem Vorteil. Sie verbessert nicht nur die Schreibleistung von All-Flash-Konfigurationen erheblich, sondern verlängert auch die Lebensdauer von SSDs und maximiert die allgemeine Kosteneffizienz des Einsatzes.

Die Komprimierungsrate kann bei sich wiederholenden virtualisierten Umgebungen 90% erreichen.





Inline-Komprimierung

QuTS hero unterstützt Inline-Komprimierung, die Datei- und Blockgrößen reduziert, indem sie mit weniger Bits als die ursprüngliche Darstellung kodiert wird, wenn die Daten in einen Freigabeordner oder eine LUN geschrieben werden. Die Aktivierung der Komprimierung erhöht auch den Lese- und Schreibdurchsatz, da weniger Blöcke gelesen oder geschrieben werden müssen.

In QuTS hero ist die Komprimierung verlustfrei, um die Datenkonsistenz zu gewährleisten. ZFS unterstützt standardmäßig die LZ4-Komprimierung, da sie im Allgemeinen die schnellste ist.

Diese Komprimierung erfordert keine übermäßigen CPU-Ressourcen und wird für die Verwendung mit Freigabeordnern und LUNs empfohlen. Die Komprimierung sollte nicht für Dateien verwendet werden, die bereits komprimiert wurden (z.B. MP3-Dateien, H.264 Videos und ZIP-Dateien), da der durch erneute Komprimierung eingesparte Speicherplatz vernachlässigbar ist.

LZ4 ist ein verlustfreier Datenkomprimierungsalgorithmus, der extrem schnelle Geschwindigkeiten bei der Komprimierung und Dekomprimierung bietet. Er ist die beste Wahl für niedrige Latenz und hohe EA, wenn die Übertragungsgeschwindigkeiten höher als 100 MB/s sind. Er ist besonders für die Inline-Komprimierung geeignet.

Komprimierungsgeschwindigkeit	Dekomprimierungsgeschwindigkeit
740 MB/S	4530 MB/S

*Getestet mit Izbench

Inline-Deduplizierung

QuTS hero verwendet zielbasierte, blockbasierte und Inline-Deduplizierung. Zielbasiert bedeutet, dass die Quellseite die Daten zuerst an die Zielseite überträgt und der Deduplizierungsprozess dann auf der Zielseite stattfindet. Blockebene bedeutet, dass kleine redundante Teile der Daten innerhalb einer Datei oder zwischen verschiedenen Dateien erkannt und entfernt werden können. Inline bedeutet, dass der Deduplizierungsprozess sofort stattfindet, während die Daten auf das Speichergerät übertragen werden, aber bevor sie geschrieben werden. Die Vorteile dieser Kombination (Zielbasiert, Blockebene und Inline) sind die Reduzierung der benötigten CPU-Ressourcen des Clients, das Erreichen höherer Deduplizierungsraten und die Reduzierung der Laufwerkszugriffsanforderungen im Vergleich zu anderen Deduplizierungsmechanismen.

Der Deduplizierungsprozess besteht aus drei Schritten: Zerlegen von Daten in Blöcke, Berechnen des Fingerabdrucks und Durchsuchen der Dedup-Tabelle. Zunächst zerlegt ZFS Dateien in kleine Datenblöcke. Wenn ein Datenblock geschrieben wird, verwendet das System eine kryptografische Prüfsumme, um den Fingerabdruck jedes Datenblocks zu berechnen. Dieser wird verwendet, um zu erkennen, ob es identische Blöcke in vorhandenen Daten gibt. ZFS verwaltet eine Zuordnungstabelle (die so genannte Dedup-Tabelle), die alle Fingerabdrücke, Speicherorte und Referenzzahlen enthält. Wenn ein Datenblock mit demselben Fingerabdruck wie ein bereits vorhandener Datenblock auf das NAS geschrieben wird, prüft ZFS die Dedup-Tabelle, weist diesem Datenblock denselben Speicherort zu und erhöht die Referenzanzahl um eins, anstatt an einen neuen Speicherort zu schreiben.

Die Inline-Deduplizierung reduziert den Bedarf an zusätzlichem Speicherplatz, erfordert aber mehr Rechenressourcen beim Schreiben von Daten, was die Systemleistung beeinträchtigen kann. Diese Auswirkung auf die Leistung wird mit zunehmender Größe der Dedup-Tabelle wachsen - ebenso wie die Schwierigkeit, die Tabelle zu durchsuchen. Um Leistungseinbußen bei der Inline-Deduplizierung zu vermeiden, verwendet QNAP den SmartDDT (Smart DeDuplication Table)-Algorithmus, um die wachsende Dedup-Tabelle einzudämmen, indem die Leistungsänderungen oder die Größe der Dedup-Tabelle überwacht werden. Wenn einige Kriterien erfüllt sind, stoppt QuTS hero den Deduplizierungsdienst. Wenn dies geschieht, sind entweder die Daten nicht sehr wiederholend und das Anwendungsszenario ist möglicherweise nicht für die Inline-Deduplizierung geeignet, oder das NAS benötigt mehr Arbeitsspeicher.

Für Umgebungen, die Deduplizierung erfordern, empfehlen wir, so viel Arbeitsspeicher wie möglich zu installieren - mindestens 16GB Arbeitsspeicher, mehr als 64GB werden dringend empfohlen.

SnapSync:

die günstige Replikationslösung für Sicherung, Datenschutz und Notfallwiederherstellung.

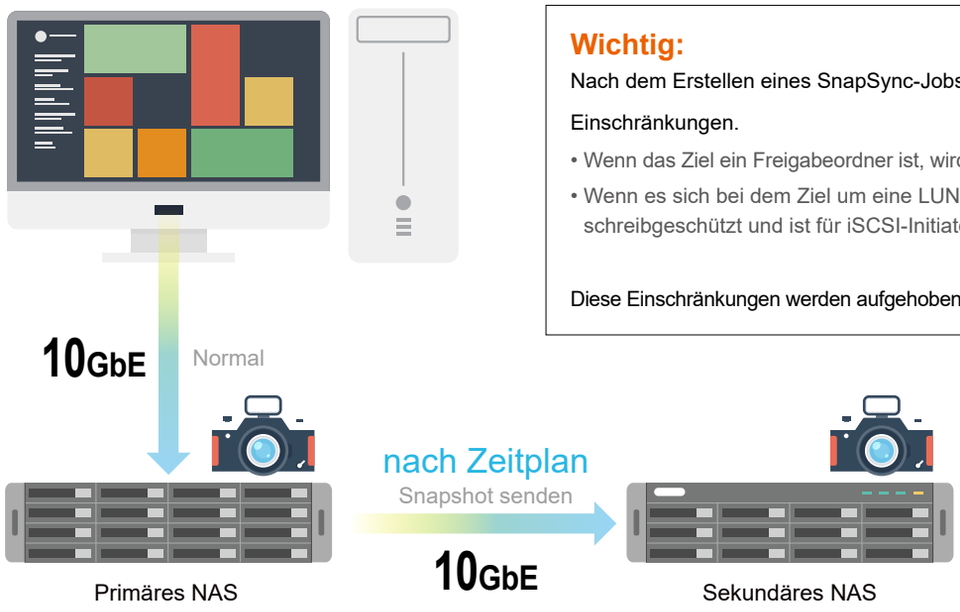
Einleitung:

SnapSync ist eine günstige und einfach zu bedienende einheitliche Replikationslösung, die in QuTS hero für Sicherung, Datenschutz und Notfallwiederherstellung integriert ist.

SnapSync wird über eine Spiegelungsbeziehung zwischen Freigabe/LUN auf primären und sekundären Speichersystemen konfiguriert. Es synchronisiert das Replikat regelmäßig, um es mit den Änderungen, die auf das primäre System geschrieben wurden, auf dem neuesten Stand zu halten. Diese gespiegelten Daten werden in einem sekundären Speichersystem erstellt. Bei einem Notfall am primären Standort können Sie die Daten vom sekundären System übernehmen. Mit SnapSync können Sie die Gesamtkosten von Notfallwi ederherstellungslösungen reduzieren und so die Investitionen leichter rechtfertigen, indem Sie Ihren Notfallwi ederherstellungsstandort für den aktiven Geschäftsbetrieb nutzen.

Funktionsweise von SnapSync:

Die SnapSync-Funktion auf Blockebene kann variable Blöcke per Snapshot sichern, was wenig Ressourcen benötigt und bandbreitenschonend ist, und wichtige Dateien können effizient an entfernten Standorten gesichert werden. Bei der inkrementellen Replikation werden identische Daten nie zweimal gesendet. SnapSync-Wiederherstellungszeiten sind auch schneller als andere Methoden, da für die Wiederherstellung nur ein Vollbackup und das letzte differentielle Backup benötigt werden.



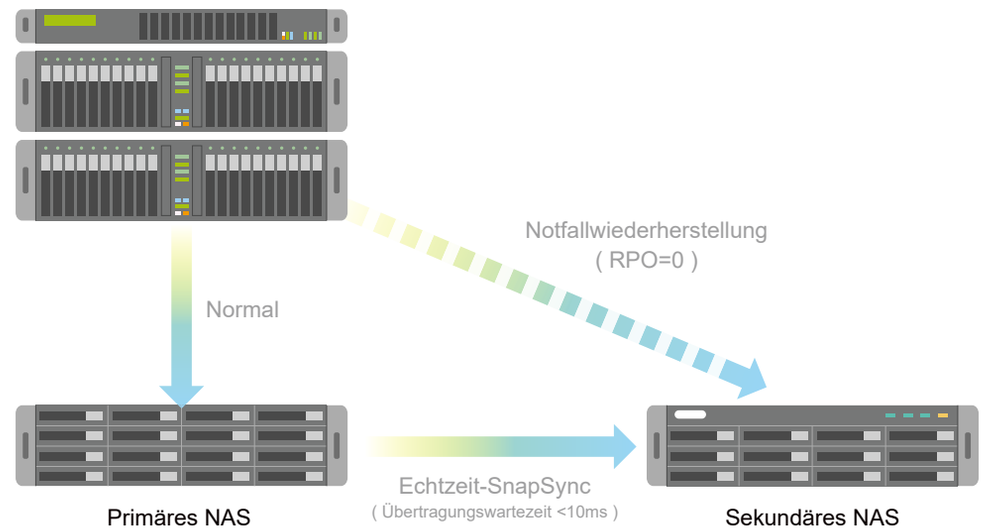
Wichtig:

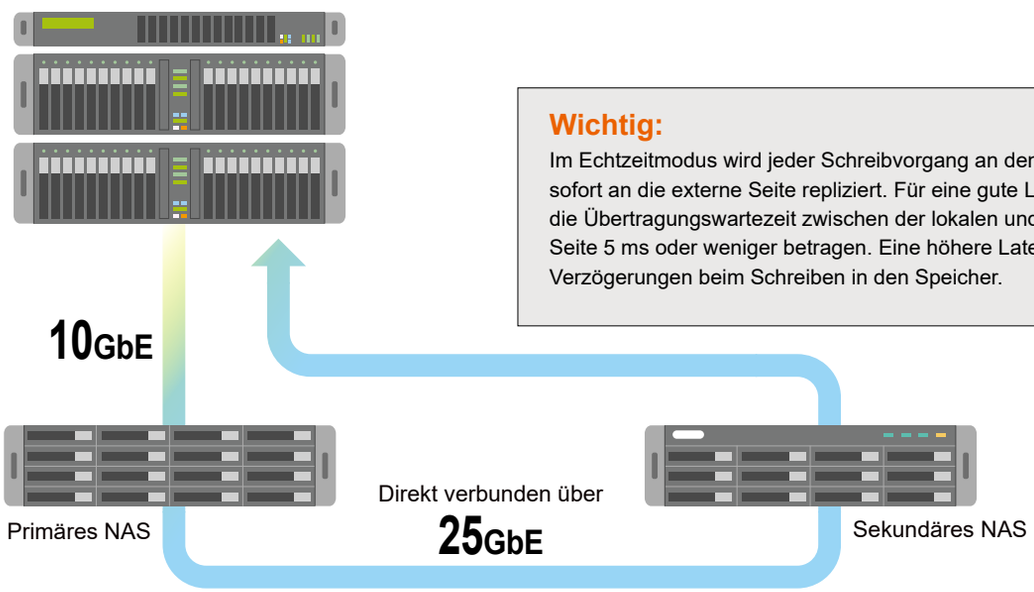
Nach dem Erstellen eines SnapSync-Jobs gelten die folgenden Einschränkungen.

- Wenn das Ziel ein Freigabeordner ist, wird der Ordner schreibgeschützt.
- Wenn es sich bei dem Ziel um eine LUN handelt, wird die LUN schreibgeschützt und ist für iSCSI-Initiatoren unzugänglich.

Diese Einschränkungen werden aufgehoben, wenn der Auftrag gelöscht wird.

Die erweiterte Echtzeit-SnapSync-Funktion ermöglicht es dem primären und sekundären NAS, jederzeit synchronisiert zu bleiben.





Wichtig:
Im Echtzeitmodus wird jeder Schreibvorgang an den lokalen Speicher sofort an die externe Seite repliziert. Für eine gute Leistung muss die Übertragungswartezeit zwischen der lokalen und der externen Seite 5 ms oder weniger betragen. Eine höhere Latenz verursacht Verzögerungen beim Schreiben in den Speicher.

Echtzeit-SnapSync stellt sicher, dass die Daten des primären und sekundären NAS vollständig geschrieben werden, bevor der Hostserver antwortet. Dadurch wird ein RPO=0 (Recovery Point Objective) gewährleistet, was auch bedeutet, dass die Daten auf dem sekundären NAS mit dem primären NAS konsistent sind. Im Notfall kann der Hostserver bei einem Ausfall des Servers sofort das Ziel auf das sekundäre NAS ändern.

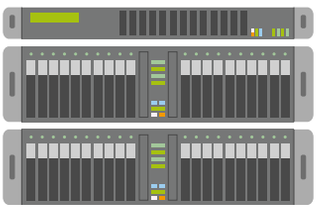
Wenn Sie ein großes Speichersystem aufbauen, wird empfohlen, das sekundäre NAS mit SnapSync einzurichten und dann Snapshot Replica & HBS3 zu verwenden, um regelmäßige Multiversionssicherungen vom sekundären NAS auf ein anderes Backup-NAS durchzuführen.

Unterstützt auch QES NAS (QES 2.1.1 v11 oder neuer) SnapSync zu QuTS hero NAS.

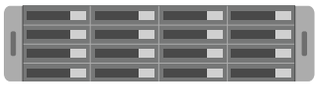
Was zu tun ist, wenn ein Notfall eintritt

Wenn ein Notfall eintritt...

Produktionsserver

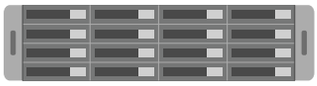


NAS (Freigabespeicher)

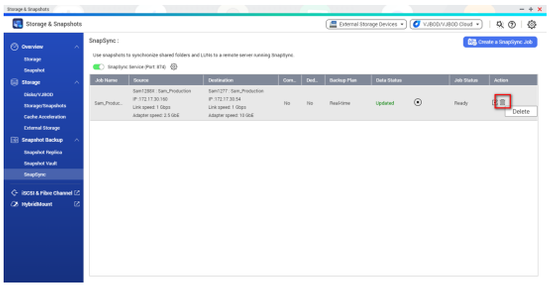


Löschen Sie zunächst die SnapSync-Aufgabe 1

Primäres NAS
192.168.100.100



Sekundäres NAS
192.168.100.200



Die exklusiven Rechte des Ordners werden beim Löschen der SnapSync-Aufgabe entfernt.

A Auf sekundäre NAS-IP wechseln

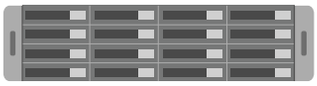
oder

B Ändern Sie die sekundäre NAS-IP auf die ursprüngliche primäre NAS-IP

Produktionsserver

Ziel einbinden = 192.168.100.200

Sekundäres NAS



192.168.100.200 => 192.168.100.100

Bewährte Praktiken für die Konfiguration von Realtime SnapSync

Im Echtzeitmodus wird jeder Schreibvorgang an den lokalen Speicher sofort an eine externe Seite repliziert. Für eine gute Leistung muss die Übertragungswartezeit zwischen der lokalen und der externen Seite 5 ms oder weniger betragen. Eine höhere Latenz führt zu Schreibverzögerungen im lokalen Speicher. Achten Sie bei der Konfiguration daher auf Folgendes:

- Es wird empfohlen, dass die Latenz der Netzwerkumgebung so weit wie möglich unter 10 Millisekunden liegt (optimal sind weniger als 5 Millisekunden)
- Die E/A-Leistung des sekundären NAS sollte die gleiche sein wie die des primären NAS.
- Wir empfehlen die direkte Verbindung des primären und sekundären NAS über 25GbE.

Bewährte Praktiken für die Konfiguration

Der Leistungsbenchmark von Echtzeit-SnapSync von QNAP Labs:

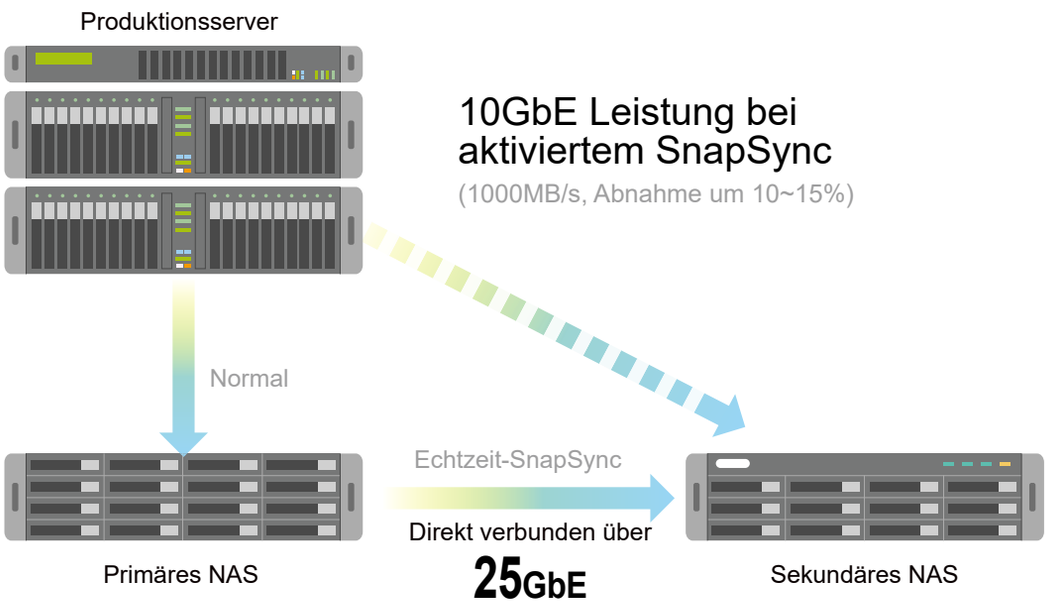
- Getestet mit dem iSCSI-Protokoll: Wenn der Benutzer eine Verbindung zum QuTS hero NAS mit 10GbE herstellt, kann die Dateiübertragung von 10GbE 1127 MB/s erreichen, bevor SnapSync aktiviert ist. Die Effizienz beträgt 1005MB/s, wenn SnapSync in Echtzeit durchgeführt wird. Die Dämpfung beträgt etwa 10~15%.

- Getestet mit dem SMB-Protokoll: Wenn der Benutzer eine Verbindung zum QuTS hero NAS mit 10GbE herstellt, kann die Dateiübertragung von 10GbE 1025 MB/s erreichen, bevor SnapSync aktiviert ist. Die Effizienz beträgt 930MB/s, wenn SnapSync in Echtzeit durchgeführt wird. Die Dämpfung beträgt etwa 10~15%.

- Getestet mit dem NFS-Protokoll: Wenn der Benutzer eine Verbindung zum QuTS hero NAS mit 10GbE herstellt, kann die Dateiübertragung von 10GbE 961 MB/s erreichen, bevor SnapSync aktiviert ist. Die Effizienz beträgt 897MB/s, wenn SnapSync in Echtzeit durchgeführt wird. Die Dämpfung beträgt etwa 7~15%.

Vor-SnapSync Schutz

Nach Aktivierung von Echtzeit-SnapSync



Testumgebung:

- EA Modus : sync = Standard & sync = keine
- Blockgröße = 128K
- Jumbo Frame (MTU) = 9000
- Wenn der obige Test mit 10GbE in der gleichen Umgebung verbunden wird, erhöht sich die Dämpfung aufgrund des Einflusses der Netzwerkverbindung und der Verzögerung nach der Aktivierung von Echtzeit-SnapSync wieder auf etwa 20%.

Vorteile:

Reduzieren Sie Ausfallzeiten und schützen Sie sich vor Datenverlusten

SnapSync bietet geplante und Echtzeit-Replikation. Dies ermöglicht Recovery Point Objectives (RPO) von Sekunden. Sie können selbst für Ihre größten schreibintensiven Arbeitslasten strenge Wiederherstellungsziele einhalten.

Reduzieren Sie die Auslastung der Netzwerkbandbreite:

SnapSync nutzt die Speichereffizienz, indem inkrementelle Blöcke mit Komprimierung/Deduplizierung über das Netzwerk gesendet werden, um die Datenübertragung zu beschleunigen und die Bandbreitennutzung zu reduzieren. Mit SnapSync können Sie einen Thin-Replikationsdatenstrom nutzen, um ein einziges Repository zu erstellen, das den aktiven Mirror verwaltet.

Verteilen Sie große Datenmengen ganz einfach:

Manchmal müssen Sie große Datenmengen übertragen: um Serverraum-Arrays zu migrieren, Außenstellen zu konsolidieren oder eine neue Niederlassung einzurichten. SnapSync bietet eine schnelle, effiziente und flexible Methode zum Verschieben von Daten. Wenn Ihr Unternehmen geografisch verstreut ist und alle Standorte Zugriff auf denselben Datensatz benötigen (z.B. Schulungsvideos oder Verkaufspakete), können Sie mit SnapSync schnell dieselben Daten an alle Standorte verteilen.

CDM (Copy Data Management) und Datenanalyse:

Die Durchführung umfangreicher Analysen mag für Ihr Unternehmen von entscheidend sein, kann aber die Leistung von Produktionsumgebungen beeinträchtigen. Mit SnapSync und Snapshot können Sie die replizierten Daten nutzen, um komplexe Analysen auf sekundären Datenkopien durchzuführen.

Anforderungen an Datenaufbewahrung, Compliance und mehrere Versionen:

Viele Anwendungsfälle erfordern lange Datenaufbewahrungsfristen. Durch die Kombination von SnapSync und Snapshots können Sie Compliance-Anforderungen zum Schutz von Daten erfüllen. Snapshot-Sicherungen können dabei helfen, die Auswirkungen von Ransomware abzuschwächen, und durch SnapSync in Echtzeit die Verfügbarkeit und die Notfallwiederherstellung zu verbessern.

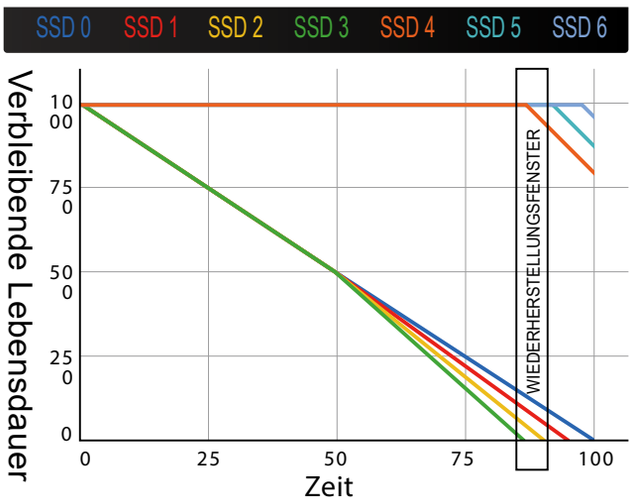
QSAL für All-Flash Arrays und SSD-RAID (QNAP-patentierte Technologie)

Der patentierte QSAL (QNAP SSD Anti-wear Leveling)-Algorithmus wird automatisch aktiviert und verhindert, dass mehrere SSDs gleichzeitig ausfallen.

Je nach SSD wird RAID 5 / 6 / 50 / 60 / TP (Triple Parity) standardmäßig aktiviert.

M0 OP = 0

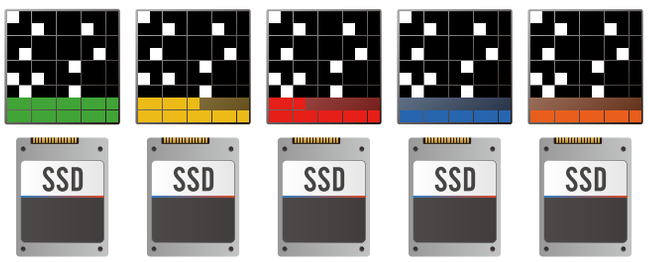
Mi OP = i x W / Mi-1, wenn (Mi-1 - Mi) > W



SSD-RAID, das von Konkurrenzprodukten verwendet wird (insbesondere solche, die nur die Belastung der ältesten SSD erhöhen), birgt ein besonders hohes Risiko von Datenverlusten zu einem späteren Zeitpunkt, da die Hälfte des Lebensendes der SSD zu nah ist, um Daten wiederherzustellen. Außerdem hat dieser Algorithmus keinen Einfluss auf das Schreiben einer großen Anzahl von großen Dateien und unterstützt nur RAID 5, was ein weiteres Risiko darstellt.

QSAL erkennt regelmäßig die Lebensdauer und Haltbarkeit von SSDs. Wenn die Lebensdauer der SSDs unter 50% liegt, wird die Overprovision-Größe dynamisch verteilt, um zu garantieren, dass jede SSD genügend Zeit hat, sich neu aufzubauen, bevor die Lebensdauer abläuft. Es erreicht die beste Balance zwischen Auslastung und Overprovisioning. QSAL wirkt sich nach der Aktivierung nicht auf die Speicherplatzauslastung aus und verursacht nur eine vernachlässigbare Leistungsminderung. Insgesamt wird der Schutz der Daten in All-Flash Arrays stark verbessert.

QSAL kann jederzeit gestartet werden und ist mit SSD Parity RAID kompatibel, das zuvor nicht konfiguriert wurde. Es wird empfohlen, QSAL zu aktivieren, bevor die Lebensdauer der SSD 50% erreicht. Wenn es zu spät aktiviert wird, kann es zu Problemen mit einer unzureichenden Wiederherstellungszeit kommen, die das Risiko einer Beschädigung des SSD-RAID mit sich bringt.

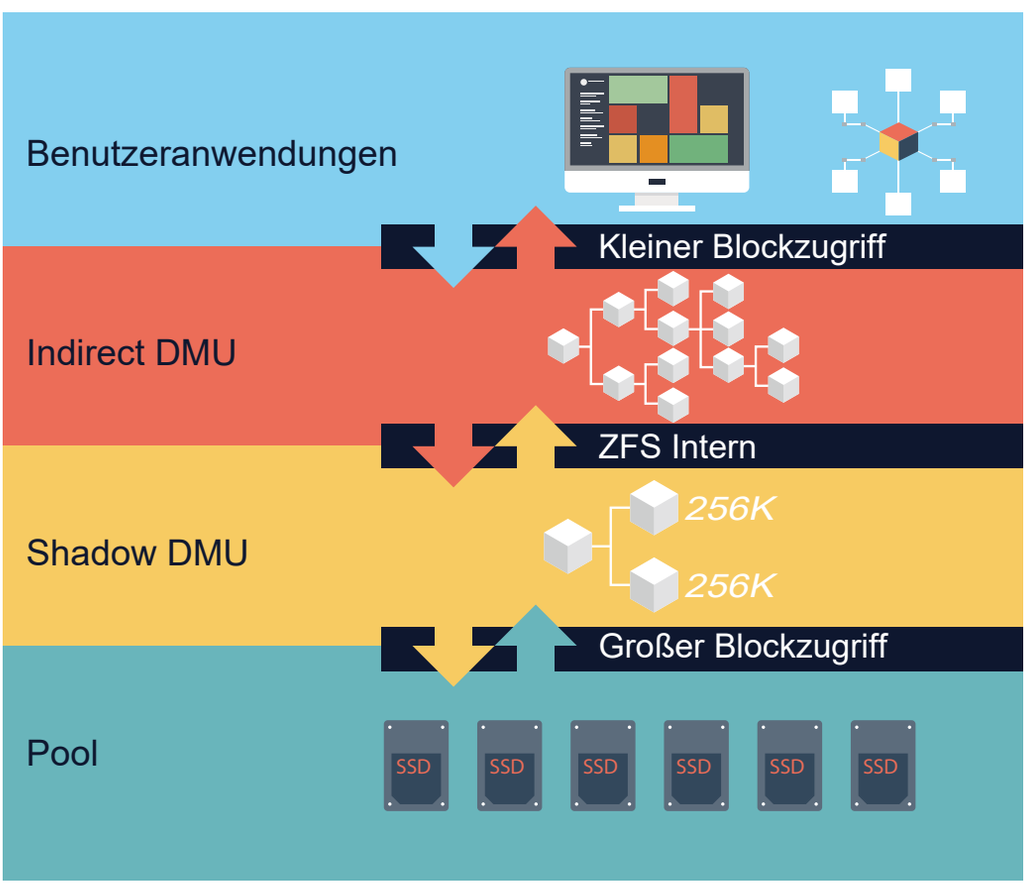


Patentnummern: US10705744B2 / US20200333959A1. Enthält auch mehrere vertrauliche Patente.

Schreib-Koaleszenz

Der Schreib-Koaleszenz-Algorithmus steuert die zufällige Schreibleistung für All-Flash Arrays

Die Schreibverstärkung ist ein unerwünschter Effekt von SSDs, der die Leistung und Ausdauer des Flash-Speichers beeinträchtigt. QuTS hero verfügt über den exklusiven Schreib-Koaleszenz-Algorithmus von QNAP, der für die Flash-Optimierung entwickelt wurde, indem er alle zufälligen Schreibvorgänge in sequenzielle Schreibvorgänge mit einem reduziertem I/O umwandelt. Er erhöht nicht nur effektiv die zufällige Schreibleistung in All-Flash-Umgebungen, sondern kann auch die SSD-Lebensdauer verbessern.



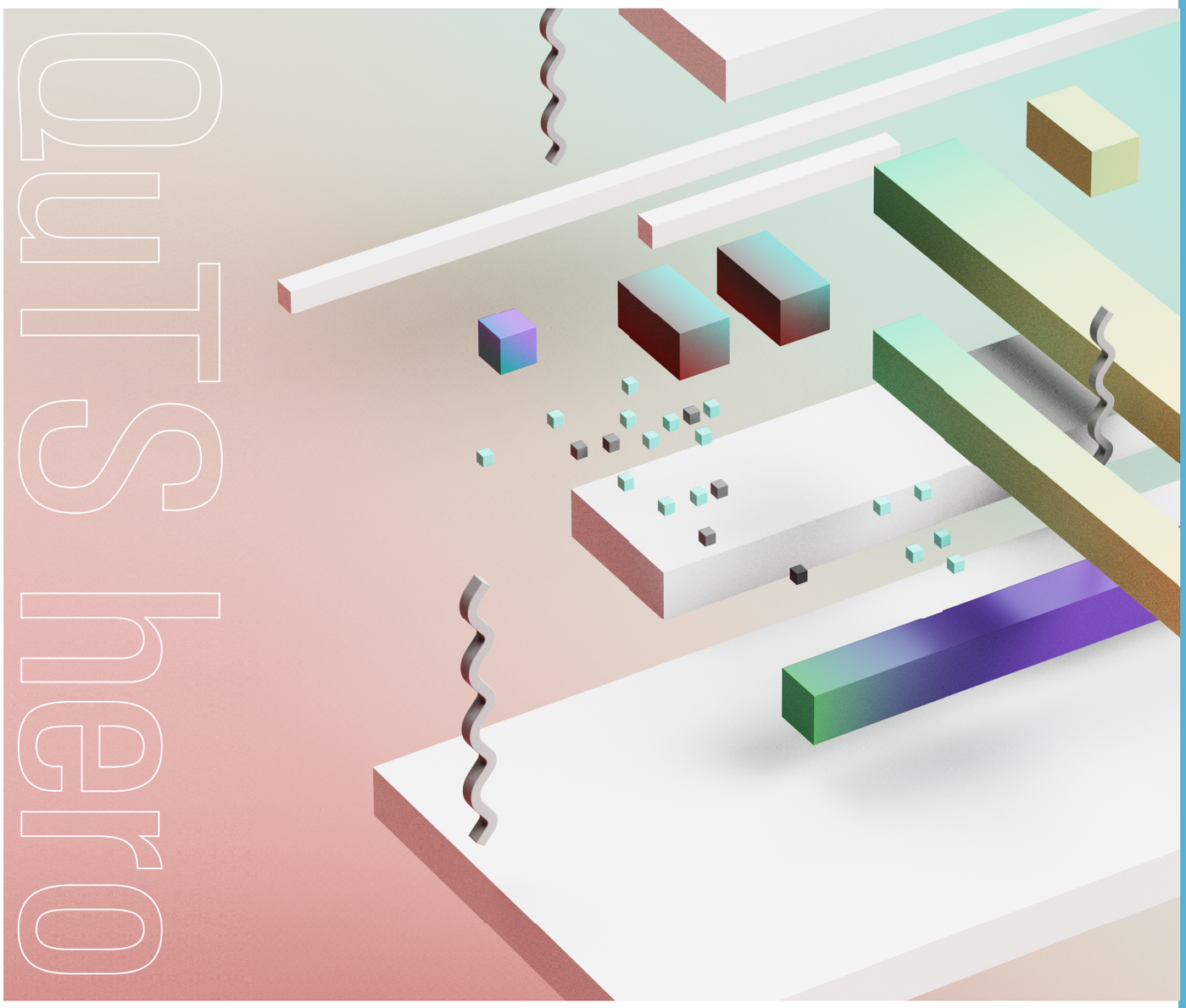
Empfohlene Lese-Cache-Konfiguration (L2ARC)

Der Arbeitsspeicher spielt eine entscheidende Rolle für die ZFS-Leistung - insbesondere bei der Highspeed-Datenübertragung, Datendeduplizierung, ARC und Caching. Es wird empfohlen, so viel Arbeitsspeicher wie möglich zu installieren, um den größten Nutzen aus der ZFS-Leistung sowie optimierten Arbeitslasten im Unternehmen zu erzielen.

Basierend auf dem in einem QuTS hero NAS installierten Arbeitsspeicher und dem Anteil, welcher der Nutzung des Lesecaches entspricht, empfehlen wir, die Größe des Lesecaches wie folgt einzustellen:

- **32GB RAM** unterstützt 1TB SSD-Größe für Lesecache
- **64GB RAM** unterstützt 2TB SSD-Größe für Lesecache
- **128GB RAM** unterstützt 4TB SSD-Größe für Lesecache
- **256GB RAM** unterstützt 8TB SSD-Größe für Lesecache

- **1TB RAM** unterstützt etwa 30TB SSD-Lesecache.
- **4TB RAM** unterstützt etwa 120TB SSD-Lesecache.


QuTS hero

Leistungsabhängigkeit zwischen Arbeitsspeichergröße und Speicherpools

Wichtig: In diesem Kapitel wird nur die grundsätzliche Abhängigkeit zwischen RAM und Speicherpools beschrieben. Wenn auf dem System andere dedizierte Dienste - wie z.B. Virtualization Station - ausgeführt werden, ist es empfehlenswert, die Nutzung zu trennen und den diesen virtuellen Maschinen zugewiesenen Arbeitsspeicher zu berücksichtigen.

NAS-Arbeitsspeicher - VM zugewiesener Arbeitsspeicher = ZFS genutzter Arbeitsspeicher

Dies ist das allgemeine Konzept für die Analyse der Arbeitsspeicher- und Speicherleistung und wie man die beste Arbeitsspeichergröße für Hochleistungsspeicher ermittelt. Anschließend wird beschrieben, wie Sie ausreichend Arbeitsspeicher auswählen, um die beste Leistung zu erzielen.

Der Arbeitsspeicher spielt in ZFS eine wichtige Rolle und hat einen großen Einfluss auf die Leistung von QuTS hero Speicherpools. Wenn ZFS mit Datenreduzierungsfunktionen (wie z.B. Verdichtung, schnelles Klonen und Deduplizierung) ausgeführt wird, benötigt das System einige Tabellen, um die Referenzanzahl dieser Datenblöcke aufzuzeichnen. Wenn die Datenmenge größer wird, muss ZFS daher ebenso viele Referenztabellen im Arbeitsspeicher halten und kann EA-Anforderungen effektiver verarbeiten.

Zweitens unterstützt ZFS die Transaktionsgruppen Funktion zum Schutz der Daten. Jede Dateiänderung (z.B. ein Schreibvorgang) ist mit einer bestimmten Transaktionsgruppe (TXG) verbunden. In regelmäßigen Abständen wird jede TXG heruntergefahren und der Pool eine Synchronisation für diese Gruppe durchführen. Ein TXG kann auch heruntergefahren werden, wenn die ARC anzeigt, dass zu viel "schmutziger" Speicher im Cache vorhanden ist. Beim Schließen eines TXG wird sofort ein anderer geöffnet und die Dateiänderungen werden dann mit dem neuen aktiven TXG verknüpft.

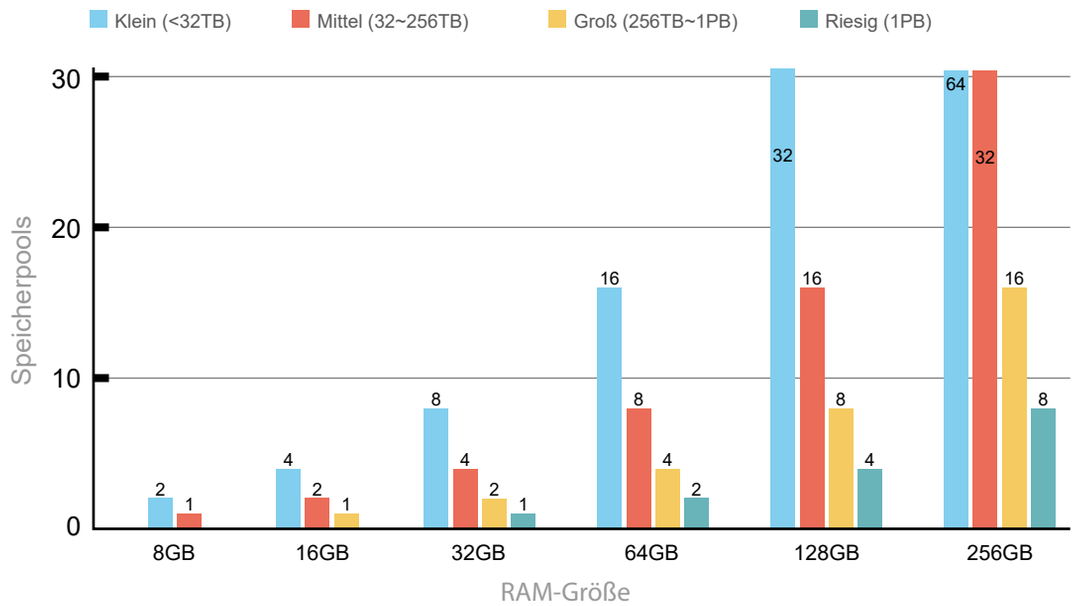
Wenn ein aktiver TXG heruntergefahren wird, während ein vorheriger TXG noch Daten mit dem Speicher synchronisiert, werden Anwendungen gedrosselt, bis die laufende Synchronisation abgeschlossen ist. In dieser Situation, in der ein TXG synchronisiert wird, während TXG + 1 aufgrund von Speicherbeschränkungen geschlossen ist und auf die Synchronisation wartet, werden Anwendungen gedrosselt, die darauf warten, auf TXG + 2 zu schreiben. Wir brauchen eine dauerhafte Sättigung des Speichers oder eine Speicherbeschränkung, um Anwendungen zu drosseln.

Bei einer Synchronisation des Speicherpools werden alle Datenblöcke der Ebene 0 auf die Festplatte übertragen. Wenn dies abgeschlossen ist, werden alle indirekten Blöcke der Ebene 1 (usw.) gesendet, bis jeder Block, der den neuen Zustand des Dateisystems repräsentiert, übertragen wurde. Zu diesem Zeitpunkt wird der Uberblock aktualisiert, um auf den neuen konsistenten Zustand des Speicherpools zu verweisen.

ZFS muss also nach Abschluss jeder Transaktion den Uberblock aktualisieren, was einen großen Overhead bedeutet. Wenn das System so viel Arbeitsspeicher wie möglich zur Verfügung stellen kann, können die Transaktionsgruppen größer sein, und der Overhead für die Aktualisierung des Uberblocks wird relativ kleiner. Somit kann mehr Arbeitsspeicher ZFS effektiv helfen, EA-Schreibanforderungen zu verbessern.

Sie können sich an unseren untenstehenden Empfehlungen orientieren und ausreichend Arbeitsspeicher für verschiedene Speicherkonfigurationen wählen.

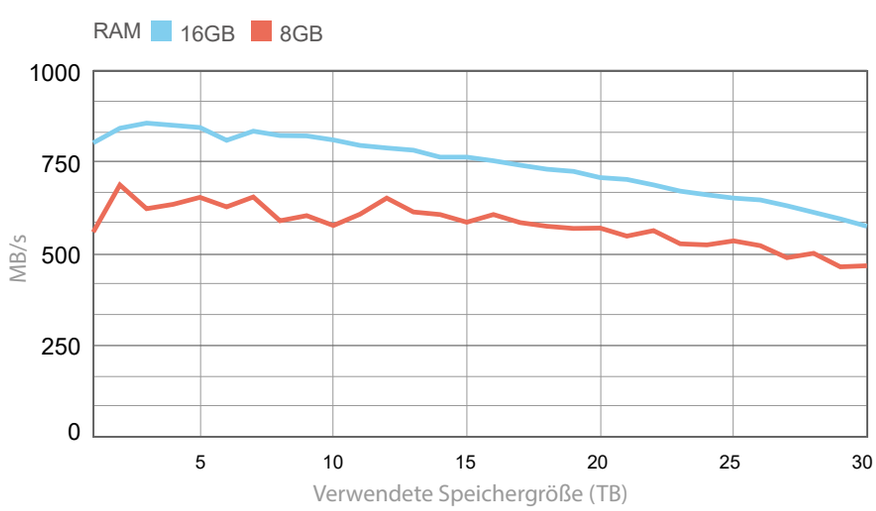
Die besten Praktiken für die Leistung zwischen RAM-Größe und Pool-Konfiguration



Wenn Ihr ZFS NAS beispielsweise über 16GB Arbeitsspeicher verfügt, wäre die empfohlene Konfiguration des Speicherpools 1 großer Pool (256 TB~1 PB), oder 2 mittelgroße Pools, oder 4 kleine Pools. Mit dieser Art von Pool-Konfiguration würde Ihr System eine gute Leistung erzielen. Wenn Sie Pools erstellen, die diese Empfehlungen überschreiten, wird Ihr NAS zwar immer noch zuverlässig arbeiten, aber mit potenziell verringerter Leistung. Hier verwenden wir unser ZFS-Einsteigermodell TS-973AX als Referenz, um die unterschiedliche Leistung zwischen 8GB (Mindestanforderung) und 16GB RAM zu zeigen.

(Wir haben bereits einen SSD-Pool und einen HDD-Pool um 35TB angelegt. Es wird der Leistungsunterschied des HDD-Pools angezeigt).

TS-h973AX-16GB im Vergleich zu 8GB - EA-Leistung (5xHDD-R0)



QuTS hero ZFS

Whitepaper

QNAP SYSTEMS, INC.

TEL : +886-2-2641-2000 FAX: +886-2-2641-0555 Email: qnapsales@qnap.com

Address : 3F, No.22, Zhongxing Rd., Xizhi Dist., New Taipei City, 221, Taiwan

QNAP may make changes to specification and product descriptions at any time, without notice.
Copyright © 2021 QNAP Systems, Inc. All rights reserved.
QNAP® and other names of QNAP Products are proprietary marks or registered trademarks of QNAP Systems, Inc.
Other products and company names mentioned herein are trademarks of their respective holders.

Netherlands (Warehouse Services)
Email: nlsales@qnap.com
TEL: +31(0)107600830

China
Email: cnsales@qnap.com
TEL: +86-400-028-0079

Thailand
Email: thsales@qnap.com
TEL: +66-2-5415988

Japan
Email: jpsales@qnap.com
FAX: 03-6435-9686

US
Email: usasales@qnap.com
TEL: +1-909-595-2782

India
Email: indiasales@qnap.com

Germany
Email: desales@qnap.com

France
Email: Frsales@qnap.com