QNAP

QuTS hero
ZFS

Documento técnico

INDEX

Página	Título
01	**Introducción**
02	**Introducción y configuración de ZFS**
02	Integridad de los datos y autorrecuperación
02	Tamaño de bloque recomendado para carpeta compartida /
03	LUN E/S sincrónica para ZIL (Automático / Todos / Ninguno)
04	Sobreaprovisionamiento de conjunto
06	**Reducción de datos (Compresión en línea / Deduplicación en línea)**
07	Compresión en línea
08	Deduplicación en línea
09	**SnapSync**
	Asequible solución de replicación para copia de seguridad, protección de datos y recuperación de desastres.
09	Introducción
09	Cómo funciona SnapSync
11	Qué hacer en caso de que se produzca un desastre
12	Prácticas recomendadas para la configuración de SnapSync en tiempo real
14	Beneficios
15	**QSAL para matrices all-flash y RAID SSD (tecnología patentada por QNAP)**
16	**Coalescencia de escritura**
	El algoritmo de coalescencia de escritura impulsa el rendimiento de escritura aleatorio para matrices all-flash
17	**Configuración de caché de lectura recomendada (L2ARC)**
19	**Dependencia de rendimiento entre el tamaño de la memoria y los conjuntos de almacenamiento**

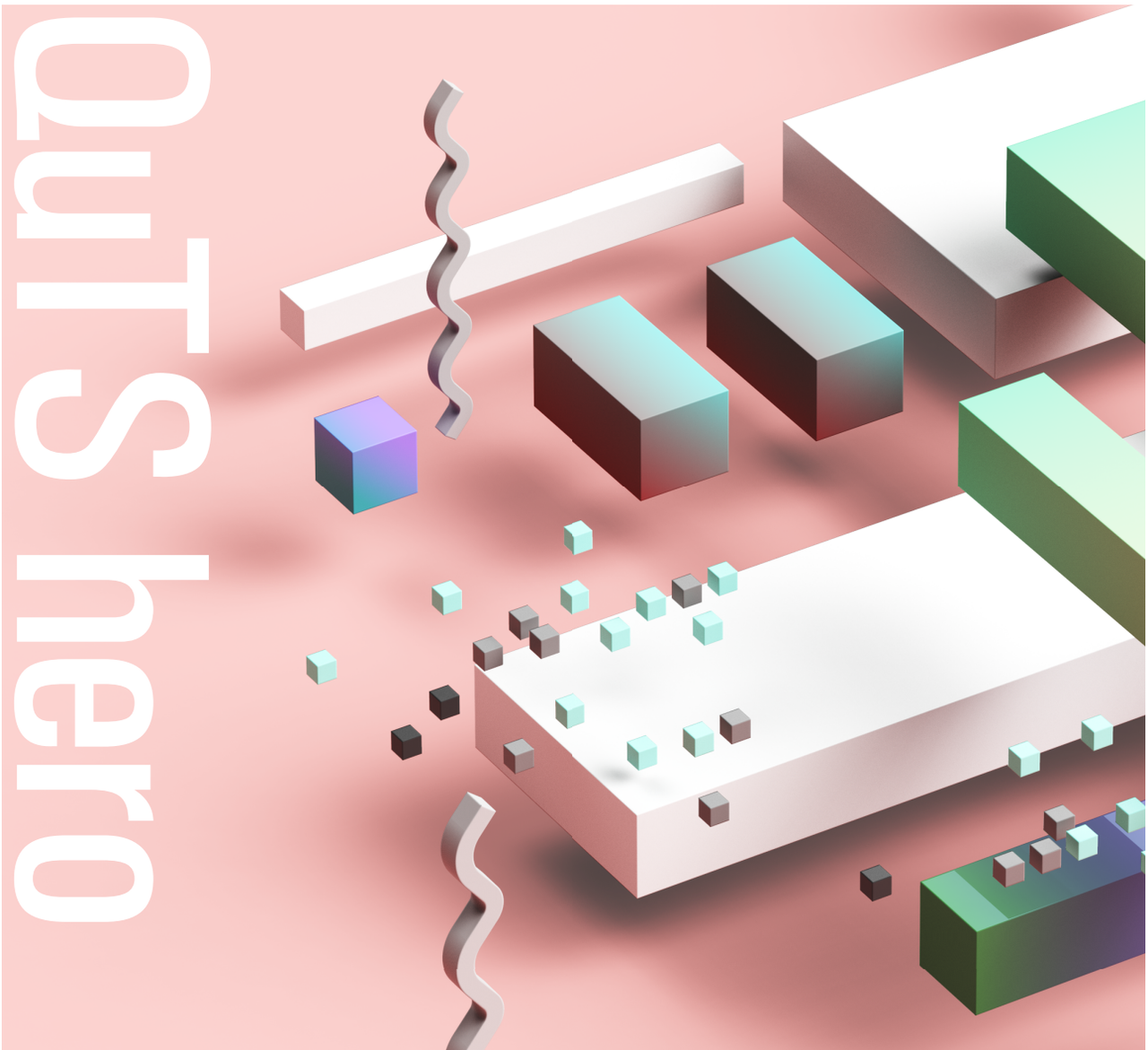
QuTS hero
ZFS

documento técnico

Introducción

En el mundo actual en el que el uso de las TI es cada vez mayor, las necesidades cambiantes pueden resultar clave para el éxito o el fracaso de las organizaciones empresariales y de las operaciones de centro de datos. En este entorno, los presupuestos están muy controlados y cualquier posible periodo de inactividad (ya sean minutos o incluso segundos) puede resultar desastroso. Por otro lado, cualquier posible aumento de la eficiencia y la productividad es de fundamental importancia para lograr el éxito.

El sistema operativo QuTS hero adopta ZFS, un sistema de archivos que prioriza la integridad de los datos para responder a la necesidad de disponer de un sistema de copia de seguridad de los datos de alto rendimiento y gran estabilidad. QuTS hero integra aplicaciones de protección de datos, reducción de datos y virtualización de alto rendimiento y admite entornos de cloud y funciones de administración de QoS (Calidad del servicio) para ofrecer un almacenamiento más asequible para empresas pequeñas, centros de datos de nivel empresarial y soluciones de operaciones de trabajo virtualizadas.

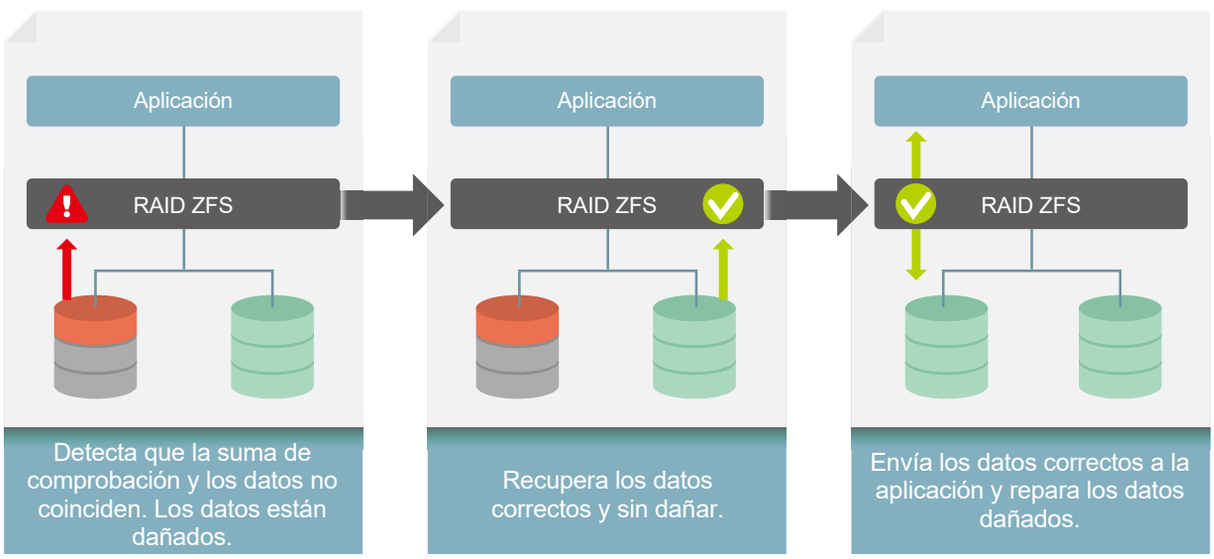


Introducción y configuración de ZFS

Integridad de los datos y autorrecuperación

La integridad y la fiabilidad de los datos son clave para garantizar el funcionamiento de las aplicaciones y las bases de datos. El sistema ZFS tiene la capacidad de evitar la corrupción silenciosa de los datos, una condición potencialmente desastrosa y difícil de detectar provocada por problemas de hardware, cables defectuosos o errores de metadatos. El sistema ZFS tiene la capacidad de recuperar automáticamente estos datos, incluyendo la comprobación de todos los bloques de datos y la reparación automática de bloques incorrectos.

Es muy aconsejable realizar regularmente una "Limpieza de conjunto" para garantizar la integridad de los datos almacenados y repara posibles problemas antes de que surjan. Puede configurar la Limpieza de conjunto para que se ejecute según una programación (por ejemplo, semanalmente o mensualmente).



Evita la corrupción silenciosa de los datos en un sistema en ejecución.

Tamaño de bloque recomendado para carpeta compartida o LUN:

128K - Edición de vídeo / Archivos de tamaño grande / Copia de seguridad
64K - Genérico / Hyper-V
32K - VMware
8k - OLTP
4K - VDI / Base de datos / Archivos de tamaño pequeño

E/S sincrónica para ZIL (Standard / Always / None):

El ZIL (Registro de intención de ZFS) es el mecanismo de registro de escritura a modo de diario que controla los cambios realizados en un archivo que no se han confirmado en el sistema de archivos. El ZIL registra las intenciones de dichos cambios en un registro, conocido como diario. En caso de que se produzca una avería del sistema o un fallo de alimentación, QuTS hero comprobará los registros del diario y aplicará los cambios programados, garantizando que el sistema de archivos se restaure rápidamente con menos posibilidades de corrupción de datos.

Al crear una carpeta compartida o un LUN, el modo de E/S del registro de escritura de ZIL debe estar configurado para priorizar la integridad de los datos o el rendimiento.

Standard (default): Las transacciones de E/S son sincrónicas o asincrónicas dependiendo de la aplicación y el tipo de solicitud de E/S.

Always: Todas las transacciones de E/S se consideran sincrónicas, lo que mejora la integridad de los datos pero reduce el rendimiento.

None: Todas las transacciones de E/S se consideran asincrónicas. Esta opción ofrece el mayor rendimiento pero tiene más riesgo de pérdida de información en caso de producirse una interrupción de la alimentación eléctrica.

Recuerde que se recomienda configurar un ZIL dedicado como un SLOG en un dispositivo de almacenamiento de alta velocidad dedicado (como una SSD NVMe). Para el ZIL, se recomienda usar una latencia ultrabaja y unidades de alta resistencia (como unidades MLC). Para la caché de lectura, debería priorizarse el rendimiento. Se recomienda dividir el ZIL y la caché de lectura para evitar un fallo prematuro de todo el grupo de SSD debido a una alta frecuencia de escritura de ZIL.

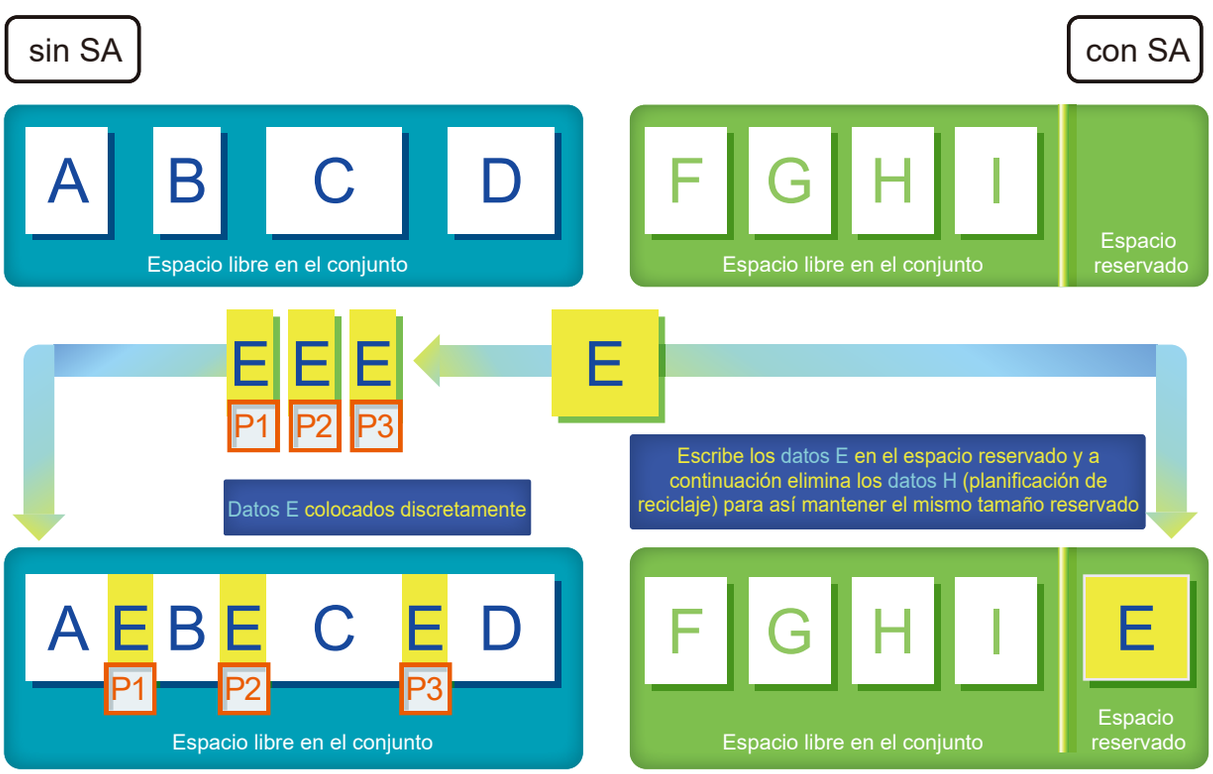
Sobreaprovisionamiento de conjunto

El rendimiento del sistema ZFS puede verse afectado cuando el conjunto de almacenamiento esté casi lleno o fragmentado. Esto se debe al comportamiento de Copia en escritura: el sistema ZFS siempre escribe los nuevos datos en un bloque diferente, actualiza los metadatos para que apunten a la nueva ubicación y luego libera el bloque antiguo en lugar de sobrescribir directamente los datos antiguos. Tras muchas rondas de escritura, la distribución de los datos en el conjunto de almacenamiento se fragmentará, lo que hará que resulte más difícil y tedioso escribir nuevos datos.

QuTS hero admite el Sobreaprovisionamiento de conjunto, que le permite especificar y reservar un porcentaje del espacio del conjunto de almacenamiento para garantizar un rendimiento de escritura aleatoria constante cuando el conjunto esté casi lleno. En la siguiente figura se ilustran los datos que se escriben en el conjunto de almacenamiento con o sin sobreaprovisionamiento de conjunto. Sin el sobreaprovisionamiento de conjunto, cuando el conjunto de almacenamiento esté casi lleno y fragmentado, los nuevos bloques de datos escritos se separarán un poco y el rendimiento se verá afectado. Con el sobreaprovisionamiento de conjunto, el sistema ZFS puede encontrar un bloque grande continuo para almacenar nuevos datos.

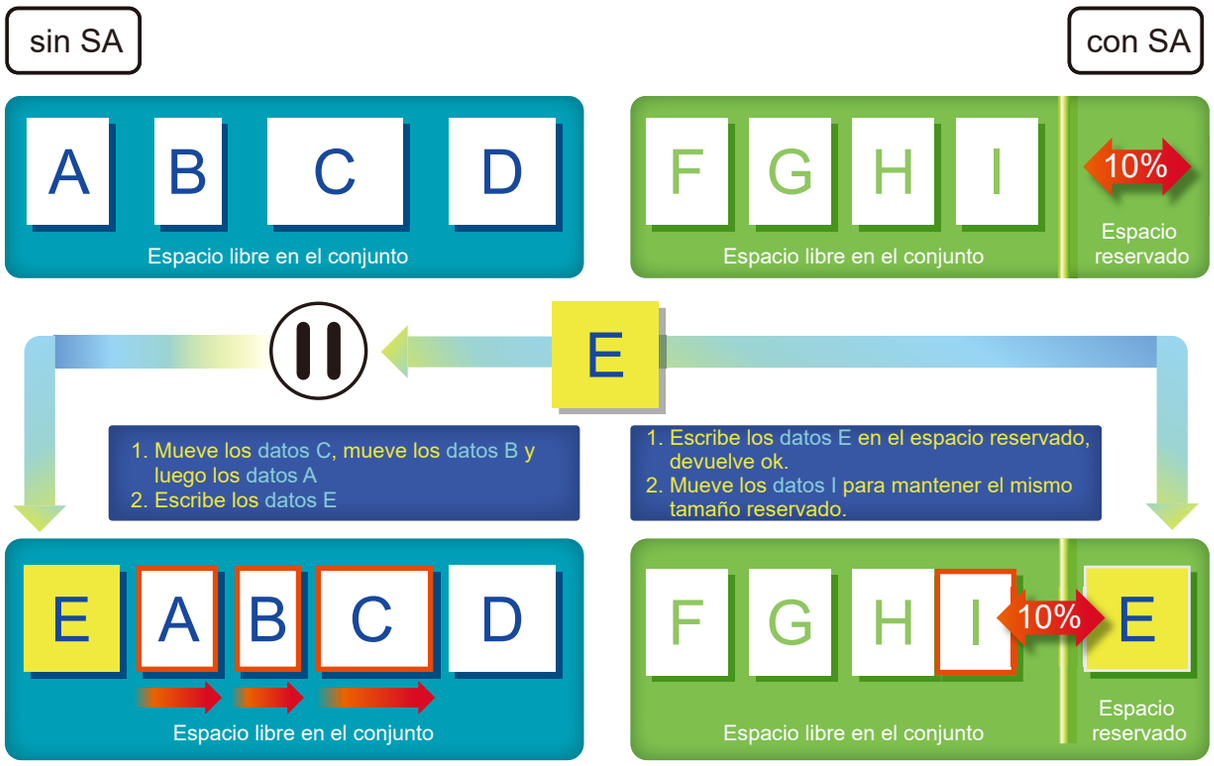
Sobreaprovisionamiento

Mejora el rendimiento de los conjuntos fragmentados (por ejemplo, cuando se escriba un bloque en una **unidad de disco duro**)



Sobreaprovisionamiento

Mejora el rendimiento de los conjuntos fragmentados (por ejemplo, al escribir un bloque en una **SSD**)



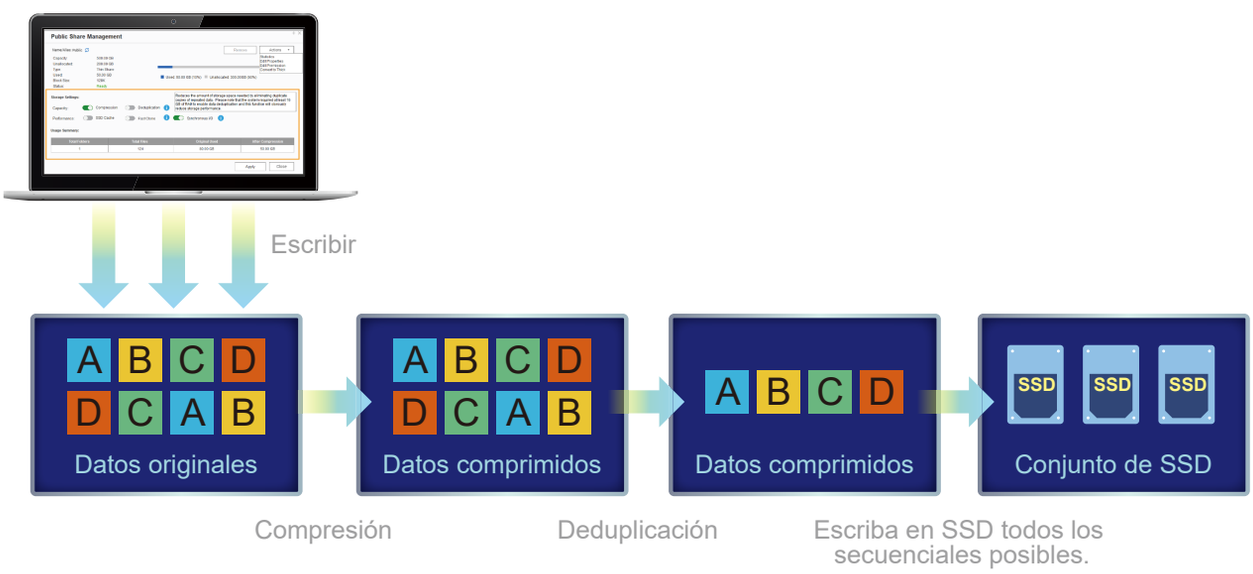
Al habilitar el Sobreaprovisionamiento de conjunto, se garantiza el rendimiento de escritura aleatoria a costa de la capacidad. Cuanto más espacio se reserve, mayor será el rendimiento. La clave para determinar si tal intercambio es necesario está en considerar la relación de lectura/escritura y los escenarios de aplicación. Cuantas más escrituras haya, mayor será el aumento de rendimiento que ofrecerá el sobreaprovisionamiento de conjunto.

Normalmente se recomienda establecer el valor del sobreaprovisionamiento de conjunto en 10-20 %. Para aplicaciones con un uso intensivo de la escritura y que exigen un alto rendimiento (como registros SQL, copia de seguridad, vigilancia e imágenes médicas), se recomienda establecer un valor más alto. Para aplicaciones con un uso intensivo de la lectura (como servidores web, OLAP y transmisión multimedia), es posible que el aumento de rendimiento que consiga el sobreaprovisionamiento de conjunto no compense la pérdida de espacio de almacenamiento y se pueda usar un valor más bajo.

Reducción de datos (Compresión en línea / Deduplicación en línea)

La Deduplicación de datos en línea permite eliminar los datos repetidos antes de que se escriban, lo que ahorra considerablemente en espacio de almacenamiento. La Compresión en línea también es compatible para ahorrar espacio al guardar archivos grandes. Estas funciones, junto con la tecnología de Compactación en línea, ofrecen la capacidad de optimizar considerablemente el uso del almacenamiento. Cuando se usan en configuraciones all-flash que almacenen datos muy repetitivos o grandes cantidades de archivos pequeños, la capacidad de reducir las necesidades de almacenamiento resulta considerablemente beneficiosa. No solo mejora significativamente el rendimiento de escritura de las configuraciones all-flash, sino que también amplía el ciclo de vida de las SSD y maximiza la rentabilidad general de la implementación.

La tasa de compresión puede alcanzar el 90 % para entornos virtualizados.



Compresión en línea

QuTS hero admite la compresión en línea, lo que reduce los tamaños de bloque y archivo codificándolos con menos bits que la representación original cuando los datos se escriben en una carpeta compartida o en un LUN. Al habilitar la compresión, también se aumenta el rendimiento de lectura y escritura, ya que se tienen que leer o escribir menos bloques.

En QuTS hero, la compresión es sin pérdidas para garantizar la coherencia de los datos. El sistema ZFS admite la compresión LZ4 de forma predeterminada, ya que normalmente es la más rápida.

Esta compresión no requiere de unos recursos de CPU excesivos y es aconsejable usarla en carpetas compartidas y LUN. La compresión no debería usarse para archivos que ya se han comprimido (como archivos MP3, vídeos H.264 y archivos zip), ya que el almacenamiento que se ahorra volviéndolos a comprimir es inapreciable.

El LZ4 es un algoritmo de compresión de datos sin pérdida que ofrece unas velocidades de compresión y descompresión extremadamente rápidas. Es la mejor opción para baja latencia y alto IO cuando las velocidades de transmisión sean superiores a 100 MB/s. Resulta especialmente adecuado para la compresión en línea.

Velocidad de compresión	Velocidad de descompresión
740 MB/S	4530 MB/S

*Probado por Izbench

Deduplicación en línea

QuTS hero utiliza una deduplicación en línea basada en el objetivo y a nivel de bloques. "Basada en el objetivo" significa que el sitio de origen transfiere primero los datos al sitio de destino y luego se produce el proceso de duplicación en el sitio de destino. "A nivel de bloques" significa que puede detectar y eliminar pequeñas partes redundantes de los datos dentro de un archivo o entre diferentes archivos. En línea significa que el proceso de deduplicación se produce justo mientras los datos se transfieren al dispositivo de almacenamiento pero antes de que se escriban. Las ventajas de esta combinación (basada en el objetivo, a nivel de bloques y en línea) es que se reducen los recursos de CPU necesarios del sitio del cliente, se consiguen unas tasas de deduplicación más altas y disminuyen las necesidades de acceso a la unidad en comparación con otros mecanismos de deduplicación.

El proceso de deduplicación incluye tres pasos: dividir los datos en bloques, calcular la huella digital y buscar en la tabla de deduplicación. En primer lugar, el sistema ZFS divide los archivos en pequeños bloques de datos. Cuando se escribe un bloque de datos, el sistema utilizará una suma de comprobación criptográfica para calcular la huella digital de cada bloque de datos, que se usa para identificar si hay algún bloque idéntico en los datos existentes. ZFS mantiene una tabla de asignación (llamada tabla de deduplicación) que contiene todas las huellas digitales, ubicaciones y total de referencias. Si se escribe un bloque de datos en el NAS con la misma huella digital que un bloque de datos existente, en lugar de escribirlo en una nueva ubicación, ZFS consultará la tabla de deduplicación, asignará la misma ubicación a este bloque de datos y aumentará el total de referencias en uno.

La deduplicación en línea reduce la necesidad de almacenamiento adicional, pero requiere de más recursos informáticos al escribir datos, lo que puede afectar al rendimiento del sistema. Este impacto en el rendimiento aumentará a medida que aumente el tamaño de la tabla de deduplicación, así como la dificultad de buscar en la tabla. Para evitar la reducción del rendimiento durante la deduplicación en línea, QNAP utiliza el algoritmo SmartDDT (Smart DeDuplication Table) para limitar el crecimiento de la tabla de deduplicación monitorizando los cambios de rendimiento o el tamaño de la tabla de deduplicación. Si se cumplen algunos criterios, QuTS hero detendrá el servicio de deduplicación. Si esto ocurre, puede ser que los datos no sean muy repetitivos y que el escenario de aplicación no sea adecuado para la deduplicación en línea o bien que el NAS necesite más memoria.

En el caso de entornos que necesitan la deduplicación, recomendamos instalar toda la memoria posible: como mínimo, 16 GB de memoria, pero se recomiendan 64 GB.

SnapSync:

la solución de replicación rentable para copia de seguridad, protección de los datos y recuperación de desastres.

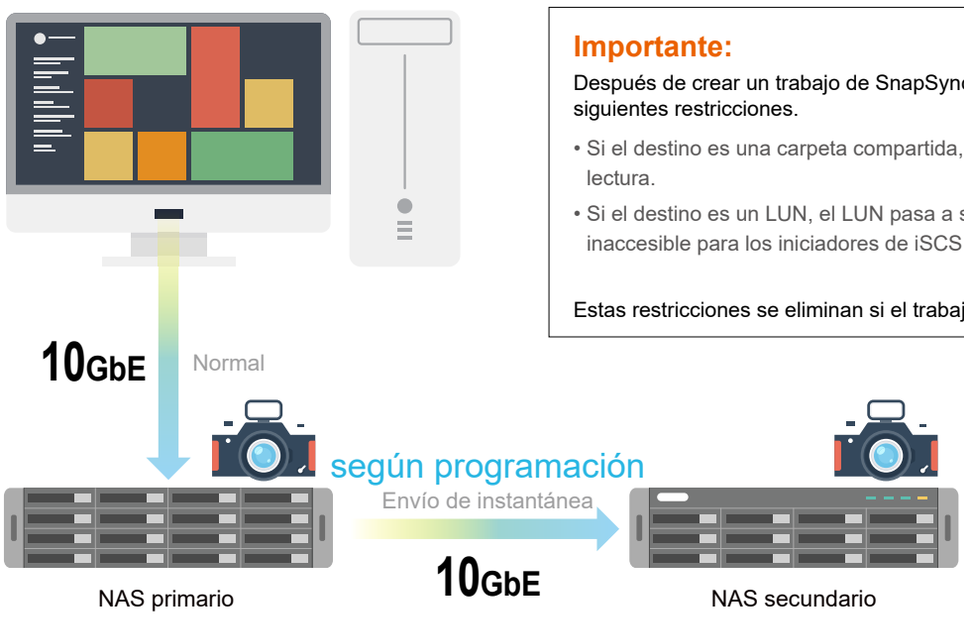
Introducción:

SnapSync es una solución de replicación unificada, asequible y fácil de usar integrada en QuTS hero para tareas de copia de seguridad, protección de datos y recuperación de desastres.

SnapSync se configura a través de la relación en espejo entre carpetas compartidas/LUN en los sistemas de almacenamiento principal y secundario. Sincroniza periódicamente la réplica para mantenerla actualizada con los cambios que se hayan escrito en el sistema principal. Estos datos replicados se crean en un sistema de almacenamiento secundario. En caso de que se produzca un desastre en el sitio principal, podrá conmutar a los datos del sistema secundario. Con SnapSync, puede reducir el coste total de las soluciones de recuperación de desastres, lo que facilita la justificación de las inversiones utilizando su sitio de recuperación de desastres de forma activa.

Cómo funciona SnapSync:

La función SnapSync de nivel de bloques permite hacer una copia de seguridad de los bloques variables a través de instantáneas, lo que resulta ligero y ahorra ancho de banda, y los archivos clave pueden copiarse eficazmente en ubicaciones remotas. Con la replicación incremental, nunca se envían dos veces los mismos datos. Los tiempos de recuperación de SnapSync también son más rápidos que otros métodos, ya que la restauración solo necesita una copia de seguridad completa y la última copia de seguridad diferencial.



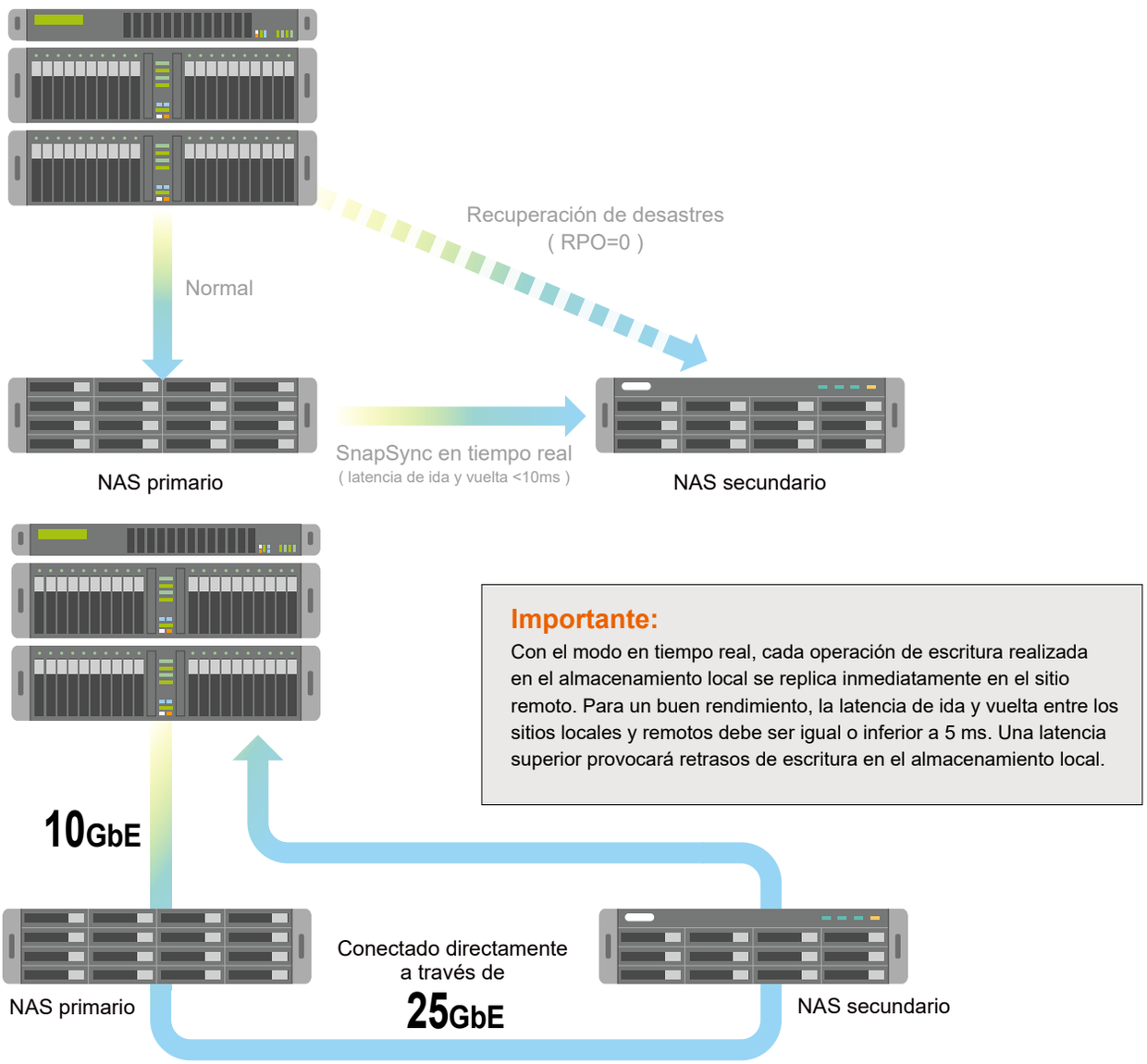
Importante:

Después de crear un trabajo de SnapSync, son de aplicación las siguientes restricciones.

- Si el destino es una carpeta compartida, la carpeta pasa a ser de solo lectura.
- Si el destino es un LUN, el LUN pasa a ser de solo lectura e inaccesible para los iniciadores de iSCSI.

Estas restricciones se eliminan si el trabajo se borra.

La función SnapSync en tiempo real más avanzada permite que el NAS principal y el NAS secundario se mantengan sincronizados en todo momento.



Importante:
Con el modo en tiempo real, cada operación de escritura realizada en el almacenamiento local se replica inmediatamente en el sitio remoto. Para un buen rendimiento, la latencia de ida y vuelta entre los sitios locales y remotos debe ser igual o inferior a 5 ms. Una latencia superior provocará retrasos de escritura en el almacenamiento local.

La función de SnapSync en tiempo real garantiza que los datos del NAS principal y el NAS secundario se escribirán por completo antes de responder al servidor host. Esto garantiza un RPO=0 (Objetivo de punto de recuperación) y también significa que los datos del NAS secundario serán coherentes con los del NAS principal. En caso de que de produzca un desastre, el servidor host puede cambiar inmediatamente el objetivo de montaje al NAS secundario cuando el servidor falle.

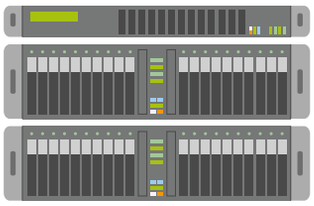
Para crear un sistema de almacenamiento de gran tamaño, se recomienda configurar el NAS secundario con SnapSync y luego usar Snapshot Replica y HBS3 para realizar regularmente copias de seguridad de varias versiones del NAS secundario a otro NAS de copia de seguridad.

También admite SnapSync de NAS QES (QES 2.1.1 v11 o posterior) para NAS QuTS hero.

Qué hacer en caso de que se produzca un desastre

Si se produce un desastre...

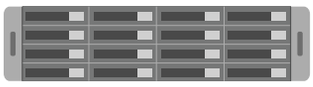
Servidor de producción



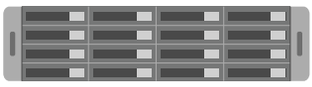
NAS (almacenamiento compartido)

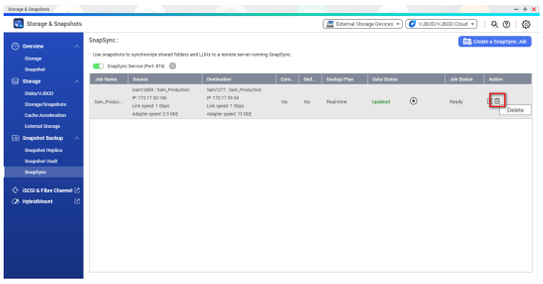
Primero, elimine la tarea de SnapSync 1

NAS primario
192.168.100.100



NAS secundario
192.168.100.200





Los permisos exclusivos de la carpeta se eliminarán al borrar la tarea de SnapSync.

A Volver a montar la IP del NAS secundario

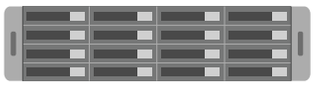
o

B Cambiar la IP del NAS secundario a la IP del NAS primario original

Servidor de producción

Montar objetivo = 192.168.100.200

NAS secundario



192.168.100.200 => 192.168.100.100

Prácticas recomendadas para la configuración de SnapSync en tiempo real

En el modo en tiempo real, cada operación de escritura realizada en el almacenamiento local se replica inmediatamente en el sitio remoto. Para un buen rendimiento, la latencia de ida y vuelta entre los sitios locales y remotos debe ser igual o inferior a 5 ms. Una latencia superior provocará retrasos de escritura en el almacenamiento local. Por tanto, a la hora de configurar, debe prestar atención a lo siguiente:

- Se recomienda que la latencia del entorno de red sea inferior a 10 milisegundos siempre que sea posible (lo ideal sería menos de 5 milisegundos)
- El rendimiento de E/S del NAS secundario debería ser igual que el del NAS primario.
- Recomendamos conectar directamente el NAS primario y el NAS secundario utilizando la conectividad 25GbE.

Prácticas recomendadas para configuración

La referencia de rendimiento de SnapSync en tiempo real de los Laboratorios de QNAP:

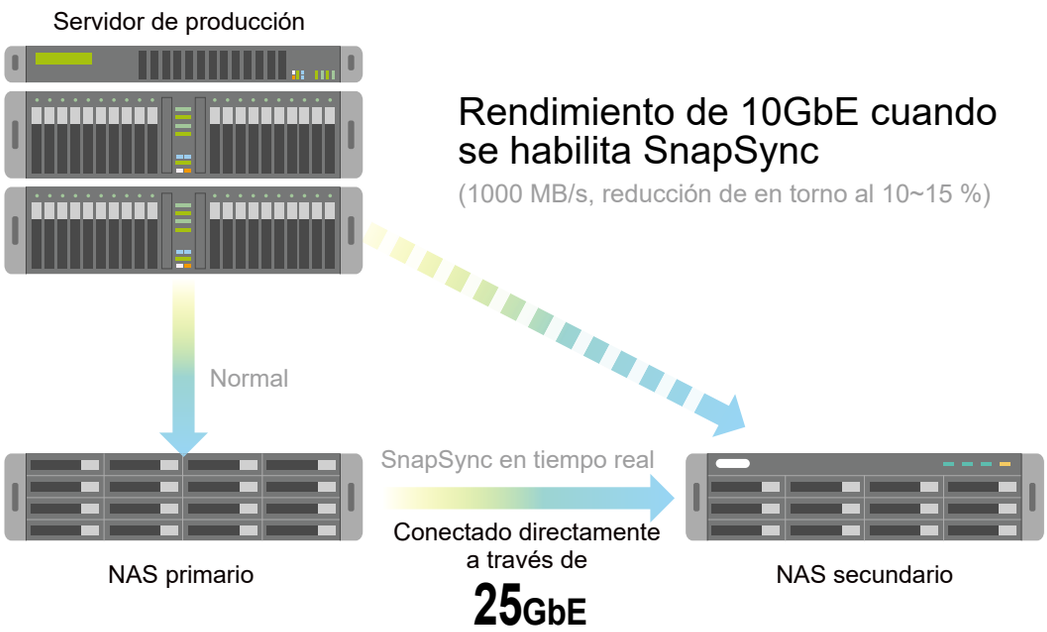
- Probado con el protocolo iSCSI: Cuando el usuario se conecte al NAS QuTS hero con la conectividad 10GbE, la transferencia de archivos de 10GbE puede alcanzar los 1127 MB/s antes de habilitar SnapSync y la eficiencia será de 1005 MB/s al utilizar SnapSync en tiempo real. Tiene una atenuación del 10~15 %.

- Probado con el protocolo SMB: Cuando el usuario se conecte al NAS QuTS hero con la conectividad 10GbE, la transferencia de archivos de 10GbE puede alcanzar los 1025 MB/s antes de habilitar SnapSync y la eficiencia será de 930 MB/s al utilizar SnapSync en tiempo real. Tiene una atenuación del 10~15 %.

- Probado con el protocolo NFS: Cuando el usuario se conecte al NAS QuTS hero con la conectividad 10GbE, la transferencia de archivos de 10GbE puede alcanzar los 961 MB/s antes de habilitar SnapSync y la eficiencia será de 897 MB/s al utilizar SnapSync en tiempo real. Tiene una atenuación del 7~15 %.

Antes de la protección de SnapSync

Tras habilitar SnapSync en tiempo real



Entorno de prueba:

- Modo IO: sincronización = estándar y sincronización = ninguno
- Tamaño del bloque = 128K
- Jumbo Frame (MTU)= 9000
- Cuando la prueba anterior se conecta con 10GbE en el mismo entorno, debido a la influencia de la conexión y el retardo de la red, la atenuación después de habilitar SnapSync en tiempo real volverá a aumentar hasta el 20 % aproximadamente.

Beneficios:

Reducción del tiempo de inactividad y protección frente a la pérdida de datos

SnapSync ofrece replicación programada y en tiempo real. Esto permite establecer Objetivos de punto de recuperación (RPO) de segundos. Podrá cumplir con los objetivos de recuperación más estrictos incluso con enormes cargas de trabajo con un uso intensivo de la escritura.

Reducción del uso del ancho de banda de la red:

SnapSync aprovecha las eficiencias de almacenamiento enviando bloques incrementales a través de la red con compresión/deduplicación para acelerar la transferencia de datos y reducir el uso del ancho de banda. Con SnapSync, puede aprovechar una transmisión de datos de replicación ligera para crear un único repositorio que mantenga la replicación activa.

Distribución sencilla de grandes cantidades de datos:

En ocasiones, es necesario enviar grandes cantidades de datos: para migrar matrices de salas de servidores, para consolidar oficinas remotas o para configurar una nueva sucursal. SnapSync ofrece un método rápido, eficiente y flexible de mover datos. Si su empresa se encuentra geográficamente dispersa y todas las ubicaciones necesitan acceder al mismo conjunto de datos (como vídeos de formación o kits de ventas), puede utilizar SnapSync para distribuir rápidamente los mismos datos a todas las ubicaciones.

CDM (Administración de copia de datos) y análisis de datos:

La realización de análisis extensos puede resultar crítica para su empresa, pero también puede afectar al rendimiento de los entornos de producción. Con SnapSync y la función de Instantáneas, puede aprovechar los datos replicados para ejecutar los análisis más complejos en copias de datos secundarias.

Necesidades de conservación de datos, cumplimiento y múltiples versiones:

Muchos casos de uso requieren de largos periodos de conservación de los datos. Si combina SnapSync con la función de Instantáneas, podrá cumplir los requisitos de cumplimiento para proteger sus datos. La copia de seguridad de instantáneas puede ayudarle a mitigar el impacto del ransomware y, gracias a SnapSync en tiempo real, se puede aumentar la disponibilidad y la recuperación de desastres.

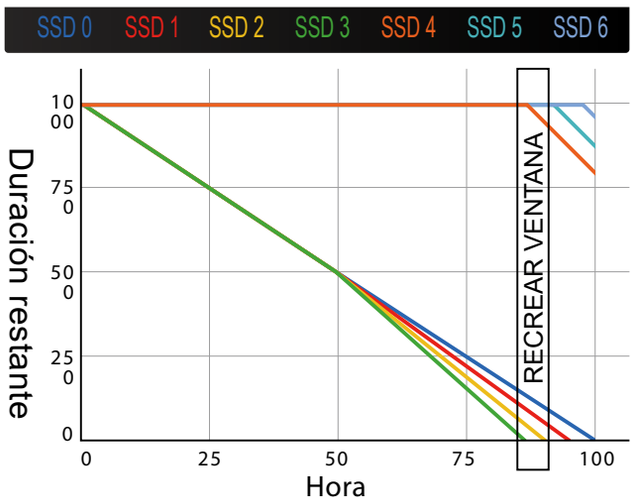
QSAL para matrices all-flash y RAID SSD (tecnología patentada por QNAP)

El algoritmo QSAL (QNAP SSD Anti-wear Leveling) se habilita automáticamente y evita que varias SSD fallen al mismo tiempo.

Se habilitarán por defecto los RAID SSD 5 / 6 / 50 / 60 / TP (Triple paridad) según corresponda.

M0 OP = 0

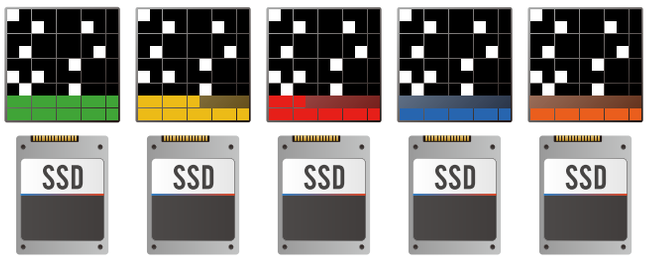
Mi OP = i x W / Mi-1, if (Mi-1 - Mi) > W



Los RAID SSD usados por los productos de la competencia (principalmente aquellos que solo aumentan la carga de la SSD más antigua) tienen un riesgo especialmente alto de pérdida de datos más adelante, ya que la mitad del ciclo de vida de la SSD es demasiado pronto para reconstruir datos. Además, este algoritmo no tiene ningún efecto en la escritura de un gran número de archivos grandes y solo admite RAID 5, lo cual supone otro riesgo.

El algoritmo QSAL detecta regularmente el ciclo de vida y la durabilidad de la SSD y, cuando el ciclo de vida de la SSD es inferior al 50 %, distribuye dinámicamente el tamaño de sobreaprovisionamiento para luego garantizar que todas las SSD tienen tiempo suficiente para reconstruir los datos antes de que finalice su ciclo de vida. Logra un equilibrio óptimo entre utilización y sobreaprovisionamiento. El algoritmo QSAL no afecta al uso del espacio tras la activación y solo provoca una insignificante reducción del rendimiento. En general, la protección de los datos en las matrices all-flash mejorará considerablemente.

El algoritmo QSAL se puede iniciar en cualquier momento y también es compatible con RAID de paridad SSD que no se hayan configurado anteriormente. Se recomienda habilitar el algoritmo QSAL antes de que el ciclo de vida de la SSD alcance el 50 %. Si se habilita demasiado tarde, pueden surgir problemas de falta de tiempo para la reconstrucción de los datos, lo que puede provocar daños en el RAID SSD.

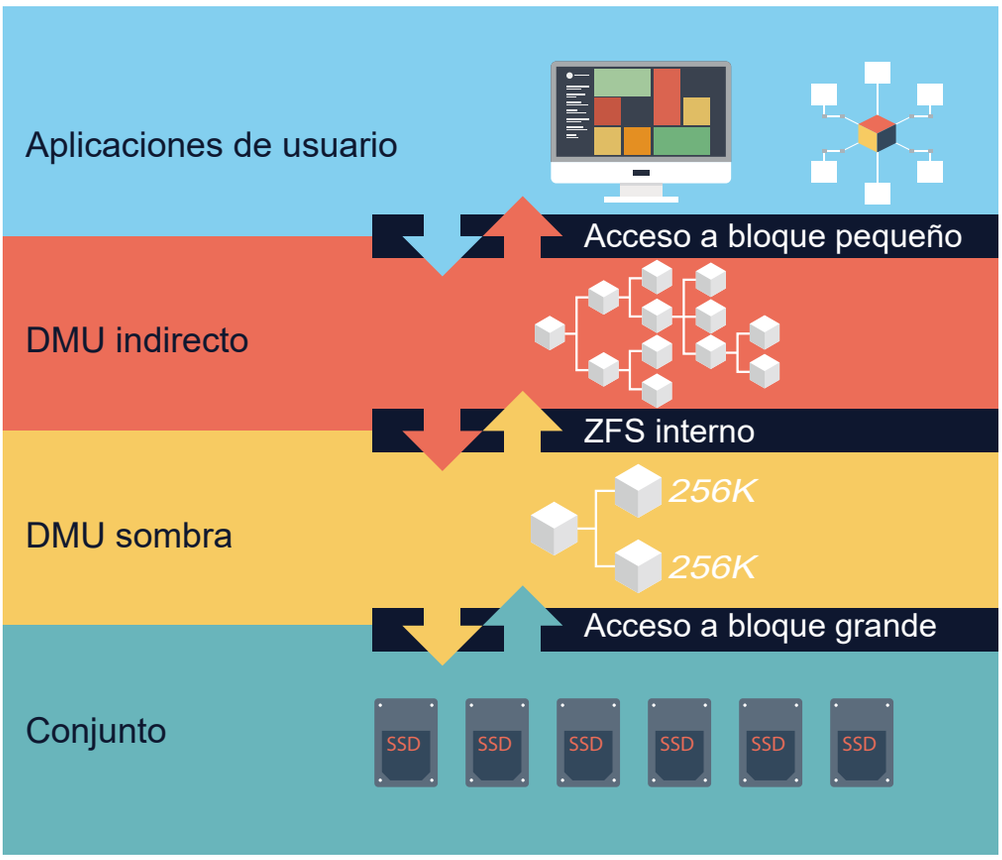


Números de patente: US10705744B2 / US20200333959A1. También incluye varias patentes confidenciales.

Coalescencia de escritura

El algoritmo de coalescencia de escritura impulsa el rendimiento de escritura para matrices all-flash.

La amplificación de escritura es un efecto indeseable de los SSD, que causa problemas de rendimiento y resistencia de la memoria flash. QuTS hero cuenta con el exclusivo algoritmo de coalescencia de escritura de QNAP, diseñado para la optimización flash al transformar todas las escrituras aleatorias en escrituras secuenciales además de reducir la E/S. No solo aumenta el rendimiento de la escritura aleatoria para los entornos all-flash, sino que también puede mejorar el ciclo de vida de la SSD.



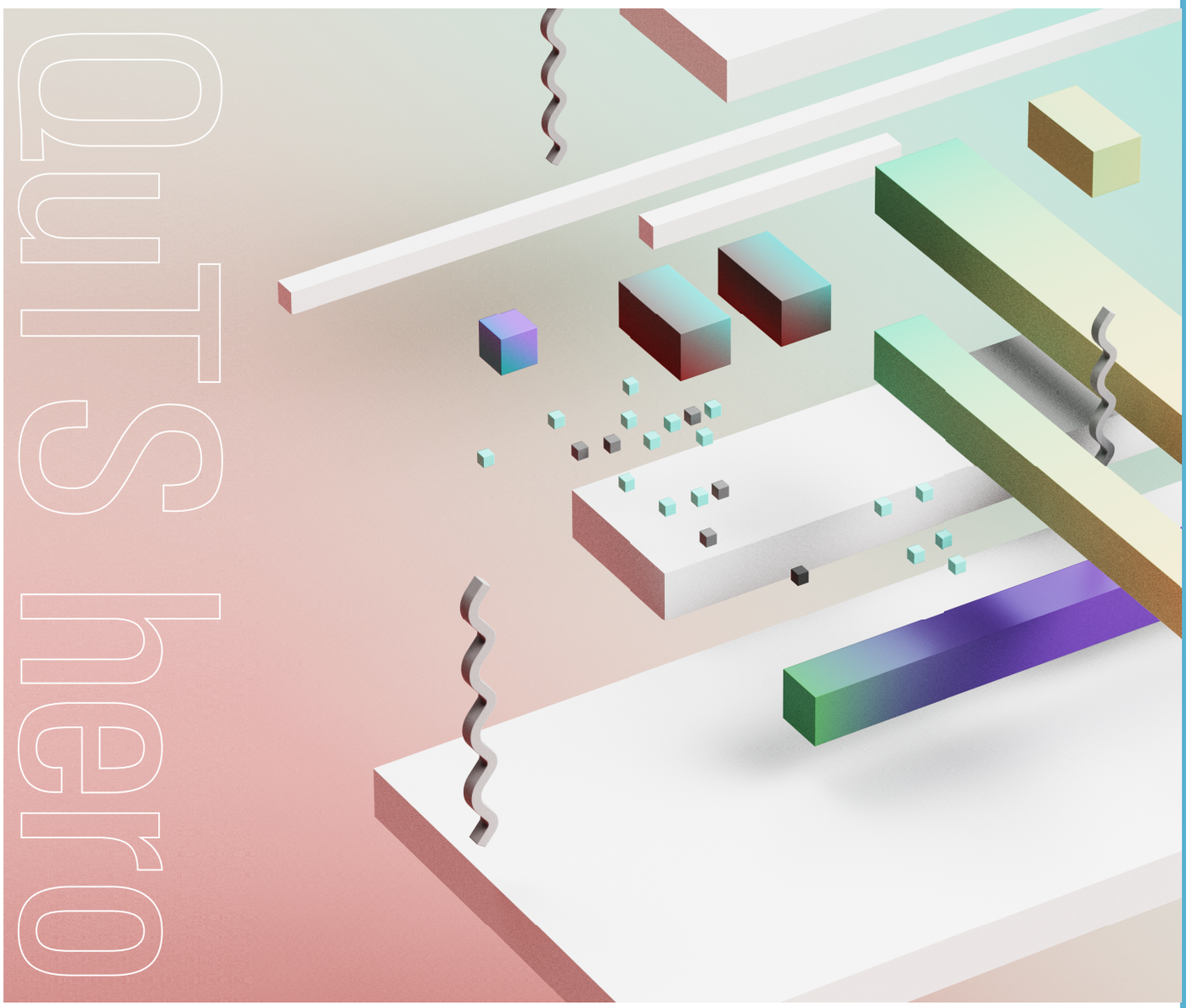
Configuración de caché de lectura recomendada (L2ARC)

La memoria desempeña un papel vital en el rendimiento ZFS, sobre todo en la transferencia de datos de alta velocidad, la deduplicación de datos, ARC y la caché. Se recomienda instalar toda la memoria posible para conseguir los máximos beneficios para el rendimiento ZFS y optimizar las cargas de trabajo.

En función de la memoria instalada en un NAS QuTS hero y la proporción que corresponda al uso de la caché de lectura, recomendamos configurar el tamaño de la caché de lectura de la siguiente manera:

- **32 GB de RAM** admite un tamaño de SSD de 1TB para la caché de lectura;
- **64 GB de RAM** admite un tamaño de SSD de 2 TB para la caché de lectura;
- **128 GB de RAM** admite un tamaño de SSD de 4 TB para la caché de lectura;
- **256 GB de RAM** admite un tamaño de SSD de 8 TB para la caché de lectura;

- **1 TB RAM** admite una SSD de caché de lectura de unos 30 TB.
- **4 TB RAM** admite una SSD de caché de lectura de unos 120 TB.


QuTS hero

Dependencia de rendimiento entre el tamaño de la memoria y los conjuntos de almacenamiento

Importante: En este capítulo solo se describe la dependencia básica entre la RAM y los conjuntos de almacenamiento. Si el sistema está ejecutando otros servicios dedicados (como Virtualization Station), se recomienda separar el uso y considerar la memoria asignada a estas máquinas virtuales.

Memoria del NAS - memoria asignada a VM = memoria usada por ZFS

Este es el concepto general para analizar el rendimiento del almacenamiento y la memoria y saber cómo identificar las prácticas recomendadas en cuanto a tamaño de la memoria para un almacenamiento de alto rendimiento. Luego describimos cómo elegir memoria suficiente para conseguir un rendimiento óptimo.

La memoria desempeña un papel vital en ZFS y afecta considerablemente al modo en que se ejecutan los conjuntos de almacenamiento de QuTS hero. Cuando el sistema ZFS se ejecuta con funciones de reducción de datos (como compactación, clonación rápida y deduplicación), el sistema necesita algunas tablas para registrar el total de referencia de estos bloques de datos. Por tanto, cuando la cantidad de datos aumente, el sistema ZFS tendrá que albergar todas las tablas de referencia posibles en la memoria y podrá gestionar las solicitudes de IO de una manera más eficaz.

En segundo lugar, el sistema ZFS admite la función de grupos de transacción para la protección de los datos. Cada modificación del archivo (p. ej., una escritura) está asociada a un grupo de transacción (TXG) concreto. En intervalos regulares, cada TXG se apagará y el conjunto emitirá una operación de sincronización para ese grupo. Un TXG también se puede apagar cuando el ARC indique que se está almacenando en la caché demasiada memoria "sucia". Cuando un TXG se cierra, se abre otro inmediatamente y las modificaciones de archivo se asociarán al nuevo TXG activo.

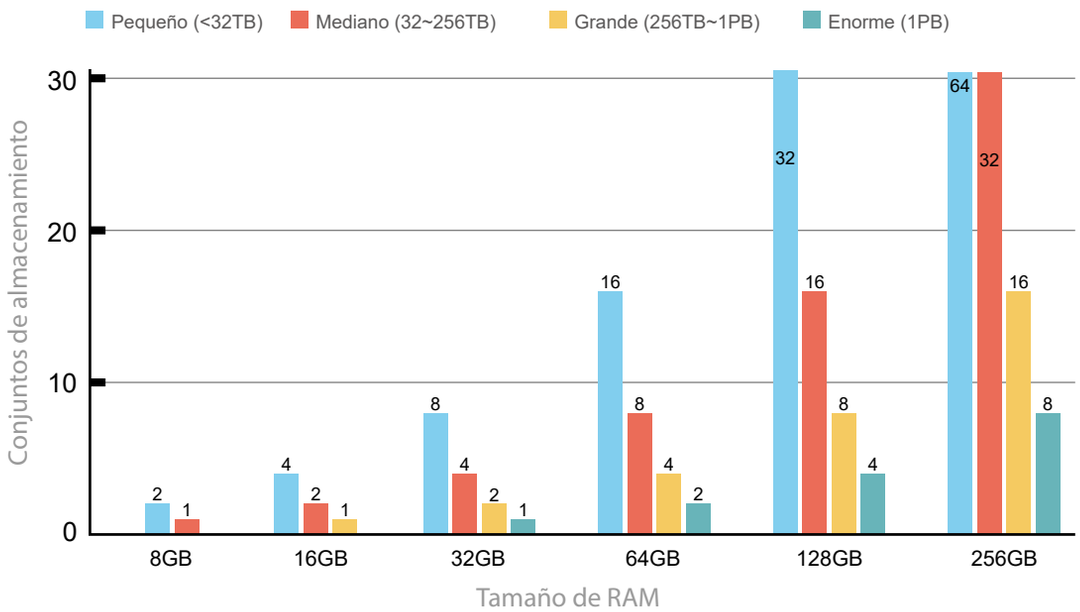
Si un TXG activo se apaga mientras uno anterior sigue sincronizando datos con el almacenamiento, las aplicaciones se bloquearán hasta que se complete la sincronización en curso. En esta situación en la que un TXG se está sincronizando, mientras que el TXG + 1 se cierra debido a limitaciones de memoria y está esperando para sincronizarse, las aplicaciones se bloquearán esperando poder escribir en TXG + 2. Necesitamos una saturación sostenida del almacenamiento o un límite de memoria para bloquear las aplicaciones.

Una sincronización del conjunto de almacenamiento implicará el envío de todos los bloques de datos de nivel 0 al disco. Una vez finalizado, todos los bloques indirectos de nivel 1 (etc.) se enviarán hasta que se completen todos los bloques que representan el nuevo estado del sistema de archivos. En ese punto, el bloque superior se actualiza para apuntar al nuevo estado coherente del conjunto de almacenamiento.

Por tanto, el sistema ZFS debe actualizar el bloque superior tras finalizar cada transacción, lo cual supone una gran carga. Si el sistema puede suministrar toda la memoria posible, los grupos de transacción pueden ser más grandes y la carga de actualizar el bloque superior será relativamente menor. Así, un mayor nivel de memoria puede ayudar eficazmente al sistema ZFS a mejorar las solicitudes de IO de escritura.

Puede consultar nuestras siguientes recomendaciones y elegir una memoria suficiente para las distintas configuraciones de almacenamiento.

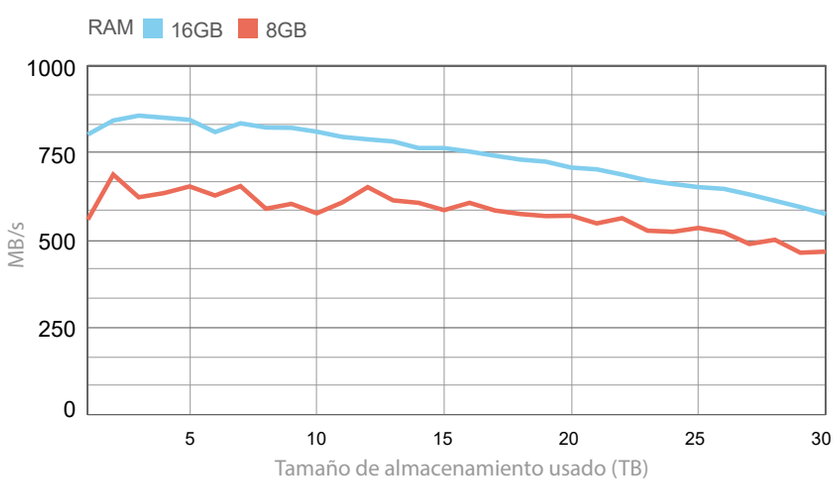
Prácticas recomendadas para garantizar el rendimiento entre el tamaño de RAM y la configuración del conjunto



Por ejemplo, si su NAS ZFS tiene 16 GB de memoria, la configuración recomendada para el conjunto de almacenamiento sería 1 conjunto grande (256 TB~1 PB), 2 conjuntos de tamaño mediano o 4 conjuntos pequeños. Este tipo de configuración de conjunto permitiría al sistema alcanzar un buen rendimiento. Si crea conjuntos que superan estas recomendaciones, su NAS funcionará de forma fiable, pero el rendimiento que puede verse reducido. Aquí usamos como referencia nuestro modelo de ZFS básico, el TS-973AX, para mostrar los distintos rendimientos entre 8 GB (requisito mínimo) y 16 GB de RAM.

(Ya hemos creado un conjunto de SSD y un conjunto de HDD de unos 35 TB. Se muestra la diferencia de rendimiento del conjunto de HDD).

TS-h973AX de 16GB vs TS-h973AX de 8 GB - Rendimiento del IO (5 HDD-R0)



QuTS hero
ZFS
documento técnico

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