



Managen von Daten mit StorageGRID

StorageGRID 11.5

NetApp
April 11, 2024

Inhalt

- Managen von Daten mit StorageGRID 1
 - Was ist ein Objekt 1
 - Schutz von Objektdaten 2
 - Das Leben eines Objekts 3

Managen von Daten mit StorageGRID

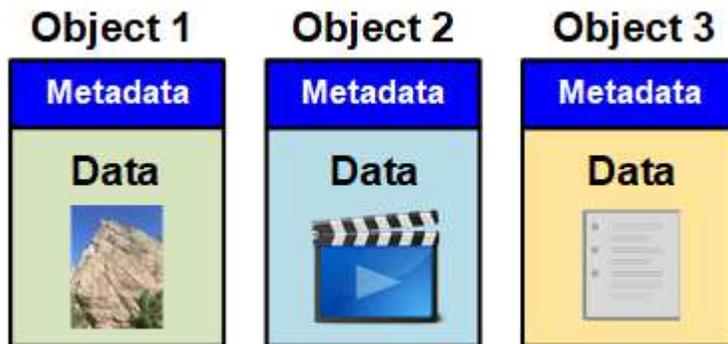
Bei der Arbeit mit dem StorageGRID System ist es hilfreich, zu verstehen, wie das StorageGRID System die Daten managt.

- "Was ist ein Objekt"
- "Schutz von Objektdaten"
- "Das Leben eines Objekts"

Was ist ein Objekt

Bei Objekt-Storage ist die Storage-Einheit ein Objekt und nicht eine Datei oder ein Block. Im Gegensatz zur Baumstruktur eines File-Systems oder Block-Storage werden die Daten im Objekt-Storage in einem flachen, unstrukturierten Layout organisiert. Objekt-Storage entkoppelt den physischen Standort der Daten von der Methode zum Speichern und Abrufen dieser Daten.

Jedes Objekt in einem objektbasierten Storage-System besteht aus zwei Teilen: Objekt-Daten und Objekt-Metadaten.



Objektdaten

Objektdaten können alles sein, z. B. ein Foto, ein Film oder eine medizinische Aufzeichnung.

Objekt-Metadaten

Objektmetadaten sind alle Informationen, die ein Objekt beschreiben. StorageGRID verwendet Objektmetadaten, um die Standorte aller Objekte im Grid zu verfolgen und den Lebenszyklus eines jeden Objekts mit der Zeit zu managen.

Objektmetadaten enthalten Informationen wie die folgenden:

- Systemmetadaten, einschließlich einer eindeutigen ID für jedes Objekt (UUID), dem Objektnamen, dem Namen des S3-Buckets oder Swift-Containers, dem Mandanten-Kontonamen oder -ID, der logischen Größe des Objekts, dem Datum und der Uhrzeit der ersten Erstellung des Objekts Und Datum und Uhrzeit der letzten Änderung des Objekts.
- Der aktuelle Speicherort der einzelnen Objektkopien oder Fragmente, deren Löschen codiert wurde

- Alle dem Objekt zugeordneten Benutzer-Metadaten.

Objektmetadaten sind individuell anpassbar und erweiterbar und bieten dadurch Flexibilität für die Nutzung von Applikationen.

Detaillierte Informationen zum StorageGRID Speichern von Objektmetadaten und -Speicherort finden Sie unter ["Management von Objekt-Metadaten-Storage"](#).

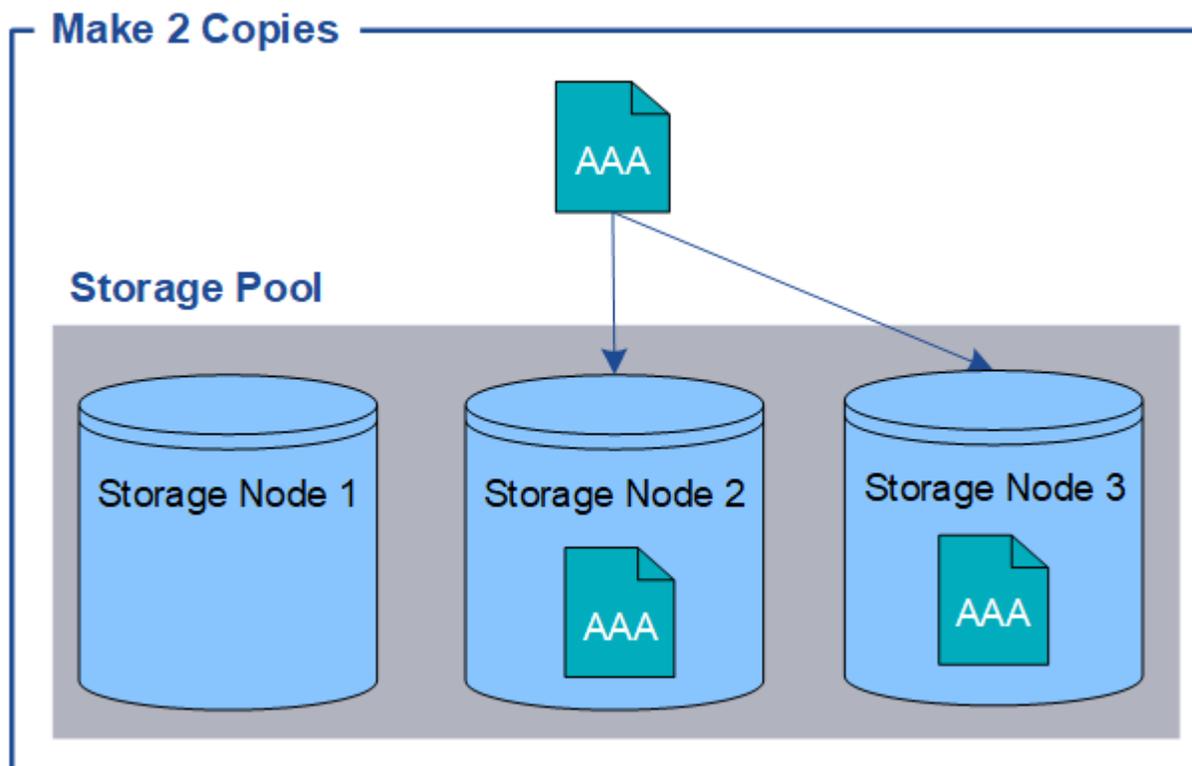
Schutz von Objektdaten

Das StorageGRID System bietet zwei Mechanismen zum Schutz von Objektdaten vor Verlust: Replizierung und Erasure Coding.

Replizierung

Wenn StorageGRID Objekte mit einer ILM-Regel (Information Lifecycle Management) übereinstimmt, die für die Erstellung replizierter Kopien konfiguriert ist, erstellt das System exakte Kopien von Objektdaten und speichert sie in Storage-Nodes, Archivierungs-Nodes oder Cloud-Storage-Pools. ILM-Regeln bestimmen die Anzahl der Kopien, die erstellt werden, wo diese Kopien gespeichert werden und wie lange sie vom System aufbewahrt werden. Falls eine Kopie verloren geht, beispielsweise aufgrund des Verlusts eines Storage-Nodes, ist das Objekt nach wie vor verfügbar, wenn eine Kopie davon an einer anderen Stelle im StorageGRID System vorhanden ist.

Im folgenden Beispiel gibt die Regel „2 Kopien erstellen“ an, dass zwei replizierte Kopien jedes Objekts in einem Speicherpool platziert werden, der drei Storage-Nodes enthält.

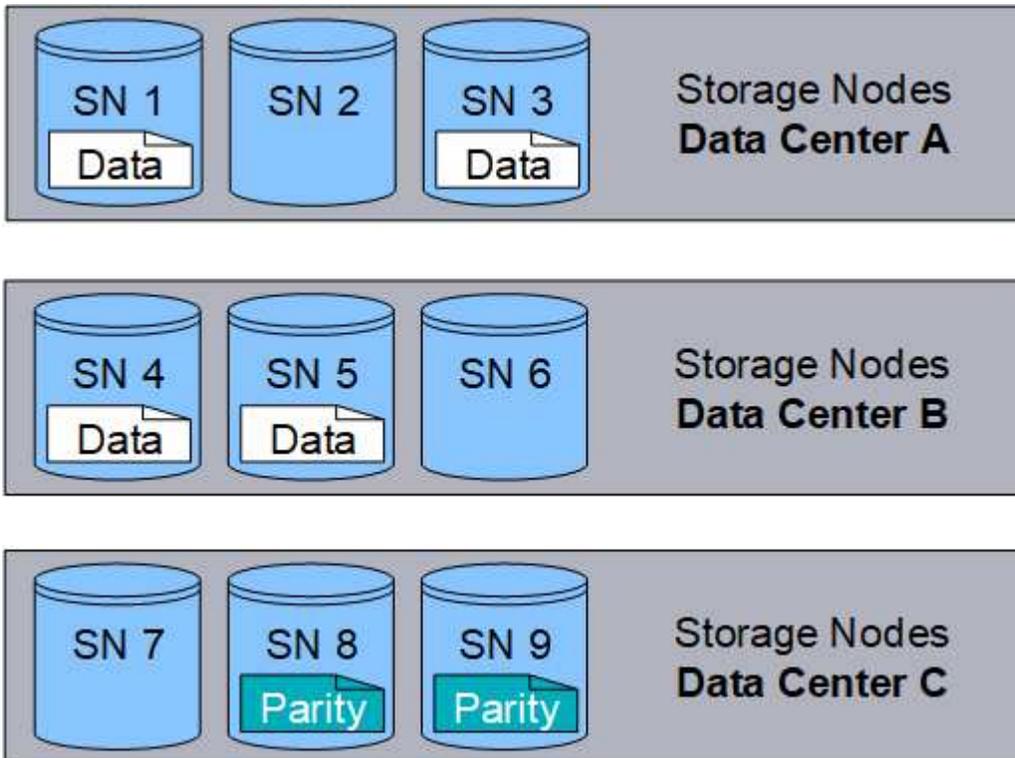


Erasure Coding

Wenn StorageGRID Objekte mit einer ILM-Regel übereinstimmt, die zur Erstellung von mit Datenkonsistenz

versehenen Kopien konfiguriert ist, werden Objektdaten in Datenfragmente zerlegt, zusätzliche Paritätsfragmente berechnet und jedes Fragment auf einem anderen Storage Node gespeichert. Wenn auf ein Objekt zugegriffen wird, wird es anhand der gespeicherten Fragmente neu zusammengesetzt. Wenn ein Daten oder ein Paritätsfragment beschädigt wird oder verloren geht, kann der Algorithmus zum Erasure Coding diese Fragmente mit einer Teilmenge der verbleibenden Daten und Paritätsfragmente neu erstellen. ILM-Regeln und Erasure Coding-Profilen bestimmen das verwendete Verfahren zum Erasure Coding-Verfahren.

Das folgende Beispiel zeigt den Einsatz von Erasure Coding für Objektdaten. In diesem Beispiel verwendet die ILM-Regel ein Codierungsschema für das Löschen von 4+2. Jedes Objekt wird in vier gleiche Datenfragmente geteilt und aus den Objektdaten werden zwei Paritätsfragmente berechnet. Jedes der sechs Fragmente ist in drei Datacentern auf einem anderen Storage Node gespeichert, um bei Node-Ausfällen oder Standortausfällen ihre Daten zu sichern.



Verwandte Informationen

["Objektmanagement mit ILM"](#)

["Mit Information Lifecycle Management"](#)

Das Leben eines Objekts

Das Leben eines Objekts besteht aus verschiedenen Etappen. Jede Phase stellt die Vorgänge dar, die mit dem Objekt auftreten.

Der Lebenszyklus eines Objekts umfasst das Aufnehmen, das Kopieren-Management, das Abrufen und Löschen von Objekten.

- **Ingest:** Der Prozess einer S3- oder Swift-Client-Anwendung, bei der ein Objekt über HTTP auf das StorageGRID-System gespeichert wird. In dieser Phase beginnt das StorageGRID-System mit der Verwaltung des Objekts.
- **Kopierverwaltung:** Der Prozess des Managements replizierter und mit Erasure Coding codierter Kopien in

StorageGRID, wie in den ILM-Regeln der aktiven ILM-Richtlinie beschrieben. Während der Kopiemanagementphase schützt StorageGRID Objektdaten vor Verlust. Dazu wird die angegebene Anzahl und der angegebene Typ von Objektkopien auf Storage-Nodes, in einem Cloud-Storage-Pool oder auf Archiv-Node erstellt und beibehalten.

- **Retrieve:** Der Prozess einer Client-Anwendung, die auf ein vom StorageGRID-System gespeichertes Objekt zugreift. Der Client liest das Objekt, das von einem Storage-Node, Cloud-Storage-Pool oder Archive Node abgerufen wird.
- **Löschen:** Der Vorgang, bei dem alle Objektkopien aus dem Raster entfernt werden. Objekte können entweder gelöscht werden, wenn eine Client-Applikation eine Löschanfrage an das StorageGRID System sendet, oder infolge eines automatischen Prozesses, der StorageGRID nach Ablauf der Nutzungsdauer des Objekts durchführt.

Verwandte Informationen

["Objektmanagement mit ILM"](#)

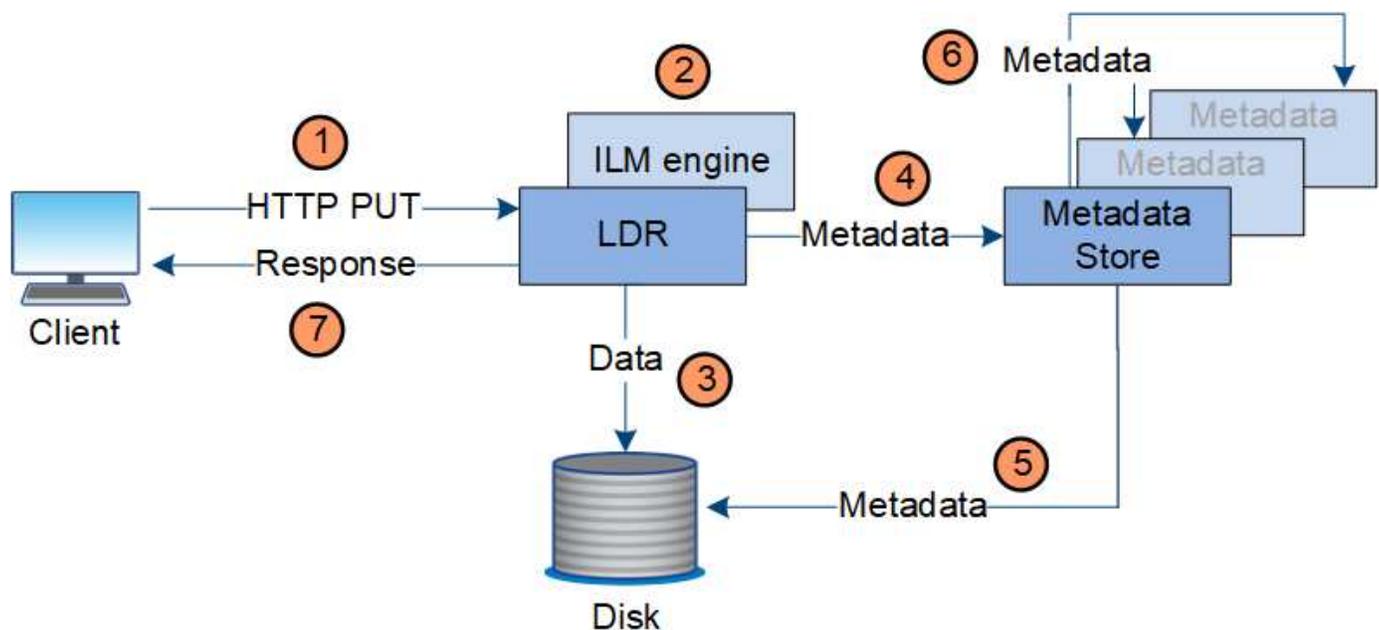
["Mit Information Lifecycle Management"](#)

Datenfluss aufnehmen

Ein Aufnahme- oder Speichervorgang besteht aus einem definierten Datenfluss zwischen dem Client und dem StorageGRID System.

Datenfluss

Wenn ein Client ein Objekt im StorageGRID-System speichert, verarbeitet der LDR-Service auf Storage Nodes die Anforderung und speichert die Metadaten und Daten auf der Festplatte.



1. Die Client-Applikation erstellt das Objekt und sendet es über eine HTTP PUT-Anforderung an das StorageGRID System.
2. Das Objekt wird anhand der ILM-Richtlinie des Systems bewertet.
3. Der LDR-Service speichert die Objektdaten als replizierte Kopie oder als Kopie mit dem Erasure Coding.

(Das Diagramm zeigt eine vereinfachte Version zum Speichern einer replizierten Kopie auf Festplatte.)

4. Der LDR-Service sendet die Objektmetadaten an den Metadatenpeicher.
5. Der Metadaten-Speicher speichert die Objekt-Metadaten auf der Festplatte.
6. Der Metadatenpeicher überträgt Kopien von Objektmetadaten an andere Storage-Nodes. Diese Kopien werden auch auf der Festplatte gespeichert.
7. Der LDR-Dienst gibt eine HTTP 200 OK-Antwort an den Client zurück, um zu bestätigen, dass das Objekt aufgenommen wurde.

Verwaltung von Kopien

Objektdaten werden von der aktiven ILM-Richtlinie und ihren ILM-Regeln gemanagt. ILM-Regeln erstellen replizierte oder Erasure-codierte Kopien, um Objektdaten vor Verlust zu schützen.

Unterschiedliche Typen und Standorte von Objektkopien können zu unterschiedlichen Zeiten der Lebensdauer des Objekts erforderlich sein. ILM-Regeln werden regelmäßig überprüft, um sicherzustellen, dass Objekte nach Bedarf platziert werden.

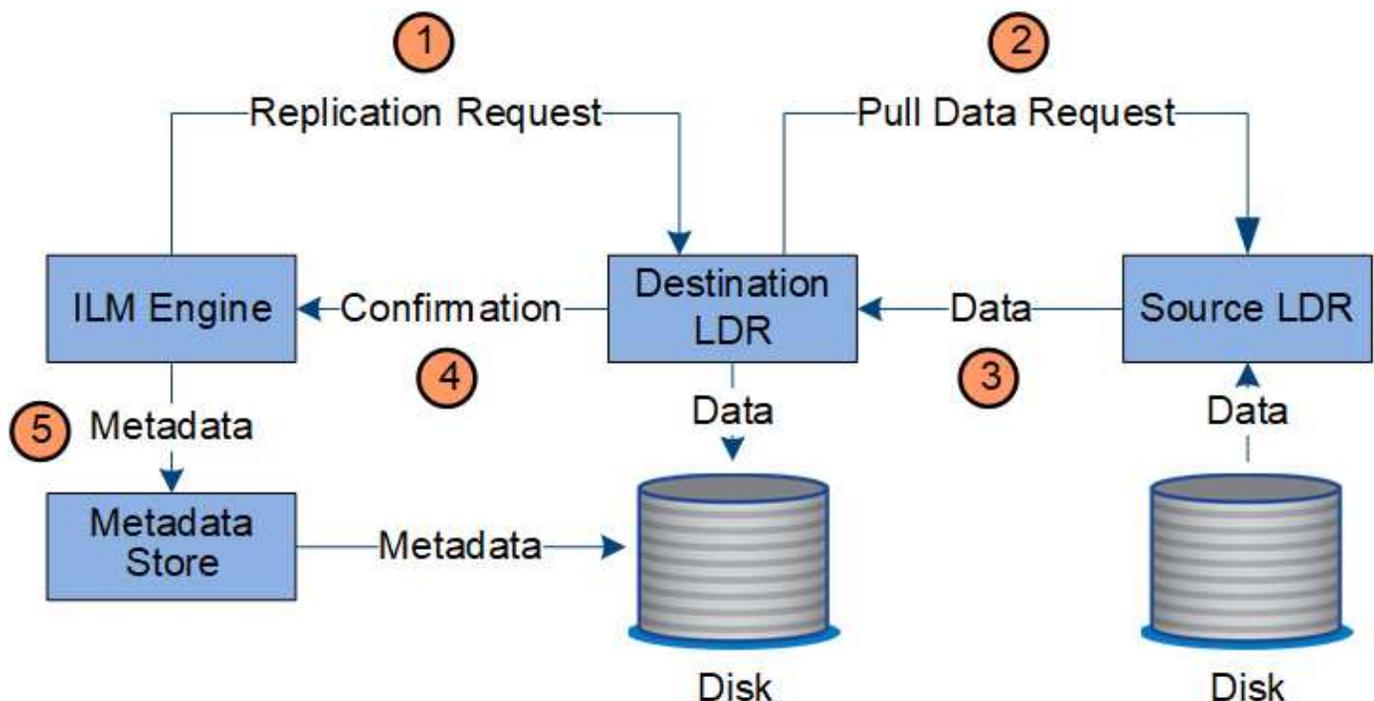
Objektdaten werden vom LDR-Service gemanagt.

Content-Schutz: Replikation

Wenn für die Anweisungen zur Content-Platzierung einer ILM-Regel replizierte Kopien von Objektdaten erforderlich sind, werden von den Storage-Nodes, die den konfigurierten Storage-Pool bilden, Kopien auf Festplatte erstellt und gespeichert.

Datenfluss

Die ILM-Engine im LDR-Service steuert die Replikation und stellt sicher, dass die korrekte Anzahl von Kopien an den richtigen Standorten und für die richtige Zeit gespeichert wird.



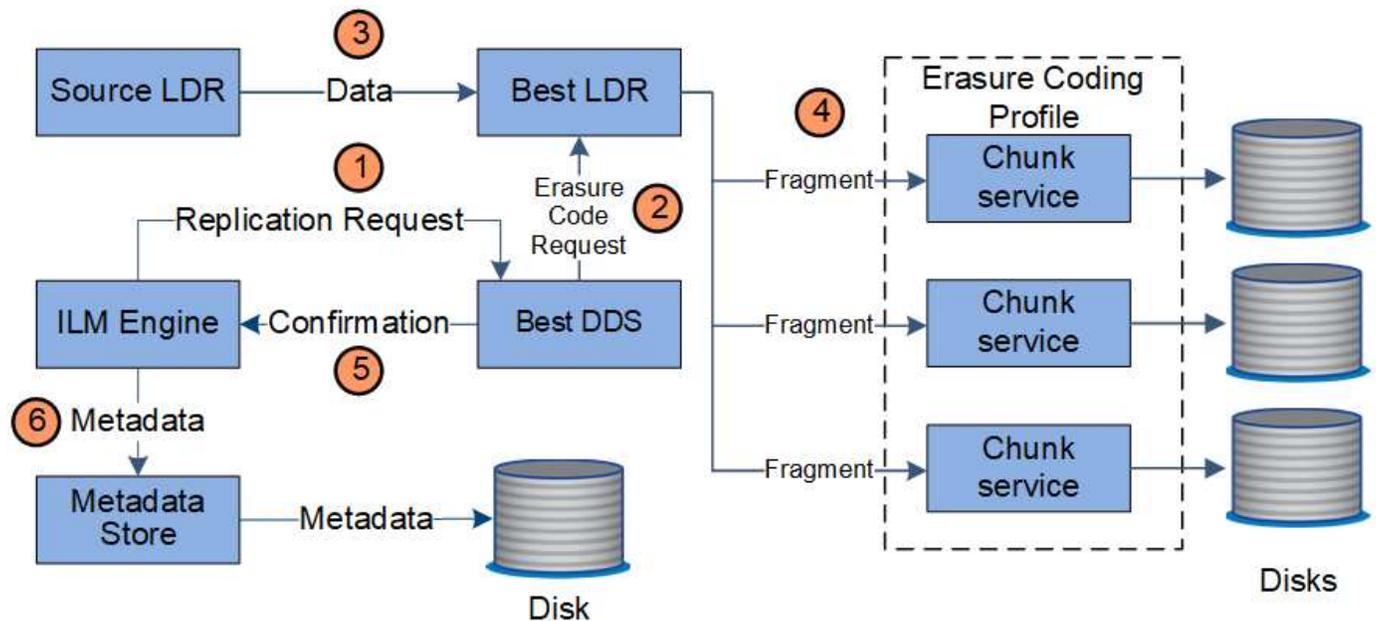
1. Die ILM-Engine fragt den ADC-Service ab, um den besten Ziel-LDR-Service innerhalb des durch die ILM-Regel festgelegten Storage-Pools zu ermitteln. Er sendet dann diesen LDR-Service einen Befehl, um die Replikation zu initiieren.
2. Der Ziel-LDR-Dienst fragt den ADC-Dienst nach dem besten Quellspeicherort ab. Anschließend sendet er eine Replikationsanfrage an den Quell-LDR-Service.
3. Der Quell-LDR-Service sendet eine Kopie an den Ziel-LDR-Service.
4. Der Ziel-LDR-Service benachrichtigt die ILM Engine, dass die Objektdaten gespeichert wurden.
5. Die ILM-Engine aktualisiert den Metadatenpeicher mit Objektspeichermetadaten.

Content Protection: Erasure Coding

Wenn eine ILM-Regel Anweisungen zur Erstellung von Erasure-codierten Kopien von Objektdaten enthält, werden Objektdaten im Rahmen des entsprechenden Erasure Coding-Schemas in Daten- und Paritätsfragmente unterteilt und diese Fragmente über die im Erasure Coding-Profil konfigurierten Storage-Nodes verteilt.

Datenfluss

Die ILM-Engine, die eine Komponente des LDR-Service ist, steuert das Erasure Coding-Verfahren und stellt sicher, dass das Erasure Coding-Profil auf Objektdaten angewendet wird.



1. Die ILM-Engine fragt den ADC-Service ab, um zu bestimmen, welcher DDS-Service den Erasure Coding-Vorgang am besten ausführen kann. Sobald die ILM-Engine ermittelt wurde, sendet sie eine „Initiierung“-Anforderung an den Service.
2. Der DDS-Dienst weist ein LDR an, den Code der Objektdaten zu löschen.
3. Der Quell-LDR-Service sendet eine Kopie an den für das Erasure Coding ausgewählten LDR-Service.
4. Nach der entsprechenden Anzahl von Paritäts- und Datenfragmenten verteilt der LDR-Service diese Fragmente auf die Storage-Nodes (Chunk-Services), aus denen sich der Speicherpool des Erasure Coding-Profiles besteht.
5. Der LDR-Service benachrichtigt die ILM-Engine und bestätigt, dass Objektdaten erfolgreich verteilt werden.

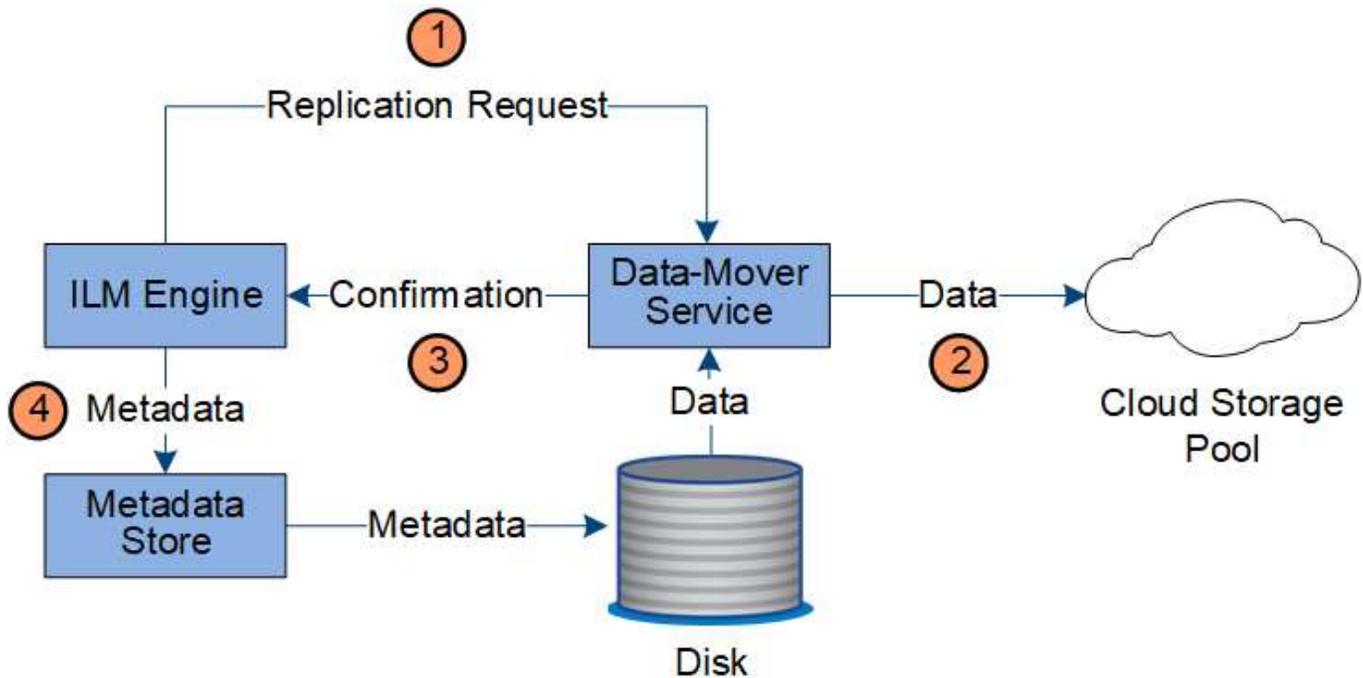
6. Die ILM-Engine aktualisiert den Metadatenpeicher mit Objektspeichermetadaten.

Content-Sicherung: Cloud Storage Pool

Wenn für die Anweisungen zur Content-Platzierung einer ILM-Regel eine replizierte Kopie von Objektdaten in einem Cloud Storage-Pool gespeichert werden muss, werden Objektdaten in den externen S3-Bucket oder Azure Blob-Storage-Container verschoben, der für den Cloud Storage Pool angegeben wurde.

Datenfluss

Die ILM-Engine, die eine Komponente des LDR-Service ist, und der Data Mover-Service steuern die Verschiebung von Objekten in den Cloud-Speicherpool.

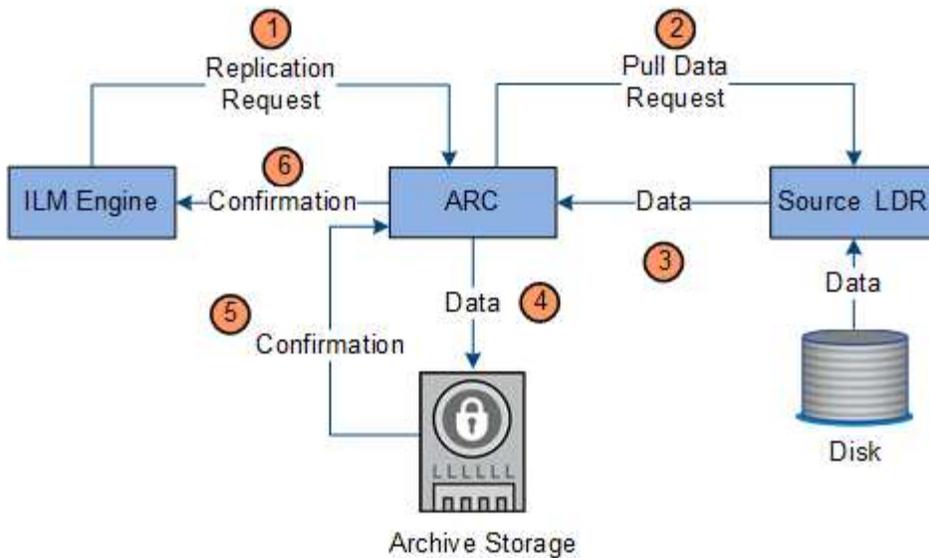


1. Die ILM-Engine wählt einen Data Mover-Service zur Replizierung in den Cloud-Storage-Pool aus.
2. Der Data Mover-Service sendet die Objektdaten an den Cloud-Speicherpool.
3. Der Data Mover-Service benachrichtigt die ILM-Engine, dass die Objektdaten gespeichert wurden.
4. Die ILM-Engine aktualisiert den Metadatenpeicher mit Objektspeichermetadaten.

Content-Schutz: Archivierung

Ein Archivierungsvorgang besteht aus einem definierten Datenfluss zwischen dem StorageGRID System und dem Client.

Wenn die ILM-Richtlinie erfordert, dass eine Kopie der Objektdaten archiviert wird, sendet die ILM-Engine, die eine Komponente des LDR-Service ist, eine Anforderung an den Archiv-Node, der wiederum eine Kopie der Objektdaten an das Ziel-Archiv-Storage-System sendet.



1. Die ILM-Engine sendet eine Anforderung an den ARC-Service, eine Kopie auf Archivmedien zu speichern.
2. Der ARC-Dienst fragt den ADC-Service nach dem besten Quellspeicherort ab und sendet eine Anfrage an den Quell-LDR-Dienst.
3. Der ARC-Dienst ruft Objektdaten aus dem LDR-Dienst ab.
4. Der ARC-Dienst sendet die Objektdaten an das Archivmedienziel.
5. Das Archivmedium benachrichtigt den ARC-Dienst, dass die Objektdaten gespeichert wurden.
6. Der ARC-Dienst benachrichtigt die ILM-Engine, dass die Objektdaten gespeichert wurden.

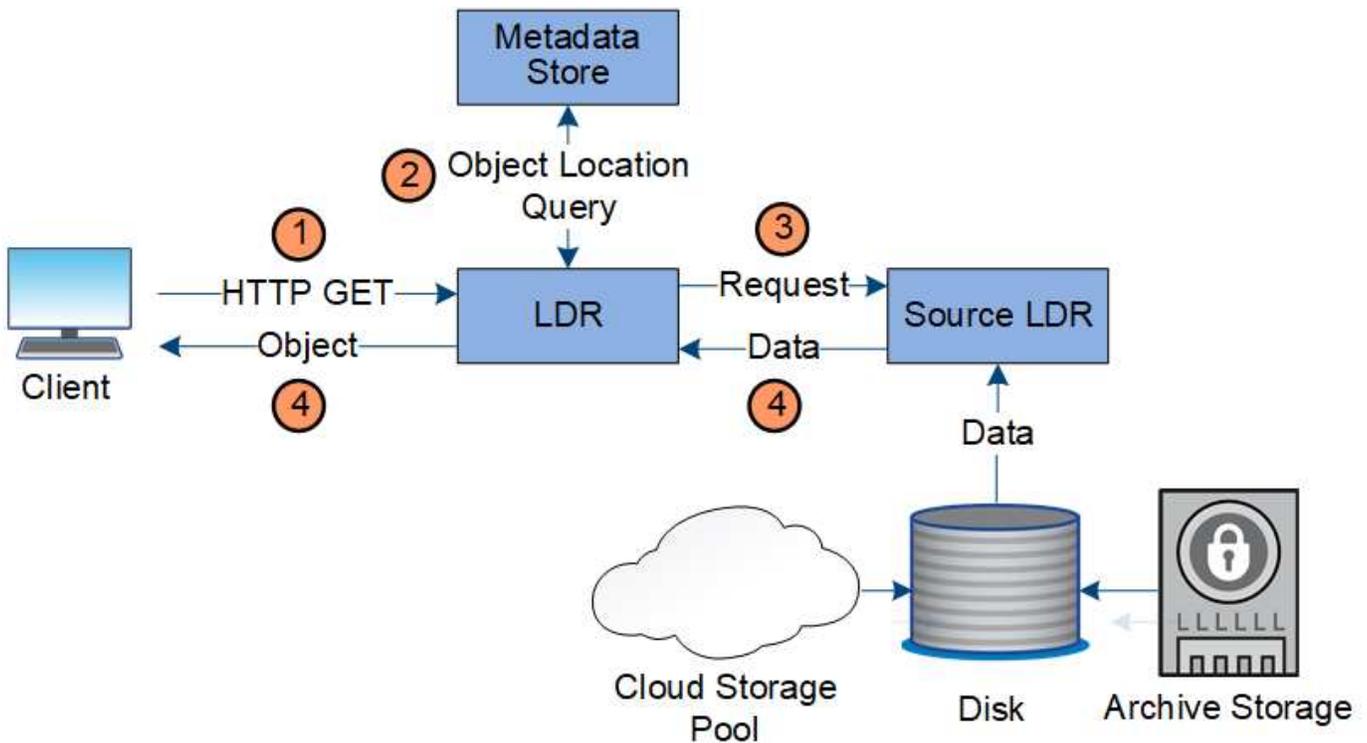
Abrufen des Datenflusses

Ein Abrufvorgang besteht aus einem definierten Datenfluss zwischen dem StorageGRID-System und dem Client. Das System verwendet Attribute, um den Abruf des Objekts von einem Storage-Node oder ggf. einem Cloud-Storage-Pool oder Archiv-Node zu verfolgen.

Der LDR-Service des Storage Node fragt den Metadatenpeicher nach dem Speicherort der Objektdaten ab und ruft ihn vom Quell-LDR-Service ab. Bevorzugt wird der Abruf von einem Storage Node durchgeführt. Wenn das Objekt auf einem Speicherknoten nicht verfügbar ist, wird die Abfrage an einen Cloud-Speicherpool oder einen Archiv-Node geleitet.



Wenn sich die einzige Objektkopie auf AWS Glacier Storage oder in der Azure Archiveebene befindet, muss die Client-Applikation eine Anfrage zur Wiederherstellung NACH S3-Objekten stellen, um eine abrufbare Kopie in dem Cloud Storage Pool wiederherzustellen.



1. Der LDR-Service erhält eine Abrufanforderung von der Client-Anwendung.
2. Der LDR-Service fragt den Metadatenpeicher nach dem Objektdatenstandort und den Metadaten ab.
3. Der LDR-Service leitet die Abfrage an den Quell-LDR-Service weiter.
4. Der Quell-LDR-Dienst gibt die Objektdaten aus dem abgefragten LDR-Dienst zurück und das System gibt das Objekt an die Client-Anwendung zurück.

Löschen des Datenflusses

Alle Objektkopien werden aus dem StorageGRID System entfernt, wenn ein Client einen Löschvorgang durchführt oder die Lebensdauer des Objekts abgelaufen ist. Dies wird automatisch entfernt. Es gibt einen definierten Datenfluss zum Löschen von Objekten.

Löschhierarchie

StorageGRID bietet verschiedene Methoden zur Steuerung der Aufbewahrung oder Löschung von Objekten. Objekte können nach Client-Anforderungen oder automatisch gelöscht werden. StorageGRID priorisiert alle S3 Object Lock-Einstellungen bei Löschanfragen von Clients, die nach ihrer Wichtigkeit über den S3-Bucket-Lebenszyklus und die Anweisungen zur ILM-Platzierung priorisiert werden.

- **S3 Object Lock:** Wenn die globale S3 Object Lock-Einstellung für das Grid aktiviert ist, können S3-Clients Buckets mit aktivierter S3-Objektsperre erstellen und dann über die S3-REST-API Aufbewahrungseinstellungen für jede Objektversion festlegen, die diesem Bucket hinzugefügt wurde.
 - Eine Objektversion, die sich unter einer gesetzlichen Aufbewahrungspflichten befindet, kann nicht mit irgendeiner Methode gelöscht werden.
 - Bevor das Aufbewahrungsdatum einer Objektversion erreicht ist, kann diese Version nicht mit einer Methode gelöscht werden.
 - Objekte in Buckets, für die S3 Objektsperre aktiviert ist, werden durch ILM „Forever“ beibehalten.

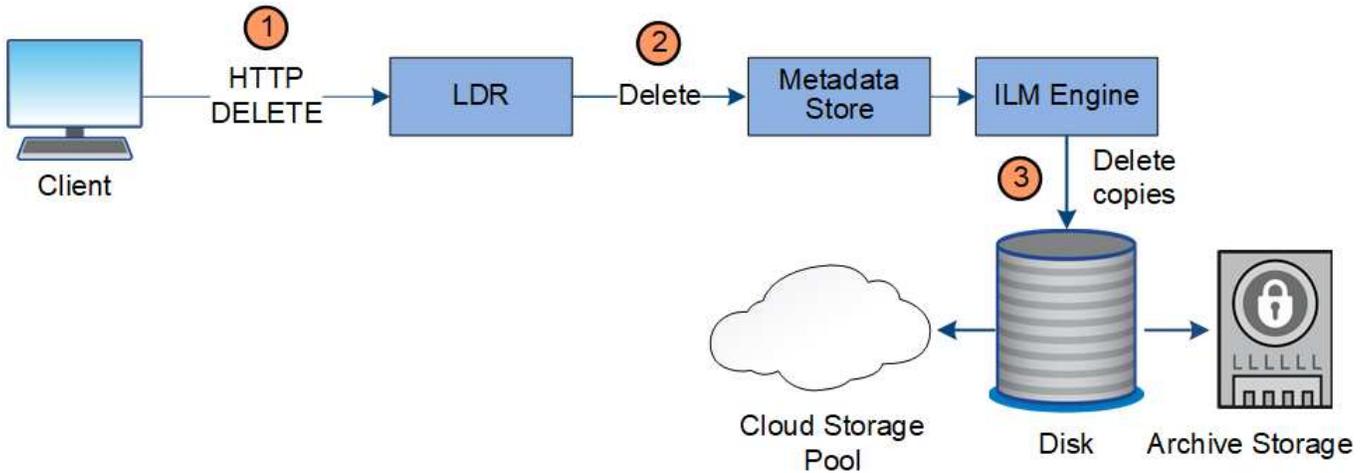
Nachdem jedoch eine Aufbewahrungsfrist erreicht ist, kann eine Objektversion durch eine Client-Anfrage oder den Ablauf des Bucket-Lebenszyklus gelöscht werden.

- **Client delete Request:** Ein S3- oder Swift-Client kann eine delete-Objekt-Anfrage stellen. Wenn ein Client ein Objekt löscht, werden alle Kopien des Objekts aus dem StorageGRID System entfernt.
- **S3-Bucket-Lebenszyklus:** S3-Clients können eine Lebenszykluskonfiguration zu ihren Buckets hinzufügen, die eine Ablaufaktion angibt. Wenn ein Bucket-Lebenszyklus vorhanden ist, löscht StorageGRID automatisch alle Kopien eines Objekts, wenn das in der Aktion „Ablaufdatum“ angegebene Datum oder die Anzahl der Tage erfüllt werden, es sei denn, der Client löscht das Objekt zuerst.
- **ILM-Platzierungsanweisungen:** Vorausgesetzt, dass für den Bucket keine S3-Objektsperre aktiviert ist und es keinen Bucket-Lebenszyklus gibt, löscht StorageGRID automatisch ein Objekt, wenn der letzte Zeitraum der ILM-Regel endet und es keine weiteren Platzierungen für das Objekt gibt.



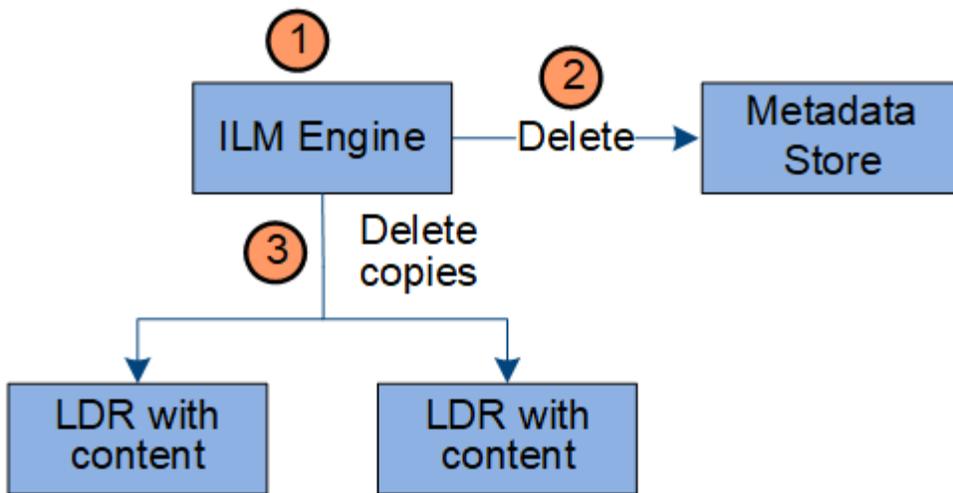
Die Aktion „Ablaufdatum“ in einem S3-Bucket-Lebenszyklus überschreibt immer die ILM-Einstellungen. Aus diesem Grund kann ein Objekt auch dann im Grid verbleiben, wenn ILM-Anweisungen zum Auflagen des Objekts verfallen sind.

Datenfluss für Clientlöschungen



1. Der LDR-Dienst erhält eine Löschanforderung von der Client-Anwendung.
2. Der LDR-Service aktualisiert den Metadatenpeicher, sodass das Objekt auf die Client-Anforderungen gelöscht wird, und weist die ILM-Engine an, alle Kopien von Objektdaten zu entfernen.
3. Das Objekt wurde aus dem System entfernt. Der Metadatenpeicher wird aktualisiert, um Objektdaten zu entfernen.

Datenfluss für ILM-Löschungen



1. Die ILM-Engine legt fest, dass das Objekt gelöscht werden muss.
2. Die ILM-Engine benachrichtigt den Metadatenpeicher. Der Metadatenpeicher aktualisiert Objektmetadaten, sodass das Objekt auf Client-Anforderungen gelöscht aussieht.
3. Die ILM-Engine entfernt alle Kopien des Objekts. Der Metadatenpeicher wird aktualisiert, um Objektmetadaten zu entfernen.

Copyright-Informationen

Copyright © 2024 NetApp. Alle Rechte vorbehalten. Gedruckt in den USA. Dieses urheberrechtlich geschützte Dokument darf ohne die vorherige schriftliche Genehmigung des Urheberrechtinhabers in keiner Form und durch keine Mittel – weder grafische noch elektronische oder mechanische, einschließlich Fotokopieren, Aufnehmen oder Speichern in einem elektronischen Abrufsystem – auch nicht in Teilen, vervielfältigt werden.

Software, die von urheberrechtlich geschütztem NetApp Material abgeleitet wird, unterliegt der folgenden Lizenz und dem folgenden Haftungsausschluss:

DIE VORLIEGENDE SOFTWARE WIRD IN DER VORLIEGENDEN FORM VON NETAPP ZUR VERFÜGUNG GESTELLT, D. H. OHNE JEGLICHE EXPLIZITE ODER IMPLIZITE GEWÄHRLEISTUNG, EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE STILLSCHWEIGENDE GEWÄHRLEISTUNG DER MARKTGÄNGIGKEIT UND EIGNUNG FÜR EINEN BESTIMMTEN ZWECK, DIE HIERMIT AUSGESCHLOSSEN WERDEN. NETAPP ÜBERNIMMT KEINERLEI HAFTUNG FÜR DIREKTE, INDIREKTE, ZUFÄLLIGE, BESONDERE, BEISPIELHAFT SCHÄDEN ODER FOLGESCHÄDEN (EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE BESCHAFFUNG VON ERSATZWAREN ODER -DIENSTLEISTUNGEN, NUTZUNGS-, DATEN- ODER GEWINNVERLUSTE ODER UNTERBRECHUNG DES GESCHÄFTSBETRIEBS), UNABHÄNGIG DAVON, WIE SIE VERURSACHT WURDEN UND AUF WELCHER HAFTUNGSTHEORIE SIE BERUHEN, OB AUS VERTRAGLICH FESTGELEGTER HAFTUNG, VERSCHULDENSUNABHÄNGIGER HAFTUNG ODER DELIKTSHAFTUNG (EINSCHLIESSLICH FAHRLÄSSIGKEIT ODER AUF ANDEREM WEGE), DIE IN IRGEND EINER WEISE AUS DER NUTZUNG DIESER SOFTWARE RESULTIEREN, SELBST WENN AUF DIE MÖGLICHKEIT DERARTIGER SCHÄDEN HINGEWIESEN WURDE.

NetApp behält sich das Recht vor, die hierin beschriebenen Produkte jederzeit und ohne Vorankündigung zu ändern. NetApp übernimmt keine Verantwortung oder Haftung, die sich aus der Verwendung der hier beschriebenen Produkte ergibt, es sei denn, NetApp hat dem ausdrücklich in schriftlicher Form zugestimmt. Die Verwendung oder der Erwerb dieses Produkts stellt keine Lizenzierung im Rahmen eines Patentrechts, Markenrechts oder eines anderen Rechts an geistigem Eigentum von NetApp dar.

Das in diesem Dokument beschriebene Produkt kann durch ein oder mehrere US-amerikanische Patente, ausländische Patente oder anhängige Patentanmeldungen geschützt sein.

ERLÄUTERUNG ZU „RESTRICTED RIGHTS“: Nutzung, Vervielfältigung oder Offenlegung durch die US-Regierung unterliegt den Einschränkungen gemäß Unterabschnitt (b)(3) der Klausel „Rights in Technical Data – Noncommercial Items“ in DFARS 252.227-7013 (Februar 2014) und FAR 52.227-19 (Dezember 2007).

Die hierin enthaltenen Daten beziehen sich auf ein kommerzielles Produkt und/oder einen kommerziellen Service (wie in FAR 2.101 definiert) und sind Eigentum von NetApp, Inc. Alle technischen Daten und die Computersoftware von NetApp, die unter diesem Vertrag bereitgestellt werden, sind gewerblicher Natur und wurden ausschließlich unter Verwendung privater Mittel entwickelt. Die US-Regierung besitzt eine nicht ausschließliche, nicht übertragbare, nicht unterlizenzierbare, weltweite, limitierte unwiderrufliche Lizenz zur Nutzung der Daten nur in Verbindung mit und zur Unterstützung des Vertrags der US-Regierung, unter dem die Daten bereitgestellt wurden. Sofern in den vorliegenden Bedingungen nicht anders angegeben, dürfen die Daten ohne vorherige schriftliche Genehmigung von NetApp, Inc. nicht verwendet, offengelegt, vervielfältigt, geändert, aufgeführt oder angezeigt werden. Die Lizenzrechte der US-Regierung für das US-Verteidigungsministerium sind auf die in DFARS-Klausel 252.227-7015(b) (Februar 2014) genannten Rechte beschränkt.

Markeninformationen

NETAPP, das NETAPP Logo und die unter <http://www.netapp.com/TM> aufgeführten Marken sind Marken von NetApp, Inc. Andere Firmen und Produktnamen können Marken der jeweiligen Eigentümer sein.