



# **NetApp und Dremio's Next Generation Hybrid Iceberg Lakehouse Lösung**

NetApp Solutions

NetApp  
December 19, 2024

# Inhalt

- NetApp und Dremio's Next Generation Hybrid Iceberg Lakehouse Lösung ..... 1
  - Die hybride Iceberg Lakehouse Lösung der nächsten Generation von NetApp und Dremio ..... 1
  - Lösungsüberblick ..... 1
  - Technologieanforderungen Erfüllt ..... 3
  - Bereitstellungsverfahren ..... 4
  - Übersicht zur Lösungsüberprüfung ..... 15
  - Anwendungsfälle Von Kunden ..... 22
  - Schlussfolgerung ..... 24

# NetApp und Dremio's Next Generation Hybrid Iceberg Lakehouse Lösung

## Die hybride Iceberg Lakehouse Lösung der nächsten Generation von NetApp und Dremio

In diesem Dokument besprechen wir die Implementierungsdetails von Dremio mit verschiedenen Datenquellen von NetApp Storage Controllern, einschließlich ONTAP S3, NAS und StorageGRID. Während der Bereitstellung verwendeten wir das TPC-DS Benchmark-Tool, um 99 SQL-Abfragen über verschiedene Quellen hinweg auszuführen. In diesem Dokument werden auch Anwendungsfälle für NetApp erläutert sowie ein Anwendungsfall mit Kunden aus dem Bereich Automobilvertrieb.

### Autoren

Karthikeyan Nagalingam, Principal Architect, NetApp, Inc. Roger Frey, VP, Alliances, Dremio Corporation. Mark Shainman, Principal Product Marketing Manager. Dremio Corporation.

## Lösungsüberblick

Die Hybrid Iceberg Lakehouse Lösung bietet einzigartige Vorteile bei der Bewältigung der Herausforderungen von Data-Lake-Kunden. Durch Nutzung der Dremio Unified Lakehouse-Plattform sowie von NetApp ONTAP-, StorageGRID- und NetApp-Lösungen können Unternehmen ihren Geschäftsbetrieb wesentlich voranbringen. Die Lösung bietet nicht nur Zugriff auf mehrere Datenquellen, einschließlich NetApp-Quellen, sondern verbessert auch die gesamte analytische Leistung und hilft Unternehmen, geschäftliche Einblicke zu gewinnen, die zu geschäftlichem Wachstum führen.

## Übersicht über NetApp

- Die NetApp Angebote wie ONTAP und StorageGRID ermöglichen die Trennung von Storage und Computing und ermöglichen somit eine optimale Ressourcenauslastung je nach spezifischen Anforderungen. Dank dieser Flexibilität sind Kunden in der Lage, ihren Storage mithilfe von NetApp Storage-Lösungen unabhängig zu skalieren
- Mithilfe der Storage Controller von NetApp können Kunden ihre Vektordatenbank mithilfe der NFS- und S3-Protokolle effizient mit Daten versorgen. Diese Protokolle erleichtern die Speicherung der Kundendaten und das Management des Vektordatenbankindexes, sodass nicht mehr mehrere Kopien von Daten benötigt werden, auf die über Datei- und Objektmethoden zugegriffen wird.
- NetApp ONTAP bietet nativen Support für NAS und Objekt-Storage bei führenden Cloud-Service-Providern wie AWS, Azure und Google Cloud. Diese breite Kompatibilität gewährleistet eine nahtlose Integration und ermöglicht damit Datenmobilität für Kunden, globale Verfügbarkeit, Disaster Recovery, dynamische Skalierbarkeit und hohe Performance.

## StorageGRID

Unser branchenführender Objekt-Storage-StorageGRID bietet eine leistungsstarke Richtlinien-Engine für automatisierte Datenablage, flexible Implementierungsoptionen und unübertroffene Beständigkeit dank

mehrstufigem Codierungsverfahren zur Fehlerkorrektur. Es verfügt über eine skalierbare Architektur, die Milliarden Objekte und Petabyte an Daten in einem Single Namespace unterstützt. Die Lösung ermöglicht eine Hybrid-Cloud-Integration und Daten-Tiering auf wichtige Cloud-Plattformen. Ein Bericht von IDC MarketScape Worldwide Object-Based Vendor Assessment 2019 belegt die Spitzenposition.

Darüber hinaus zeichnet sich StorageGRID durch das Management unstrukturierter Daten nach Bedarf mit softwaredefiniertem Objekt-Storage, Geo-Redundanz und Multi-Site-Funktionen aus. Sie umfasst richtlinienbasiertes Information Lifecycle Management und bietet Funktionen zur Cloud-Integration wie Spiegelung und Suche. Es verfügt über verschiedene Zertifizierungen, darunter Common Criteria, NF203 Digital Safe Component, ISO/IEC 25051, KPMG und Cohasset Compliance Assessment.

Zusammengefasst bietet NetApp StorageGRID leistungsstarke Funktionen, Skalierbarkeit, Hybrid-Cloud-Integration und Compliance-Zertifizierungen für ein effizientes Management unstrukturierter Daten jeder Größenordnung.

## **NetApp ONTAP**

NetApp ONTAP ist eine robuste Storage-Lösung, die eine breite Palette an Enterprise-Funktionen bietet. Es umfasst Snapshot, das applikationskonsistente und manipulationssichere sofortige Backups bietet. SnapRestore ermöglicht eine nahezu sofortige Wiederherstellung von Backups nach Bedarf, während SnapMirror integrierte Funktionen für Remote Backup und Disaster Recovery bietet. Darüber hinaus bietet die Lösung Autonomous Ransomware Protection (ARP) und sorgt mit Funktionen wie der Verifizierung durch mehrere Administratoren, der Verschlüsselung von Daten im Ruhezustand mit FIPS-Zertifizierung, Verschlüsselung von Daten auf der Übertragungsstrecke, Multi-Faktor-Authentifizierung (MFA) und rollenbasierter Zugriffssteuerung (Role Based Access Control, RBAC) für Datensicherheit. Umfassende Funktionen für Protokollierung, Revision, integriertes und externes Verschlüsselungsmanagement, sicheres Löschen und sicheres Management mehrerer Mandanten sorgen für zusätzliche Datensicherheit und Compliance.

NetApp ONTAP verfügt außerdem über SnapLock, das die Einhaltung gesetzlicher Vorschriften ermöglicht und ein hohes Maß an Integrität, Performance und Aufbewahrung bei niedrigen Gesamtbetriebskosten bietet. Es ist vollständig in NetApp ONTAP® 9 integriert und bietet Schutz vor böswilligen Handlungen, böswilligen Administratoren und Ransomware.

Die Lösung umfasst NSE/NVE-Verschlüsselung für die Verschlüsselung von Daten im Ruhezustand und im Übertragungsprozess, Multi-Faktor-Administratorzugriff und die Verifizierung durch mehrere Administratoren. Active IQ bietet auf KI-Basis prädiktive Analysen und Korrekturmaßnahmen. QoS sorgt dabei für Workload-Kontrolle der Servicequalität. Die Management- und Automatisierungsintegration ist intuitiv dank SysMgr/GUI/CLI/API. FabricPool ermöglicht automatisches Daten-Tiering, und die Lösung bietet Effizienz durch Inline-Datenkomprimierung, Deduplizierung und Data-Compaction. NetApp garantiert dem Kunden das kostenlose Erreichen der Workload-Effizienzziele.

NetApp ONTAP unterstützt verschiedene Protokolle, darunter NVMe/FC, FC, NVMe/TCP, iSCSI, NFS, SMB und S3, wodurch es sich zu einer Unified-Storage-Lösung entwickelt. Insgesamt bietet NetApp ONTAP umfangreiche Enterprise-Funktionen sowie robuste Sicherheit, Compliance, Effizienz und Vielseitigkeit, um verschiedenste Storage-Anforderungen zu erfüllen.

## **Dremio Übersicht**

Dremio ist die Unified Lakehouse Plattform für Self-Service-Analysen und KI. Die Dremio Unified Analytics Plattform bringt die Benutzer näher an die Daten heran und bietet Flexibilität, Skalierbarkeit und Performance am See zu einem Bruchteil der Kosten älterer Data Warehouse-Lösungen. Dremio ermöglicht „Shift-Left“-Analysen zur Eliminierung komplexer und teurer Datenintegration und ETL und ermöglicht nahtlose Analysen auf Unternehmensebene, ohne Daten verschieben zu müssen. Dremio bietet außerdem:

- Benutzerfreundliche Self-Service-Analyse über eine universelle semantische Ebene und eine eng integrierte, hochleistungsfähige SQL-Abfrageengine, die das Verbinden, Steuern und Analysieren aller Daten sowohl in der Cloud als auch On-Premises vereinfacht.
- Die nativen Lakehouse-Management-Funktionen von Dremio in Apache Iceberg vereinfachen die Datenerkennung und automatisieren die Datenoptimierung. Sie ermöglichen hochperformante Analysen mit von Git inspirierter Data-Versionierung.
- Dremio basiert auf Open Source und offenen Standards und ermöglicht es Unternehmen, sich nicht an eine Bindung zu binden und für Innovationen positioniert zu bleiben. Enterprise-Unternehmen vertrauen auf Dremio als am einfachsten zu verwendende Lakehouse-Plattform mit dem besten Preis-Leistungs-Verhältnis für alle Workloads.

## Welchen Nutzen bietet die hybride Iceberg Lakehouse Lösung von Dremio und NetApp den Kunden?

- **Verbessertes Datenmanagement und Zugänglichkeit:** Dremio ist bekannt für seine Data Lakehouse-Plattform, mit der Unternehmen Daten direkt aus ihren Data Lakes mit hoher Geschwindigkeit abfragen können. NetApp hingegen ist ein führender Anbieter von Cloud-Datenservices und Storage-Lösungen. Dieses gemeinsame Angebot bietet Kunden eine umfassende Lösung zum Speichern, Managen, Abrufen und Analysieren ihrer Unternehmensdaten auf effiziente und effiziente Weise.
- **Performance-Optimierung:** Mit der Expertise von NetApp in den Bereichen Datenspeicher und Dremio in der Data Processing- und Datenoptimierung bietet die Partnerschaft eine Lösung, die die Performance von Datenvorgängen verbessert, Latenz reduziert und schnellere Geschäftseinblicke ermöglicht. Dremio hat sogar der internen IT-Analyseinfrastruktur von NetApp Performance-Vorteile gebracht.
- **Skalierbarkeit:** Sowohl Dremio als auch NetApp bieten eine skalierbare Lösung. Diese gemeinsame Lösung bietet Kunden hochskalierbare Storage-, Datenmanagement- und Analyseumgebungen. In einer hybriden Iceberg Lakehouse-Umgebung bietet die Dremio SQL-Abfrageengine in Kombination mit NetApp StorageGRID eine unvergleichliche Skalierbarkeit, Parallelität und Abfrageleistung, die den analytischen Anforderungen jedes Unternehmens gerecht wird.
- **Datensicherheit und Governance:** Beide Unternehmen haben einen starken Fokus auf Datensicherheit und Governance. Gemeinsam bieten sie robuste Sicherheits- und Data Governance-Funktionen, die den Schutz der Daten und die Einhaltung der Data Governance-Anforderungen sicherstellen. Funktionen wie rollenbasierte und fein abgestufte Zugriffskontrollen, umfassende Audits, End-to-End-Datenlieferungsline, einheitliches Identitätsmanagement und SSO mit einem umfassenden Compliance- und Sicherheits-Framework sorgen dafür, dass die analytischen Datenumgebungen in Unternehmen sicher und geregelt sind.
- **Kosteneffizienz:** Durch die Integration der Data Lake Engine von Dremio in die Storage-Lösungen von NetApp können Kunden die Kosten für Datenmanagement und Datenverschiebung senken. Unternehmen sind außerdem in der Lage, von bestehenden Data Lake-Umgebungen auf eine modernere Lakehouse-Lösung mit NetApp und Dremio umzusteigen. Diese Hybrid-Lösung von Iceberg Lakehouse bietet eine ultraschnelle Abfrageperformance und marktführende Abfrageparallelität, wodurch die TCO gesenkt und die Zeit bis zu geschäftlichen Einblicken verkürzt wird.

## Technologieanforderungen Erfüllt

Die unten aufgeführten Hardware- und Softwarekonfigurationen wurden für in diesem Dokument durchgeführte Validierungen verwendet. Diese Konfigurationen dienen als Richtlinie zur Einrichtung Ihrer Umgebung. Beachten Sie jedoch, dass die einzelnen Komponenten je nach Kundenanforderungen variieren können.

## Hardwareanforderungen

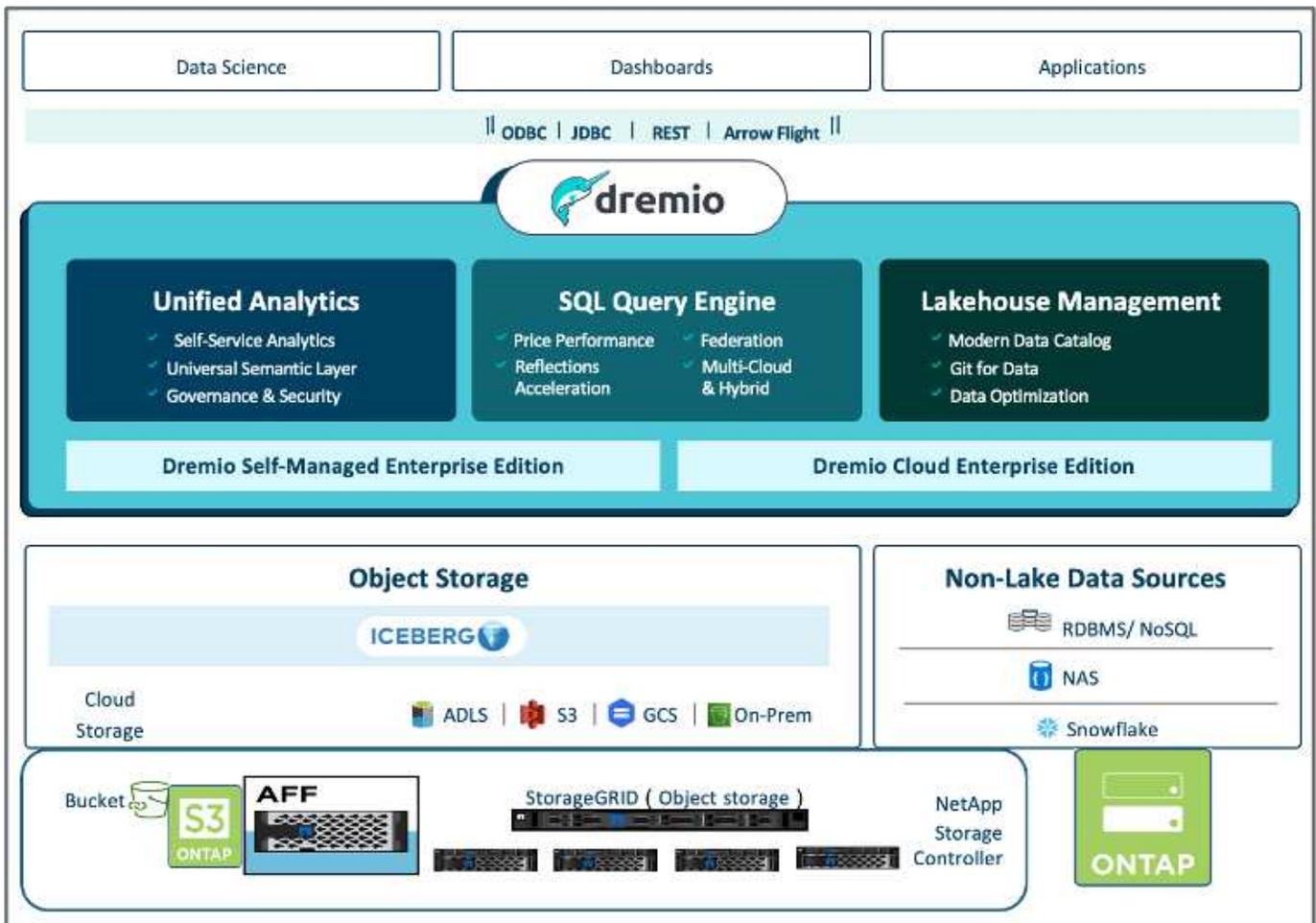
Trennt	Details
NetApp AFF Storage-Array HA-Paar	<ul style="list-style-type: none"><li>• A800</li><li>• ONTAP 9.14.1</li><li>• 48 X 3,49 TB SSD-NVM</li><li>• Zwei S3-Buckets: Dremio Metadaten und Kundendaten</li></ul>
4 X FUJITSU PRIMERGY RX2540 M4	<ul style="list-style-type: none"><li>• 64 CPUs</li><li>• Intel® Xeon® Gold 6142 CPU @ 2,60 GHz</li><li>• 256 GM physischer Speicher</li><li>• 1 x 100-GbE-Netzwerk-Port</li></ul>
Netzwerkbetrieb	<ul style="list-style-type: none"><li>• 100 GbE</li></ul>
StorageGRID	* 1 x SG100, 3xSGF6024 * 3 x 24 x 7,68 TB * zwei S3-Buckets: Dremio-Metadaten und Kundendaten.

## Softwareanforderungen

Software	Details
Dremio	<ul style="list-style-type: none"><li>• Version: 25.0.3-202405170357270647-d2042e1b</li><li>• Enterprise Edition</li></ul>
Lokal	<ul style="list-style-type: none"><li>• Dremio Cluster mit 5 Nodes</li><li>• 1 Master-Koordinator und 4 Ausführende</li></ul>

## Bereitstellungsverfahren

Bei dieser Validierung der Referenzarchitektur wurde eine Dremio-Konfiguration verwendet, die aus einem Coordinator und vier Executoren besteht



## NetApp Einrichtung

- Initialisierung des Speichersystems
- Erstellung von Storage Virtual Machines (SVM)
- Zuweisung logischer Netzwerkschnittstellen
- NFS-, S3-Konfiguration und -Lizenzierung

Folgen Sie bitte den Schritten unten für NFS (Network File System): 1. Erstellen Sie ein Flex-Group-Volume für NFSv4 oder NFSv3. In unserer Einrichtung für diese Validierung haben wir 48 SSDs, 1 SSD für das Root-Volume des Controllers und 47 SSDs für NFSv4] verwendet. Überprüfen Sie, ob die NFS-Exportrichtlinie für das Flex Group Volume Lese-/Schreibberechtigungen für das Dremio Server-Netzwerk hat.

1. Erstellen Sie auf allen Dremio Servern einen Ordner und mounten Sie das Flex Group Volume über eine logische Schnittstelle (LIF) auf jedem Dremio Server in diesen Ordner.

Folgen Sie bitte den Schritten unten für S3 (Simple Storage Service):

1. Richten Sie einen Object-Store-Server mit aktiviertem HTTP ein, und der Admin-Status wird mit dem Befehl „vserver Object-Store-Server create“ auf „up“ gesetzt. Sie haben die Möglichkeit, HTTPS zu aktivieren und einen benutzerdefinierten Listener-Port festzulegen.
2. Erstellen Sie mit dem Befehl „vserver object-Store-Server user create -user <username>“ einen Object-Store-Server-Benutzer.
3. Um den Zugriffsschlüssel und den geheimen Schlüssel zu erhalten, können Sie den folgenden Befehl

ausführen: "Set diag; vserver object-Store-Server user show -user <username>". Diese Schlüssel werden jedoch während der Benutzererstellung bereitgestellt oder können über REST-API-Aufrufe abgerufen werden.

4. Erstellen Sie eine Object-Store-Server-Gruppe mit dem in Schritt 2 erstellten Benutzer und gewähren Sie Zugriff. In diesem Beispiel haben wir „FullAccess“ bereitgestellt.
5. Erstellen Sie zwei S3-Buckets, indem Sie den entsprechenden Typ auf „S3“ setzen. Eine für die Dremio-Konfiguration und eine für Kundendaten.

## Einrichtung von ZooKeeper

Sie können die von Dremio bereitgestellte Zookeeper-Konfiguration verwenden. In dieser Validierung haben wir einen separaten Zookeeper verwendet. Wir haben die in diesem Weblink genannten Schritte befolgt <https://medium.com/@ahmetfurkandemir/distributed-hadoop-cluster-1-spark-with-all-dependencies-03c8ec616166>

## Dremio-Einrichtung

Wir folgten diesem Weblink, um Dremio via Tar Ball zu installieren.

1. Erstellen Sie eine Dremio-Gruppe.

```
sudo groupadd -r dremio
```

2. Erstellen Sie einen dremio-Benutzer.

```
sudo useradd -r -g dremio -d /var/lib/dremio -s /sbin/nologin dremio
```

3. Dremio-Verzeichnisse erstellen.

```
sudo mkdir /opt/dremio
sudo mkdir /var/run/dremio && sudo chown dremio:dremio /var/run/dremio
sudo mkdir /var/log/dremio && sudo chown dremio:dremio /var/log/dremio
sudo mkdir /var/lib/dremio && sudo chown dremio:dremio /var/lib/dremio
```

4. Laden Sie die tar-Datei von herunter <https://download.dremio.com/community-server/>
5. Entpacken Sie Dremio in das Verzeichnis /opt/dremio.

```
sudo tar xvf dremio-enterprise-25.0.3-202405170357270647-d2042e1b.tar.gz
-C /opt/dremio --strip-components=1
```

6. Erstellen Sie einen symbolischen Link für den Konfigurationsordner.

```
sudo ln -s /opt/dremio/conf /etc/dremio
```

## 7. Richten Sie Ihre Servicekonfiguration ein (systemd Setup).

- a. Kopieren Sie die Unit-Datei für den dremio-Daemon von /opt/dremio/share/dremio.service nach /etc/systemd/System/dremio.service.
- b. System neu starten

```
sudo systemctl daemon-reload
```

- c. Aktivieren Sie dremio, um beim Start zu starten.

```
sudo systemctl enable dremio
```

## 8. Dremio auf Koordinator konfigurieren. Weitere Informationen finden Sie unter Dremio Configuration

- a. Dremio.conf

```
root@hadoopmaster:/usr/src/tpcds# cat /opt/dremio/conf/dremio.conf

paths: {
  # the local path for dremio to store data.
  local: ${DREMIO_HOME}"/dremiocache"

  # the distributed path Dremio data including job results,
  downloads, uploads, etc
  #dist: "hdfs://hadoopmaster:9000/dremiocache"
  dist: "dremioS3:///dremioconf"
}

services: {
  coordinator.enabled: true,
  coordinator.master.enabled: true,
  executor.enabled: false,
  flight.use_session_service: false
}

zookeeper: "10.63.150.130:2181,10.63.150.153:2181,10.63.150.151:2181"
services.coordinator.master.embedded-zookeeper.enabled: false
root@hadoopmaster:/usr/src/tpcds#
```

- b. core-site.xml

```
root@hadoopmaster:/usr/src/tpcds# cat /opt/dremio/conf/core-site.xml
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
```

Licensed under the Apache License, Version 2.0 (the "License");  
you may not use this file except in compliance with the License.  
You may obtain a copy of the License at

<http://www.apache.org/licenses/LICENSE-2.0>

Unless required by applicable law or agreed to in writing, software  
distributed under the License is distributed on an "AS IS" BASIS,  
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or  
implied.

See the License for the specific language governing permissions and  
limitations under the License. See accompanying LICENSE file.

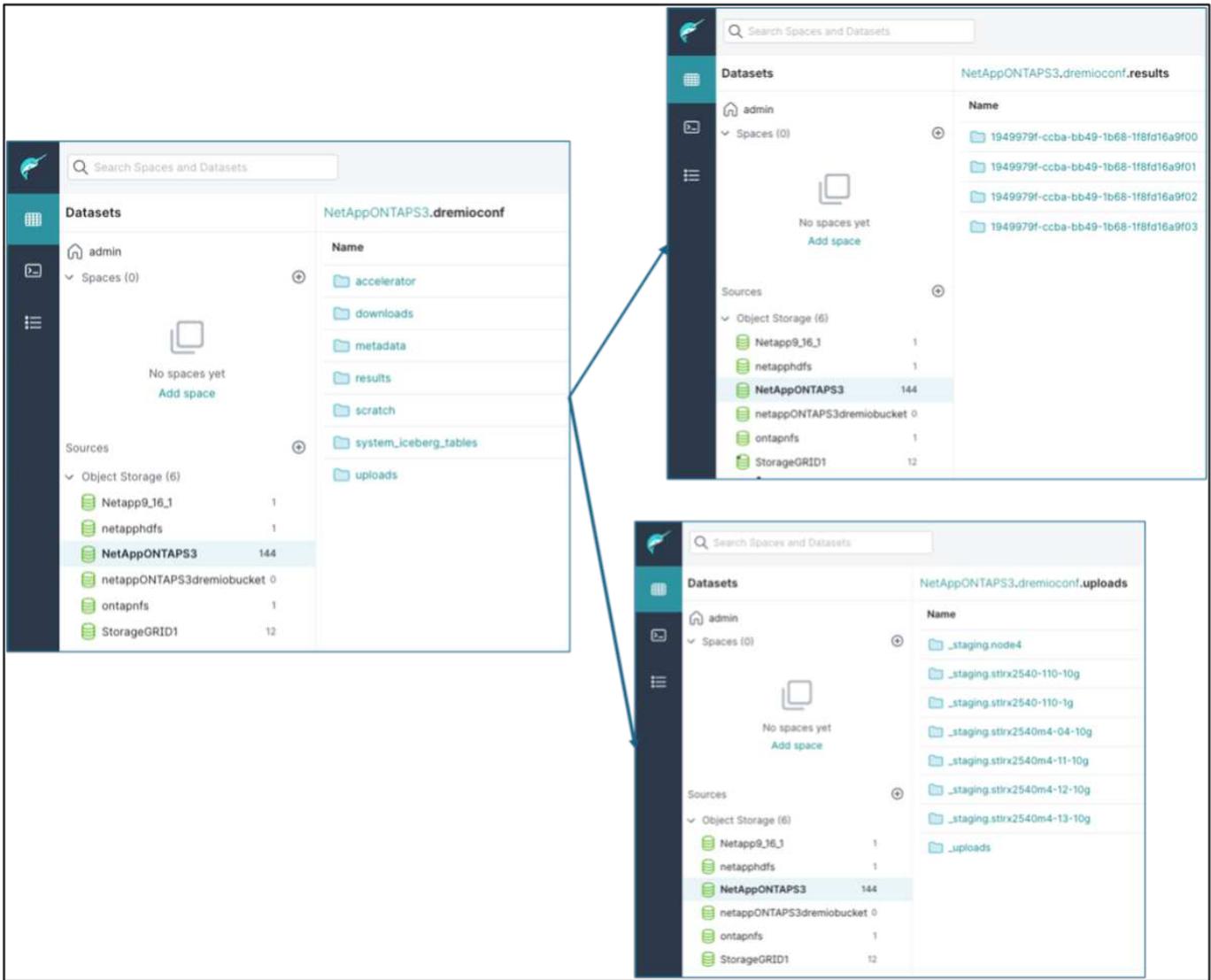
-->

<!-- Put site-specific property overrides in this file. -->

```
<configuration>
  <property>
    <name>fs.dremioS3.impl</name>
    <value>com.dremio.plugins.s3.store.S3FileSystem</value>
  </property>
  <property>
    <name>fs.s3a.access.key</name>
    <value>24G4C1316APP2BIPDE5S</value>
  </property>
  <property>
    <name>fs.s3a.endpoint</name>
    <value>10.63.150.69:80</value>
  </property>
  <property>
    <name>fs.s3a.secret.key</name>
    <value>Zd28p43rgZaU44PX_ftT279z9nt4jBSro97j87Bx</value>
  </property>
  <property>
    <name>fs.s3a.aws.credentials.provider</name>
    <description>The credential provider type.</description>
    <value>org.apache.hadoop.fs.s3a.SimpleAWSCredentialsProvider</value>
  </property>
  <property>
    <name>fs.s3a.path.style.access</name>
    <value>>false</value>
  </property>
  <property>
    <name>hadoop.proxyuser.dremio.hosts</name>
    <value>*</value>
```

```
</property>
<property>
  <name>hadoop.proxyuser.dremio.groups</name>
  <value>*</value>
</property>
<property>
  <name>hadoop.proxyuser.dremio.users</name>
  <value>*</value>
</property>
<property>
  <name>dremio.s3.compat</name>
  <description>Value has to be set to true.</description>
  <value>>true</value>
</property>
<property>
  <name>fs.s3a.connection.ssl.enabled</name>
  <description>Value can either be true or false, set to true
to use SSL with a secure Minio server.</description>
  <value>>false</value>
</property>
</configuration>
root@hadoopmaster:/usr/src/tpcds#
```

9. Die Dremio Konfiguration wird im NetApp Objektspeicher gespeichert. In unserer Validierung befindet sich der Bucket „dremioconf“ in einem ONTAP S3-Bucket. Das folgende Bild zeigt einige Details aus dem Ordner „scratch“ und „Uploads“ des S3-Buckets „dremioconf“.



1. Konfigurieren Sie Dremio auf Ausführenden. In unserem Setup haben wir 3 Ausführende.
  - a. Dremio.conf

```

paths: {
  # the local path for dremio to store data.
  local: "${DREMIO_HOME}"/dremiocache"

  # the distributed path Dremio data including job results,
  downloads, uploads, etc
  #dist: "hdfs://hadoopmaster:9000/dremiocache"
  dist: "dremioS3:///dremioconf"
}

services: {
  coordinator.enabled: false,
  coordinator.master.enabled: false,
  executor.enabled: true,
  flight.use_session_service: true
}

zookeeper: "10.63.150.130:2181,10.63.150.153:2181,10.63.150.151:2181"
services.coordinator.master.embedded-zookeeper.enabled: false

```

b. core-site.xml – wie bei der Coordinator-Konfiguration.



NetApp empfiehlt StorageGRID als primäre Objekt-Storage-Lösung für Datalake- und Lakehouse-Umgebungen. Darüber hinaus wird NetApp ONTAP für die Datei-/Objekt-Dualität eingesetzt. Im Rahmen dieses Dokuments haben wir Tests zu ONTAP S3 als Reaktion auf eine Kundenanfrage durchgeführt und es funktioniert erfolgreich als Datenquelle.

## Einrichtung mehrerer Quellen

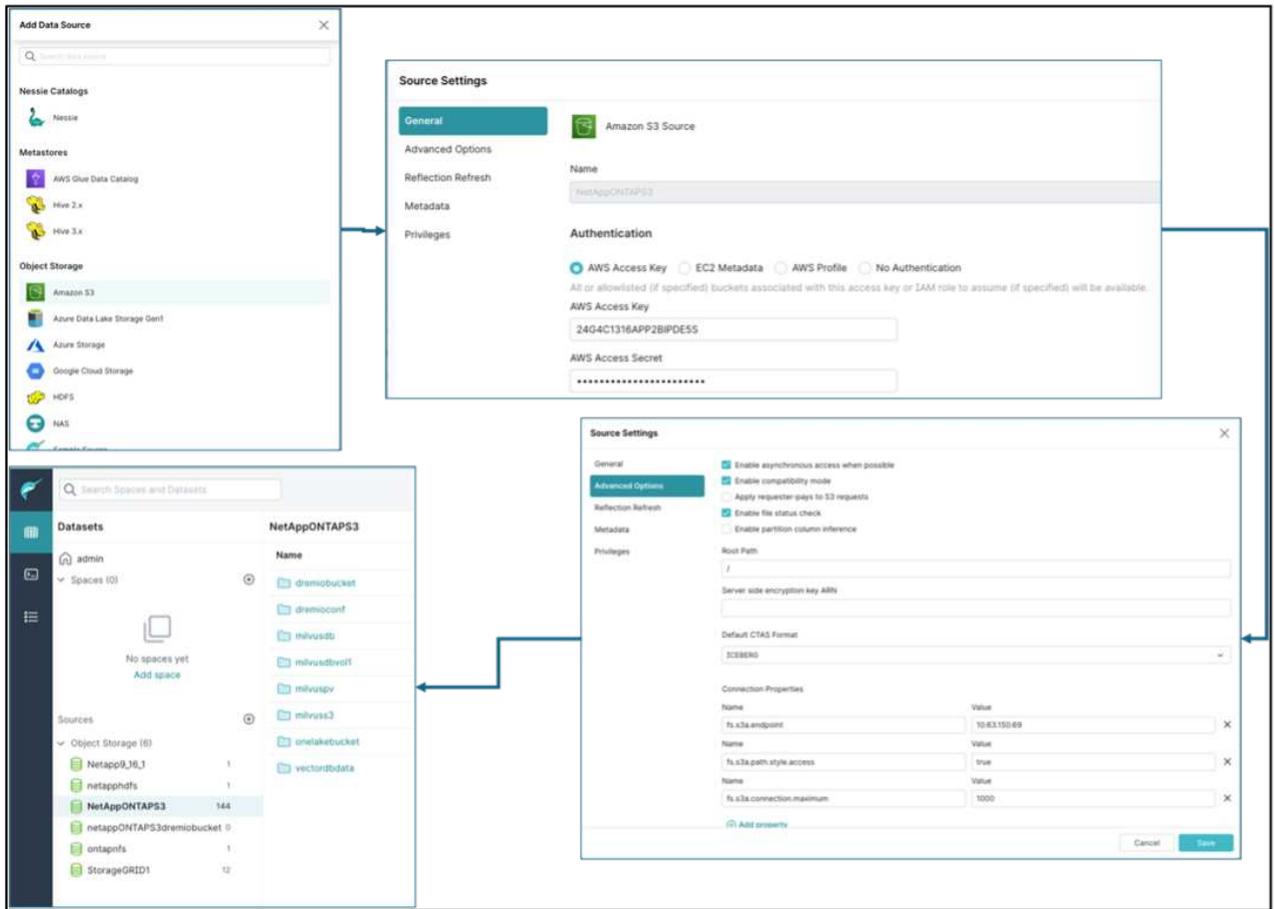
1. ONTAP S3 und StorageGRID als S3-Quelle in Dremio konfigurieren.
  - a. Dremio Dashboard → Datasets → sources → add source.
  - b. Aktualisieren Sie im allgemeinen Abschnitt den AWS-Zugriff und den geheimen Schlüssel
  - c. Aktualisieren Sie in der erweiterten Option Kompatibilitätsmodus aktivieren die Verbindungseigenschaften mit den folgenden Details. Die Endpunkt-IP/Name des NetApp Storage Controllers entweder aus ONTAP S3 oder StorageGRID.

```

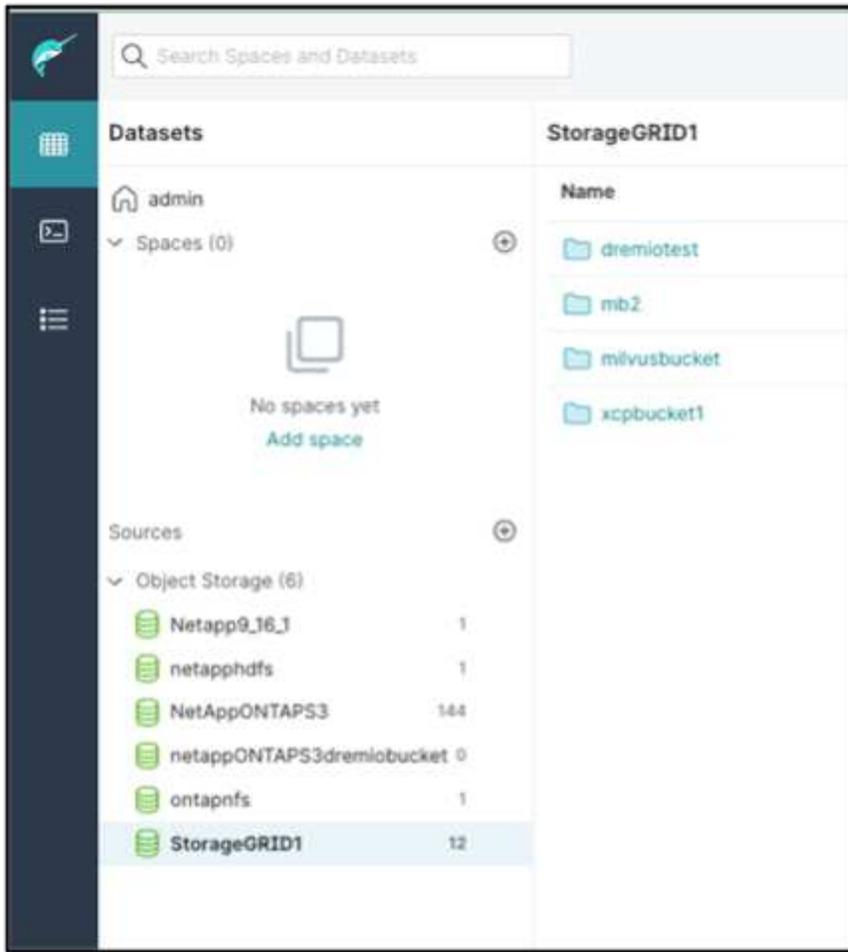
fs.s3a.endpoint = 10.63.150.69
fs.s3a.path.style.access = true
fs.s3a.connection.maximum=1000

```

- d. Lokales Caching aktivieren, wenn möglich, Max. Prozentsatz des verfügbaren Gesamtcache, wenn möglich zu verwenden = 100
- e. Zeigen Sie anschließend die Liste der Buckets aus dem NetApp Objekt-Storage an.



f. Beispielsicht von StorageGRID Bucket-Details



2. Konfigurieren Sie NAS ( speziell NFS ) als Quelle in Dremio.
  - a. Dremio Dashboard → Datasets → sources → add source.
  - b. Geben Sie im Abschnitt „Allgemein“ den Namen und den NFS-Mount-Pfad ein. Stellen Sie sicher, dass der NFS-Mount-Pfad auf dem gleichen Ordner auf allen Knoten im Dremio Cluster gemountet ist.

### Add Data Source

Search data source

**Nessie Catalogs**

- Nessie

**Metastores**

- AWS Glue Data Catalog
- Hive 2.x
- Hive 3.x

**Object Storage**

- Amazon S3
- Azure Data Lake Storage Gen1
- Azure Storage
- Google Cloud Storage
- HDFS
- NAS**

### New NAS Source

**General**

Advanced Options

Reflection Refresh

Metadata

Privileges

**NAS Source**

Name

ontapnfs

**Connection**

Mount Path ⓘ

/dremionfsdata

Search Spaces and Datasets

**Datasets**

admin

Spaces (0)

No spaces yet  
[Add space](#)

**Sources**

Source Name	Count
Netapp9_16_1	1
netapphdfs	1
NetAppONTAPS3	144
netappONTAPS3dremiobucket	0
<b>ontapnfs</b>	<b>1</b>
StorageGRID1	12

**ontapnfs**

Name

- csvfile\_from\_dataset
- results

+

```

root@hadoopmaster:~# for i in hadoopmaster hadoopnode1 hadoopnode2
hadoopnode3 hadoopnode4; do ssh $i "date;hostname;du -hs
/opt/dremio/data/spill/ ; df -h //dremionfsdata "; done
Fri Sep 13 04:13:19 PM UTC 2024
hadoopmaster
du: cannot access '/opt/dremio/data/spill/': No such file or directory
Filesystem                Size      Used Avail Use% Mounted on
10.63.150.69:/dremionfsdata 2.1T    921M   2.0T   1% /dremionfsdata
Fri Sep 13 04:13:19 PM UTC 2024
hadoopnode1
12K /opt/dremio/data/spill/
Filesystem                Size      Used Avail Use% Mounted on
10.63.150.69:/dremionfsdata 2.1T    921M   2.0T   1% /dremionfsdata
Fri Sep 13 04:13:19 PM UTC 2024
hadoopnode2
12K /opt/dremio/data/spill/
Filesystem                Size      Used Avail Use% Mounted on
10.63.150.69:/dremionfsdata 2.1T    921M   2.0T   1% /dremionfsdata
Fri Sep 13 16:13:20 UTC 2024
hadoopnode3
16K /opt/dremio/data/spill/
Filesystem                Size      Used Avail Use% Mounted on
10.63.150.69:/dremionfsdata 2.1T    921M   2.0T   1% /dremionfsdata
Fri Sep 13 04:13:21 PM UTC 2024
node4
12K /opt/dremio/data/spill/
Filesystem                Size      Used Avail Use% Mounted on
10.63.150.69:/dremionfsdata 2.1T    921M   2.0T   1% /dremionfsdata
root@hadoopmaster:~#

```

## Übersicht zur Lösungsüberprüfung

In diesem Abschnitt haben wir SQL-Testabfragen aus mehreren Quellen ausgeführt, um die Funktionalität zu überprüfen, den Spillover auf den NetApp-Speicher zu testen und zu überprüfen.

### SQL-Abfrage zu Objekt-Storage

1. Stellen Sie den Speicher auf 250 GB pro Server in dremio.env ein

```

root@hadoopmaster:~# for i in hadoopmaster hadoopnode1 hadoopnode2
hadoopnode3 hadoopnode4; do ssh $i "hostname; grep -i
DREMIO_MAX_MEMORY_SIZE_MB /opt/dremio/conf/dremio-env; cat /proc/meminfo
| grep -i memtotal"; done
hadoopmaster
#DREMIO_MAX_MEMORY_SIZE_MB=120000
DREMIO_MAX_MEMORY_SIZE_MB=250000
MemTotal:          263515760 kB
hadoopnode1
#DREMIO_MAX_MEMORY_SIZE_MB=120000
DREMIO_MAX_MEMORY_SIZE_MB=250000
MemTotal:          263515860 kB
hadoopnode2
#DREMIO_MAX_MEMORY_SIZE_MB=120000
DREMIO_MAX_MEMORY_SIZE_MB=250000
MemTotal:          263515864 kB
hadoopnode3
#DREMIO_MAX_MEMORY_SIZE_MB=120000
DREMIO_MAX_MEMORY_SIZE_MB=250000
MemTotal:          264004556 kB
node4
#DREMIO_MAX_MEMORY_SIZE_MB=120000
DREMIO_MAX_MEMORY_SIZE_MB=250000
MemTotal:          263515484 kB
root@hadoopmaster:~#

```

- Überprüfen Sie den Speicherort für die Verschüttung ( $\{DREMIO\_HOME\}/dremiocache$ ) in der Datei `dremio.conf` und den Speicherdetails.

```

paths: {
  # the local path for dremio to store data.
  local: "${DREMIO_HOME}"/dremiocache"

  # the distributed path Dremio data including job results, downloads,
  uploads, etc
  #dist: "hdfs://hadoopmaster:9000/dremiocache"
  dist: "dremioS3:///dremioconf"
}

services: {
  coordinator.enabled: true,
  coordinator.master.enabled: true,
  executor.enabled: false,
  flight.use_session_service: false
}

zookeeper: "10.63.150.130:2181,10.63.150.153:2181,10.63.150.151:2181"
services.coordinator.master.embedded-zookeeper.enabled: false

```

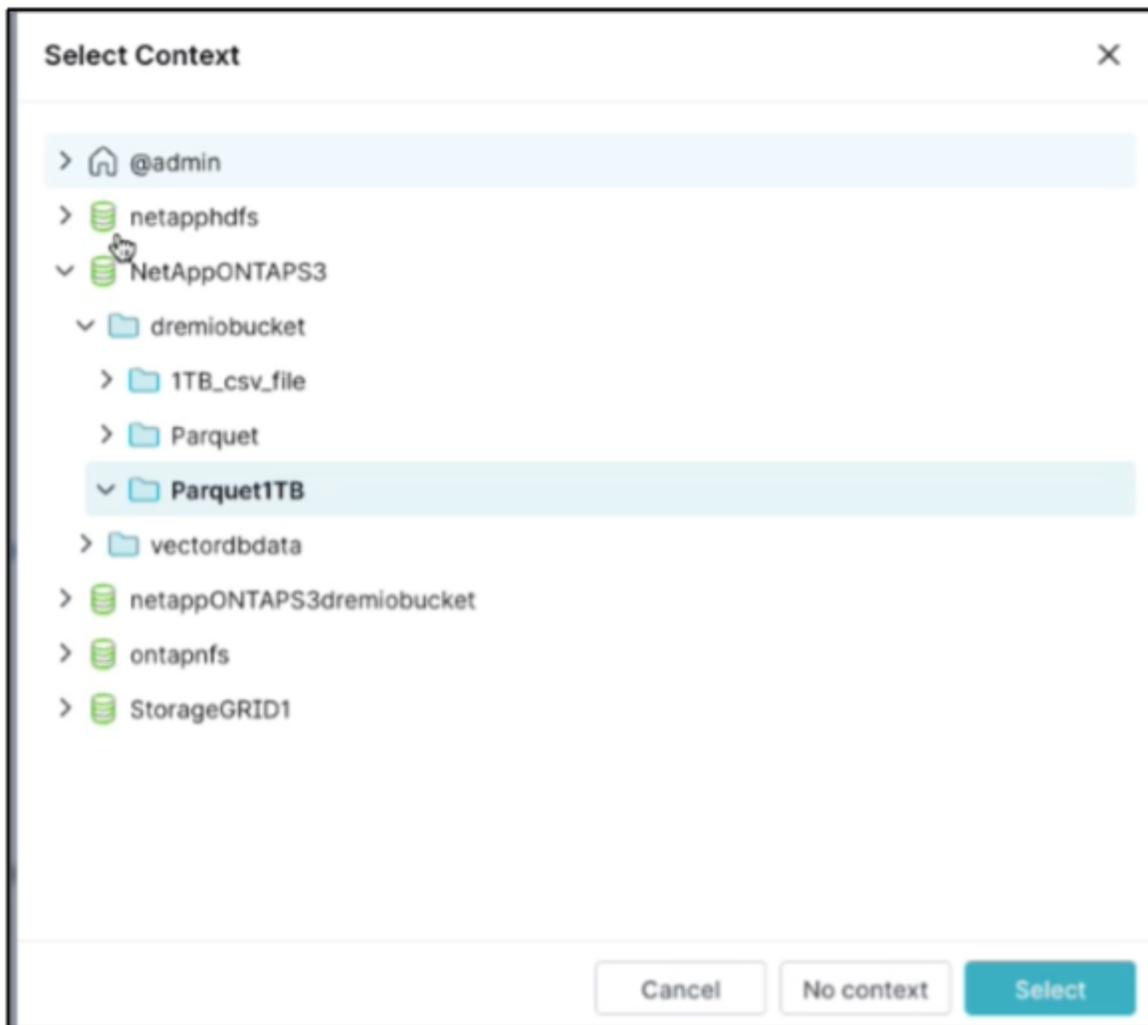
### 3. Verweisen Sie den Dremio-Überlauf auf den NetApp-NFS-Storage

```

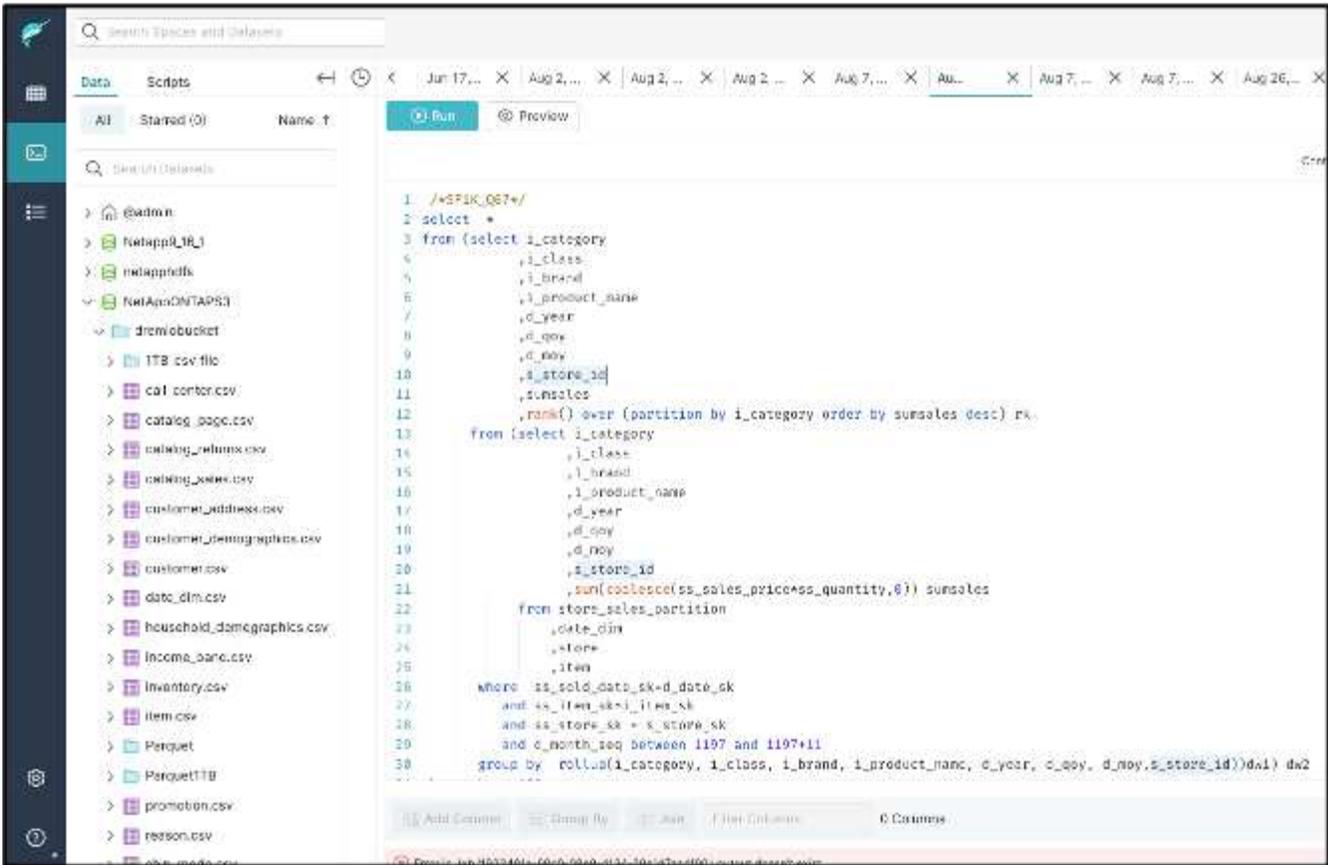
root@hadoopnode1:~# ls -ltrh /dremiocache
total 4.0K
drwx----- 3 nobody nogroup 4.0K Sep 13 16:00 spilling_stlrx2540m4-12-
10g_45678
root@hadoopnode1:~# ls -ltrh /opt/dremio/dremiocache/
total 8.0K
drwxr-xr-x 3 dremio dremio 4.0K Aug 22 18:19 spill_old
drwxr-xr-x 4 dremio dremio 4.0K Aug 22 18:19 cm
lrwxrwxrwx 1 root root 12 Aug 22 19:03 spill -> /dremiocache
root@hadoopnode1:~# ls -ltrh /dremiocache
total 4.0K
drwx----- 3 nobody nogroup 4.0K Sep 13 16:00 spilling_stlrx2540m4-12-
10g_45678
root@hadoopnode1:~# df -h /dremiocache
Filesystem                Size      Used Avail Use% Mounted on
10.63.150.159:/dremiocache_hadoopnode1  2.1T    209M   2.0T   1% /dremiocache
root@hadoopnode1:~#

```

### 4. Wählen Sie den Kontext aus. In unserem Test haben wir den Test gegen TPCDS generierte Parquetdateien in ONTAP S3 durchgeführt. Dremio Dashboard → SQL Runner → Kontext → NetAppONTAPS3→Parquet1TB



1. Führen Sie den TPC-DS query67 vom Dremio Dashboard aus



1. Überprüfen Sie, ob der Job auf allen Ausführern ausgeführt wird. Dremio Dashboard → Jobs → <jobid> → RAW Profil → select EXTERNAL\_SORT → Hostname

Raw Profile

04-xx-04 - FILTER

04-xx-05 - WINDOW

04-xx-06 - EXTERNAL\_SORT

Thread	Setup Time	Process Time	Wait Time	Max Batches	Max Records	Peak Memory	Hostname	Record Processing Rate	Operator State	Last Schedule Time
04-00-06	0.000s	0.000s	0.000s	0	0	128KB	str2540-110-10g	0	CAN_CONSUME	16:35:54
04-01-06	0.000s	0.000s	0.000s	0	0	128KB	str2540m4-04-10g	0	CAN_CONSUME	16:35:54
04-02-06	0.000s	0.000s	0.000s	0	0	128KB	str2540m4-12-10g	0	CAN_CONSUME	16:35:54
04-03-06	0.017s	0.000s	0.000s	0	0	128KB	str2540m4-13-10g	0	CAN_CONSUME	16:35:54
04-04-06	0.000s	0.000s	0.000s	0	0	128KB	str2540-110-10g	0	CAN_CONSUME	16:35:54
04-05-06	0.000s	0.000s	0.000s	0	0	128KB	str2540m4-04-10g	0	CAN_CONSUME	16:35:54
04-06-06	0.027s	0.000s	0.000s	0	0	128KB	str2540m4-12-10g	0	CAN_CONSUME	16:35:54
04-07-06	0.000s	0.000s	0.000s	0	0	128KB	str2540m4-13-10g	0	CAN_CONSUME	16:35:54

1. Wenn die SQL-Abfrage ausgeführt wird, können Sie den Split-Ordner auf Daten-Caching in NetApp-Speicher-Controller überprüfen.

```

root@hadoopnode1:~# ls -ltrh /dremlcache
total 4.0K
drwx----- 3 nobody nogroup 4.0K Sep 13 16:00 spilling_stlrx2540m4-12-10g_45678
root@hadoopnode1:~# ls -ltrh /dremlcache/spilling_stlrx2540m4-12-10g_45678/
total 4.0K
drwxr-xr-x 2 root daemon 4.0K Sep 13 16:23 1726243167416

```

2. Die SQL-Abfrage wurde mit Überlauf abgeschlossen

Job ID	User	Dataset	Query Type	Queue	Start Time	Duration	SQL
a0a5-9dab-2b16-e2ec24459900-19335115	admin	store_sales_partition	UI (run)	High Cost User Q...	08/26/2024, 12:35:53	00:08:25	/*SF1K_Q67*/ select * from (select i_category ,i_class ,i_brand ,i_product_name ,d_year ,d_m...
19383201-5cd9-0a48-1e38-e2f5a4146900	admin	store_sales_partition	JDBC Client	High Cost User Q...	08/22/2024, 19:42:54	00:08:23	/*SF1K_Q67*/ select * from (select i_category ,i_class ,i_brand ,i_product_name ,d_year ,d_m...
19384af3-285d-a01c-5277-46b8810ss1200	admin	store_sales_partition	JDBC Client	High Cost User Q...	08/22/2024, 18:00:44	00:08:23	/*SF1K_Q67*/ select * from (select i_category ,i_class ,i_brand ,i_product_name ,d_year ,d_m...
1938650f-0f9a-a285-9ea3-073aaa3c7a00	admin	store_sales_partition	JDBC Client	High Cost User Q...	08/22/2024, 16:09:20	00:08:26	/*SF1K_Q67*/ select * from (select i_category ,i_class ,i_brand ,i_product_name ,d_year ,d_m...
19387983-2031-1d4f-ca9e-576c287ba000	admin	store_sales_partition	UI (run)	High Cost User Q...	08/22/2024, 14:42:04	00:07:26	/*SF1K_Q67*/ select * from (select i_category ,i_class ,i_brand ,i_product_name ,d_year ,d_m...
19387a04-3ac3-34bd-13e5-d7f53894a000	admin	store_sales_partition	UI (run)	High Cost User Q...	08/22/2024, 14:22:51	00:07:48	/*SF1K_Q67*/ select * from (select i_category ,i_class ,i_brand ,i_product_name ,d_year ,d_m...

3. Zusammenfassung des Auftrags.

Jobs > 19335115-a0a5-9dab-2b16-e2ec24459900
Overview
SQL

### Summary

Status: COMPLETED

Total Memory: 287.16 GB

CPU Used: 02h:18m:52s

Query Type: UI (run)

Start Time: 08/26/2024 12:35:53

Duration: 08m:25s

Wait on Client: <1s

User: admin

Queue: High Cost User Queries

Input: 21.32 GB / 563.2M Rows

Output: 6.92 KB / 100 Rows

---

**Total Execution Time** 08m:25s (100%)

Pending	2ms (0.001%)
Metadata Retrieval	22ms (0.00%)
Planning	140ms (0.03%)
Queued	30ms (0.01%)
Execution Planning	116ms (0.02%)
Starting	569ms (0.11%)
Running	8m:24s (99.83%)

### Submitted SQL

```

1 /*SF1K_Q67*/
2 select *
3 from (select i_category
4         ,i_class
5         ,i_brand
6         ,i_product_name
7         ,d_year
8         ,d_qoy
9         ,d_moy

```

### Queried Datasets

- store\_sales\_partition  
NetAppONTAPS3.dremiobucket.Parquet1TB
- date\_dim  
NetAppONTAPS3.dremiobucket.Parquet1TB
- store  
NetAppONTAPS3.dremiobucket.Parquet1TB

Show more >

### Scans

- store\_sales\_partition
- date\_dim
- store
- item

4. Überprüfen Sie die Größe der verschütteten Daten

## EXTERNAL\_SORT 04-06



<b>Runtime</b>	1.68m (100%)
Startup	49.09ms (0.05%)
Processing	39.62s (39.36%)
IO Wait	1.02m (60.6%)

### Overview/Main

Batches Processed:	104333
Records Processed:	387.6M
Peak Memory:	199 MB
Bytes Sent:	44 GB
Number of Threads:	180

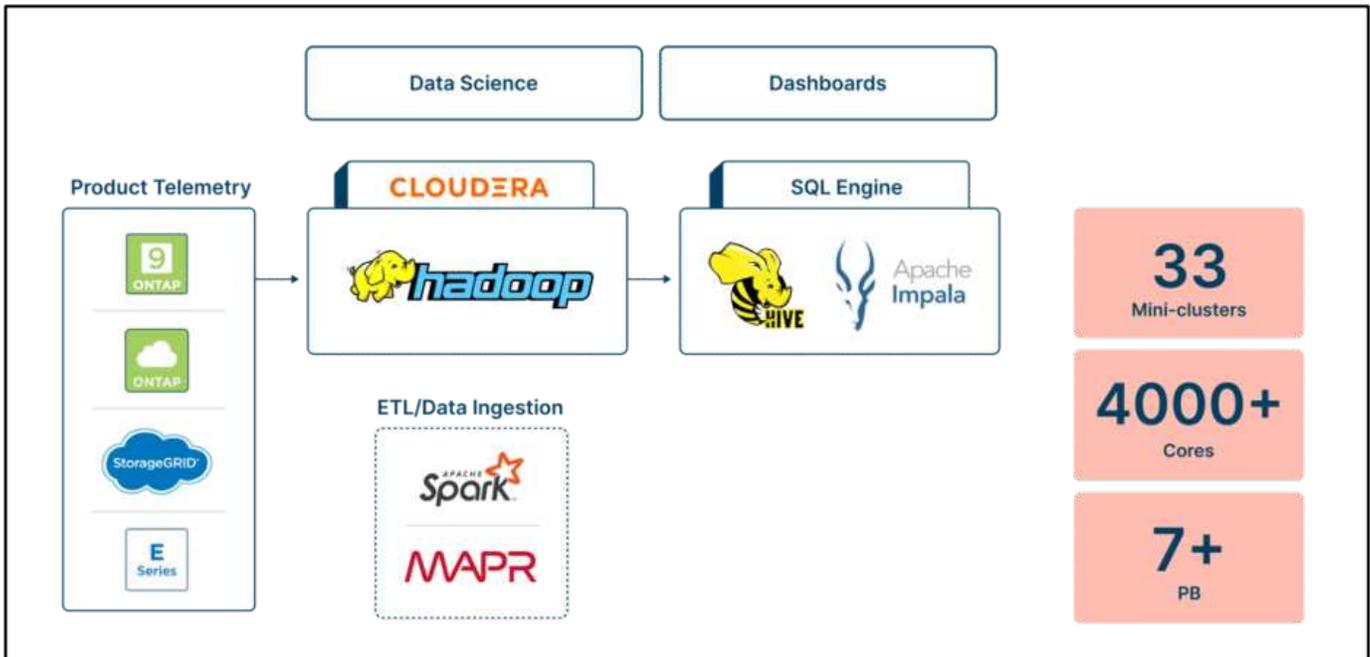
### Operator Statistics

Merge Time Nanos:	0ns
Spill Count:	360
Spill Time Nanos:	37.68m
Total Spilled Data Size:	<b>20,339,702,765</b>
Batches Spilled:	97,854

Dasselbe Verfahren gilt für NAS- und StorageGRID-Objektspeicher.

## Anwendungsfälle Von Kunden

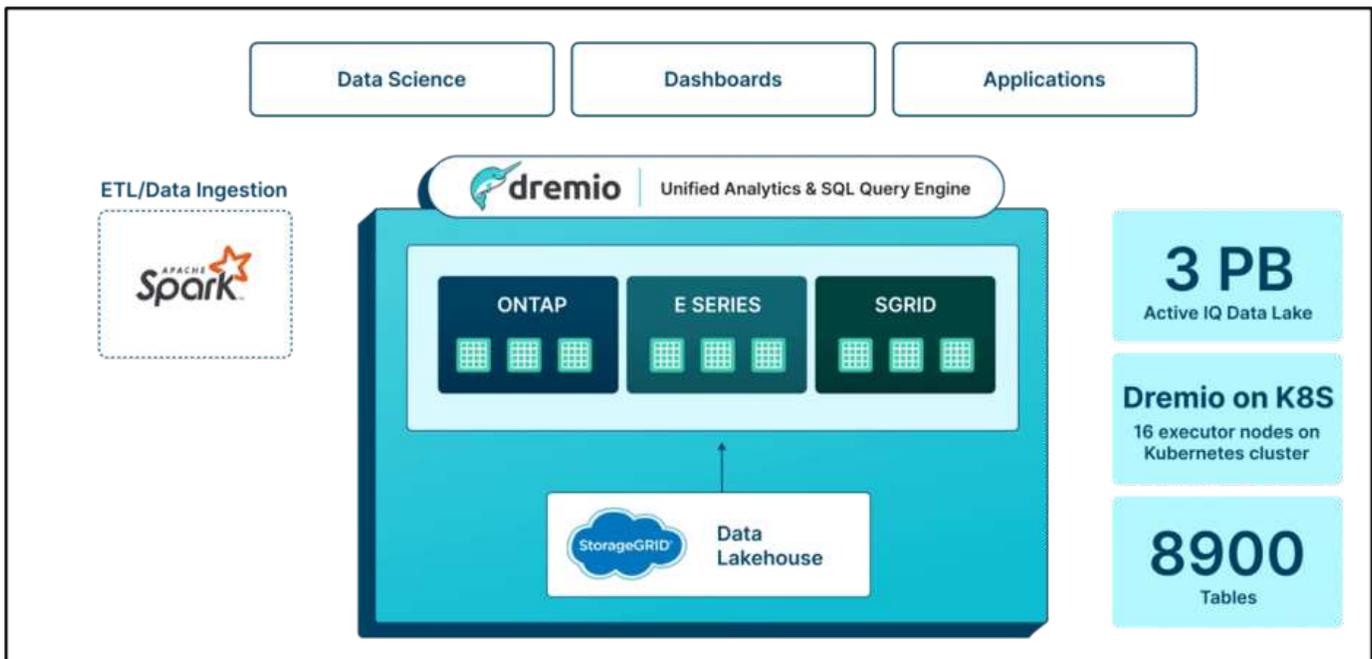
### Anwendungsfall für NetApp ActiveIQ



**Herausforderung:** Die eigene interne Active IQ-Lösung von NetApp, die ursprünglich für die Unterstützung zahlreicher Anwendungsfälle konzipiert wurde, hat sich zu einem umfassenden Angebot sowohl für interne Benutzer als auch für Kunden entwickelt. Die zugrunde liegende Hadoop/MapR-basierte Back-End-Infrastruktur stellte jedoch aufgrund des schnellen Datenwachstums und der Notwendigkeit eines effizienten Datenzugriffs vor Herausforderungen im Hinblick auf die Kosten und die Performance. Die Storage-Skalierung erforderte das Hinzufügen unnötiger Computing-Ressourcen und führte zu höheren Kosten.

Außerdem war das Management des Hadoop Clusters zeitaufwändig und erforderte spezielle Fachkenntnisse. Probleme mit der Daten-Performance und dem Datenmanagement verkomplizieren die Situation noch weiter: Abfragen benötigen durchschnittlich 45 Minuten und aufgrund von Fehlkonfigurationen Hungersnöte. Um diese Herausforderungen zu bewältigen, suchte NetApp nach einer Alternative zur vorhandenen alten Hadoop Umgebung und stellte fest, dass eine neue moderne Lösung auf Basis von Dremio die Kosten senken, Storage und Computing voneinander entkoppeln, die Performance verbessern, das Datenmanagement vereinfachen, feingranulare Kontrollen bieten und Disaster Recovery-Funktionen bereitstellen würde.

**Lösung:**



Mit Dremio konnte NetApp seine Hadoop-basierte Dateninfrastruktur in mehreren Phasen modernisieren und so eine Roadmap für einheitliche Analysen bereitstellen. Im Gegensatz zu anderen Anbietern, die erhebliche Änderungen an Data Processing erforderten, konnte Dremio nahtlos in vorhandene Pipelines integriert werden, was Zeit und Kosten während der Migration einspart. Durch den Umstieg auf eine vollständig containerisierte Umgebung senkt NetApp den Management-Overhead, erhöht die Sicherheit und erhöht die Ausfallsicherheit. Die Einführung offener Ökosysteme wie Apache Iceberg und Arrow von Dremio sorgte für Zukunftssicherheit, Transparenz und Erweiterbarkeit.

Als Ersatz für die Hadoop/Hive Infrastruktur bot Dremio Funktionalität für sekundäre Anwendungsfälle über die semantische Schicht an. Während die bestehenden Spark-basierten ETL- und Datenaufnahmemechanismen blieben, bot Dremio eine einheitliche Zugriffsebene für eine einfachere Datenerkennung und -Exploration ohne Duplizierung. Bei diesem Ansatz wurden die Datenreplizierungsfaktoren deutlich reduziert und der Storage und das Computing wurden voneinander getrennt.

**Vorteile:** Mit Dremio erzielte NetApp deutliche Kosteneinsparungen, indem der Compute-Verbrauch und der Festplattenspeicherbedarf in seinen Datenumgebungen minimiert wurden. Der neue Active IQ Data Lake besteht aus 8,900 Tabellen mit 3 Petabyte Daten, verglichen mit der bisherigen Infrastruktur mit über 7 Petabyte. Bei der Migration zu Dremio war außerdem der Wechsel von 33 Mini-Clustern und 4,000 Kernen zu 16 Executor-Nodes auf Kubernetes-Clustern erforderlich. Selbst bei deutlich geringeren Computing-Ressourcen konnte NetApp eine bemerkenswerte Performance-Steigerung verzeichnen. Durch den direkten Datenzugriff über Dremio sank die Abfragezeit von 45 Minuten auf 2 Minuten. Dadurch wurden 95 % schnellere Erkenntnisse für die vorausschauende Wartung und Optimierung erzielt. Die Migration führte außerdem zu einer Senkung der Compute-Kosten um mehr als 60 %, mehr als 20-mal schnelleren Abfragen und mehr als 30 % Einsparungen bei den Gesamtbetriebskosten (TCO).

## Auto Parts Sales – Anwendungsfall Für Kunden.

**Herausforderungen:** Innerhalb dieses globalen Autoteile-Vertriebsunternehmens konnten die Geschäfts- und Unternehmensfinanzplanungs- und -Analysegruppen keine konsolidierte Ansicht der Umsatzberichte erhalten und waren gezwungen, die einzelnen Geschäftsabsatzkennzahlen-Berichte zu lesen und zu konsolidieren. Daher treffen Kunden Entscheidungen auf der Grundlage von Daten, die mindestens einen Tag alt waren. Die Vorlaufzeiten für neue Erkenntnisse aus Big-Data-Analysen dauern in der Regel länger als vier Wochen. Die Fehlerbehebung bei Daten-Pipelines würde noch länger dauern und die bereits lange Zeitachse um zusätzliche drei Tage oder länger erweitern. Der langsame Entwicklungsprozess von Berichten und die hohe Berichtsleistung zwangen die Analysten, kontinuierlich auf die Verarbeitung oder das Laden von Daten zu

warten, anstatt neue geschäftliche Erkenntnisse zu gewinnen und neues Geschäftsverhalten voranzutreiben. Diese problematischen Umgebungen bestanden aus zahlreichen unterschiedlichen Datenbanken für verschiedene Geschäftsbereiche, woraus sich zahlreiche Datensilos ergaben. Die langsame und fragmentierte Umgebung kompliziert die Data Governance, da es zu viele Möglichkeiten für Analysten gab, ihre eigene Version der Wahrheit gegen eine einzige Quelle der Wahrheit zu finden. Die Kosten für die Datenplattform betragen über 1,9 Millionen US-Dollar, und die Personalkosten. Für die Wartung der alten Plattform und das Ausfüllen von Datenanfragen waren sieben Außendiensttechniker (Field Technical Engineers, FTEs) pro Jahr erforderlich. Angesichts der wachsenden Datenanforderungen war das Data Intelligence-Team nicht in der Lage, die ältere Umgebung an künftige Anforderungen anzupassen

**Lösung:** Kostengünstige Speicherung und Verwaltung großer Iceberg-Tabellen im NetApp-Objektspeicher. Erstellen Sie Datendomänen mithilfe der semantischen Ebene von Dremio, sodass Geschäftsanwender problemlos Datenprodukte erstellen, suchen und gemeinsam nutzen können.

**Vorteile für den Kunden:** • Verbesserte und optimierte vorhandene Datenarchitektur und reduzierte die Zeit für Einblicke von vier Wochen auf wenige Stunden • reduzierte die Fehlerbehebungszeit von drei Tagen auf nur wenige Stunden • reduzierte die Kosten für die Datenplattform und das Management um mehr als 380.000 US-Dollar • (2) jährliche Einsparungen durch die FTEs of Data Intelligence

## Schlussfolgerung

Zusammenfassend hat dieser technische Bericht umfassende Details zur Implementierung von q Hybrid Iceberg Lakehouse mit Dremio in Verbindung mit verschiedenen Datenquellen von NetApp Storage Controllern, darunter ONTAP S3, NAS und StorageGRID, bereitgestellt. Der Bereitstellungsprozess wurde erfolgreich ausgeführt, und das TPC-DS Benchmarking-Tool wurde verwendet, um 99 SQL-Abfragen über die verschiedenen Datenquellen hinweg durchzuführen. Der Bericht hat auch Kundennutzungsfälle innerhalb von NetApp untersucht und die Vielseitigkeit und Effektivität von Dremio bei der Erfüllung unterschiedlicher geschäftlicher Anforderungen demonstriert. Darüber hinaus wurde ein konkreter Anwendungsfall mit einem Autoteile-Vertriebskunde untersucht, der die praktische Anwendung und die Vorteile von Dremio für die Datenanalyse und die Erkenntnisse hervorhebt.

Insgesamt dient dieses Dokument als wertvolle Ressource zum Verständnis der Implementierung und Verwendung von Dremio mit NetApp Storage Controllern und stellt dessen Funktionen und Potenziale zur Förderung von datengetriebenen Entscheidungsprozessen und der Optimierung in verschiedenen Branchen vor.

## Wo Sie weitere Informationen finden

Sehen Sie sich die folgenden Dokumente und/oder Websites an, um mehr über die in diesem Dokument beschriebenen Informationen zu erfahren:

- Installation von ZooKeeper

<https://medium.com/@ahmetfurkandemir/distributed-hadoop-cluster-1-spark-with-all-dependencies-03c8ec616166>

- Dremio

<https://docs.dremio.com/current/get-started/cluster-deployments/deployment-models/standalone/standalone-tarball/>

- Dremio mit StorageGRID konfigurieren

<https://docs.netapp.com/us-en/storagegrid-enable/tools-apps-guides/configure-dremio-storagegrid.html#configure-dremio-data-source>

- Anwendungsfall für NetApp

<https://www.dremio.com/customers/netapp/>

## 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.