



NetApp 與 Dremio 的新一代混合冰山湖屋解決方案 NetApp Solutions

NetApp
September 26, 2024

目錄

NetApp 與 Dremio 的新一代混合冰山湖屋解決方案	1
NetApp 與 Dremio 次世代混合冰山湖屋解決方案	1
解決方案總覽	1
技術需求	3
部署程序	3
解決方案驗證總覽	14
客戶使用案例	21
結論	23

NetApp 與 Dremio 的新一代混合冰山湖屋解決方案

NetApp 與 Dremio 次世代混合冰山湖屋解決方案

在本文件中、我們將討論 Dremio 的部署詳細資料、以及 NetApp 儲存控制器的不同資料來源、包括 ONTAP S3、NAS 和 StorageGRID。在部署期間、我們使用 TPC-DS 基準測試工具、跨各種來源執行 99 個 SQL 查詢。本文件也探討 NetApp 中的客戶使用案例、以及涉及汽車零件銷售客戶的使用案例。

- 作者 *

NetApp、Inc. 首席架構設計師 Karthkeyan Nagalingam、Dremio Corporation 聯盟副總裁 Roger Frey。Mark Shainman、Dremio Corporation 首席產品行銷經理。

解決方案總覽

混合式冰山 Lakehouse 解決方案提供獨特的優勢、可因應資料湖客戶所面臨的客戶挑戰。藉由運用 Dremio Unified Lakehouse 平台和 NetApp ONTAP、StorageGRID 和 NetApp 雲端解決方案、公司可以為其業務營運增加重大價值。此解決方案不僅提供多個資料來源的存取、包括 NetApp 來源、也能提升整體分析效能、協助公司推動業務洞見、進而推動業務成長。

NetApp總覽

- NetApp 的產品（例如 ONTAP 和 StorageGRID）可分離儲存和運算、根據特定需求提供最佳的資源使用率。這種靈活度可讓客戶使用 NetApp 儲存解決方案來擴充儲存設備
- 藉由運用 NetApp 的儲存控制器、客戶可以使用 NFS 和 S3 傳輸協定、將資料有效地提供給向量資料庫。這些通訊協定可協助客戶儲存資料並管理向量資料庫索引、不需要透過檔案和物件方法存取多個資料複本。
- NetApp ONTAP 在 AWS、Azure 和 Google Cloud 等頂尖雲端服務供應商之間、提供 NAS 和物件儲存的原生支援。這種廣泛的相容性可確保無縫整合、實現客戶資料的行動力、全球存取能力、災難恢復、動態擴充性和高效能。

StorageGRID

我們領先業界的物件儲存 StorageGRID 提供強大的原則引擎、可自動放置資料、提供靈活的部署選項、以及無與倫比的耐用性與分層銷毀編碼。它擁有可擴充的架構、可在單一命名空間中支援數十億個物件和 PB 的資料。此解決方案可實現混合雲整合、讓資料分層化至主要雲端平台。在 2019 年 IDC Marketscape 全球物件型廠商評估中、IT 獲公認為領導者。

此外、StorageGRID 還具備軟體定義的物件儲存、地理區備援和多站台功能、可協助您大規模管理非結構化資料。它整合了原則型資訊生命週期管理、並提供鏡射和搜尋等雲端整合功能。本公司擁有多項認證、包括通用標準、NF203 數位安全元件、ISO/IEC 25051、KPMG、以及 Cohasset 法規遵循評估。

總而言之、NetApp StorageGRID 提供強大的功能、擴充性、混合雲整合、以及法規遵循認證、可有效大規模管理非結構化資料。

NetApp ONTAP

NetApp ONTAP 是一款功能強大的儲存解決方案、提供各種企業功能。其中包含 Snapshot、可提供應用程式一致且防竄改的即時備份。SnapRestore 可隨需即時還原備份、而 SnapMirror 則提供整合式遠端備份與災難恢復功能。此解決方案也整合了自動勒索軟體保護（ARP）、以多重系統管理員驗證、FIPS 認證的靜態資料加密、傳輸中資料加密、多因素驗證（MFA）、以及角色型存取控制（RBAC）等功能、確保資料安全。全面的記錄、稽核、內建與外部金鑰管理、安全清除、以及安全管理多個租戶、進一步強化資料安全性與法規遵循。

NetApp ONTAP 也採用 SnapLock、以低總持有成本提供符合法規要求的資料保留功能、提供高層級的完整性、效能和保留。它與 NetApp ONTAP® 9 完全整合、可防範惡意行為、惡意管理員和勒索軟體。

此解決方案包含 NSE/NVE 加密、可用於執行中和資料靜止加密、多因素管理存取、以及多重管理驗證。Active IQ 提供 AI 導向的預測分析和修正行動、而 QoS 則可確保服務工作負載的品質控制。透過 SysMgr/GUI/CLI/API、管理與自動化整合是直覺式的。FabricPool 可實現自動資料分層、而此解決方案可透過內嵌資料壓縮、重複資料刪除和壓縮等方式、提供效率。NetApp 保證客戶能免費達成工作負載效率目標。

NetApp ONTAP 支援各種通訊協定、包括 NVMe / FC、FC、NVMe / TCP、iSCSI、NFS、SMB 和 S3、使其成為統一化儲存解決方案。總的來說、NetApp ONTAP 提供廣泛的企業功能、強大的安全性、法規遵循、效率和多用途性、以滿足各種儲存需求。

Dremio 總覽

Dremio 是 Unified Lakehouse 平台、用於自助分析和 AI。Dremio Unified Analytics Platform 以舊有資料倉儲解決方案的一小部分成本、提供 Lakehouse 的靈活度、擴充性和效能、讓使用者更接近資料。Dremio 啟用「左移」分析功能、可消除複雜且昂貴的資料整合和 ETL、提供無縫的企業級分析、無需資料移動。Dremio 也具備下列特色：

- 透過通用語義層和緊密整合、高效能的 SQL 查詢引擎、提供易於使用的自助服務分析功能、讓您更輕鬆地在雲端和內部部署中連線、管理及分析所有資料。
- Dremio 的 Apache iceberg- 原生湖屋管理功能可簡化資料探索、並自動化資料最佳化、以 Git 為靈感的資料版本管理功能提供高效能分析。
- 在開放原始碼和開放式標準的基礎上、Dremio 讓公司能夠避免束縛、並保持創新的定位。企業公司相信 Dremio 是最容易使用的 Lakehouse 平台、能在所有工作負載中提供最佳的價格效能。

Dremio 和 NetApp Hybrid 冰山 Lakehouse 解決方案能為客戶帶來什麼價值？

- *** 改善資料管理與存取能力 ***：Dremio 以其資料湖平台聞名、可讓組織以高速從資料湖直接查詢資料。另一方面、NetApp 則是雲端資料服務與資料儲存解決方案的領導供應商。這項聯合方案可為客戶提供全方位的解決方案、讓客戶以高效率且有效率的方式儲存、管理、存取及分析企業資料。
- *** 效能最佳化 ***：憑藉 NetApp 在資料儲存和 Dremio 在 Data Processing 與資料最佳化方面的專業知識、這項合作夥伴關係提供了一套解決方案、可改善資料作業效能、縮短延遲時間、並加快業務洞察速度。Dremio 甚至為 NetApp 本身的內部 IT 分析基礎架構提供效能優勢。
- *** 擴充性 ***：Dremio 和 NetApp 都提供專為擴充而設計的解決方案。這套聯合解決方案為客戶提供可高度擴充的資料儲存、資料管理和分析環境。在混合式冰山 Lakehouse 環境中、與 NetApp StorageGRID 搭配使用的 Dremio SQL 查詢引擎可提供無與倫比的擴充性、並行性和查詢效能、能夠處理任何企業的分析需求。
- *** 資料安全與治理 ***：兩家公司都非常重視資料安全與治理。它們共同提供強大的安全性和資料治理功能、確保資料受到保護、並符合資料治理需求。功能包括角色型和精細存取控制、全方位稽核、端點對端點資料系列、統一化身識別管理、以及具廣泛法規遵循與安全架構的 SSO、確保公司的分析資料環境安全無虞且受到妥善管理。
- *** 成本效益 ***：透過將 Dremio 的資料湖引擎與 NetApp 的儲存解決方案整合、客戶可以降低與資料管理和資

料移動相關的成本。組織也能從舊版資料湖環境移轉到由 NetApp 和 Dremio 組成的更現代化的湖屋解決方案。這款混合式冰山 Lakehouse 解決方案提供高速查詢效能、以及領先市場的查詢並行處理功能、可降低 TCO 並縮短企業洞悉時間。

技術需求

下列硬體和軟體組態已用於本文件中所執行的驗證。這些組態可作為協助您設定環境的準則、但請注意、具體元件可能會因個別客戶需求而異。

硬體需求

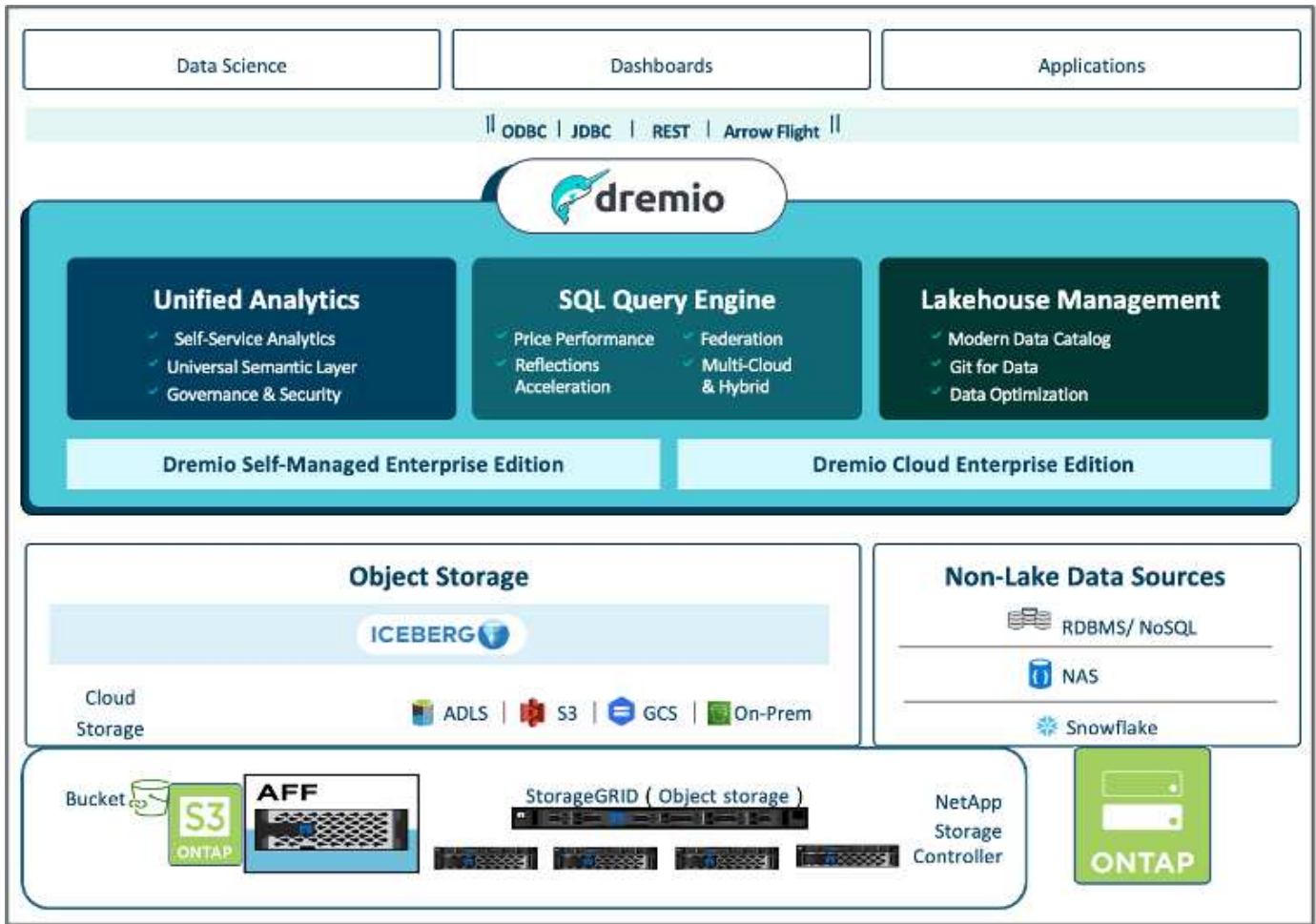
硬體	詳細資料
NetApp AFF 儲存陣列 HA 配對	<ul style="list-style-type: none"> • A800 • ONTAP 9.14.1. • 48 x 3.49TB SSD - NVM • 兩個 S3 工作箱：Dremio 中繼資料和客戶資料。
4 x Fujitsu PRIMERGY RX2540 M4	<ul style="list-style-type: none"> • 64 個 CPU • Intel® Xeon® 黃金級 6142 CPU @ 2.60GHz • 256 GM 實體記憶體 • 1 個 100GbE 網路連接埠
網路	<ul style="list-style-type: none"> • 100 GbE
StorageGRID	* 1 x SG100、3 x SGF6024 * 3 x 24 x 7.68TB * 兩個 S3 鏟斗：Dremio 中繼資料和客戶資料。

軟體需求

軟體	詳細資料
夢中	<ul style="list-style-type: none"> • 版本 - 25.0.3-202405170357270647-d2042e1b • Enterprise Edition
內部部署	<ul style="list-style-type: none"> • 5 節點 Dremio 叢集 • 1 位主要協調員和 4 位執行者

部署程序

在此參考架構驗證中、我們使用由一位協調者和四位執行者組成的 Dremio 組態



NetApp 設定

- 儲存系統初始化
- 建立儲存虛擬機器 (SVM)
- 邏輯網路介面指派
- NFS、S3 組態和授權

請遵循下列 NFS (網路檔案系統) 步驟：1. 為 NFSv4 或 NFSv3 建立 Flex Group Volume。在我們的驗證設定中、我們使用了 48 個 SSD、1 個 SSD 專用於控制器的根磁碟區、47 個 SSD 散佈於 NFSv4 中]。確認 Flex Group Volume 的 NFS 匯出原則具有 Dremio 伺服器網路的讀取 / 寫入權限。

1. 在所有 Dremio 伺服器上、建立一個資料夾、並透過每個 Dremio 伺服器上的邏輯介面 (LIF)、將 Flex Group Volume 掛載至此資料夾。

請依照下列步驟執行 S3 (Simple Storage Service)：

1. 使用「vserver object-store-server create」命令設定物件儲存區伺服器、並將管理狀態設定為「up」。您可以選擇啟用 HTTPS 並設定自訂接聽程式連接埠。
2. 使用「vserver object-store-server user create -user <username>」命令建立物件儲存伺服器使用者。
3. 若要取得存取金鑰和秘密金鑰、您可以執行下列命令：「Set diag; vserver object-store-server user show -user <username>」。不過、在使用者建立程序期間、將會提供這些金鑰、也可以使用 REST API 呼叫來擷取這些金鑰。

4. 使用在步驟 2 中建立的使用者建立物件儲存區伺服器群組、並授予存取權。在此範例中、我們提供了「FullAccess」。
5. 將 S3 儲存區的類型設定為「S3」、即可建立兩個 S3 儲存區。一個用於 Dremio 組態、另一個用於客戶資料。

zookeeper 設定

您可以使用 Dremio 提供的 zookeeper 組態。在這項驗證中、我們使用了獨立的 zookeeper。我們遵循本網頁連結中所述的步驟 <https://medium.com/@ahmetfurkandemir/distributed-hadoop-cluster-1-spark-with-all-dependencies-03c8ec616166>

Dremio 設定

我們跟隨這個網頁連結、透過 tar ball 安裝 Dremio。

1. 建立 Dremio 群組。

```
sudo groupadd -r dremio
```

2. 建立 dremio 使用者。

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

3. 建立 Dremio 目錄。

```
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. 從下載 tar 檔案 <https://download.dremio.com/community-server/>

5. 將 Dremio 解壓縮至 /opt/dremio 目錄。

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

6. 為組態資料夾建立符號連結。

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

7. 設定您的服務組態 (SystemD 設定)。

- a. 從 /opt/dremio/share/dremio.service 將 dremio 守護程序的單元文件複製到

/etc/systemd/system/dremio.service ◦

b. 重新啟動系統

```
sudo systemctl daemon-reload
```

c. 啟用 dremio 以在開機時啟動 ◦

```
sudo systemctl enable dremio
```

8. 在協調器上設定 Dremio ◦ 如需詳細資訊、請參閱 Dremio 組態

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.

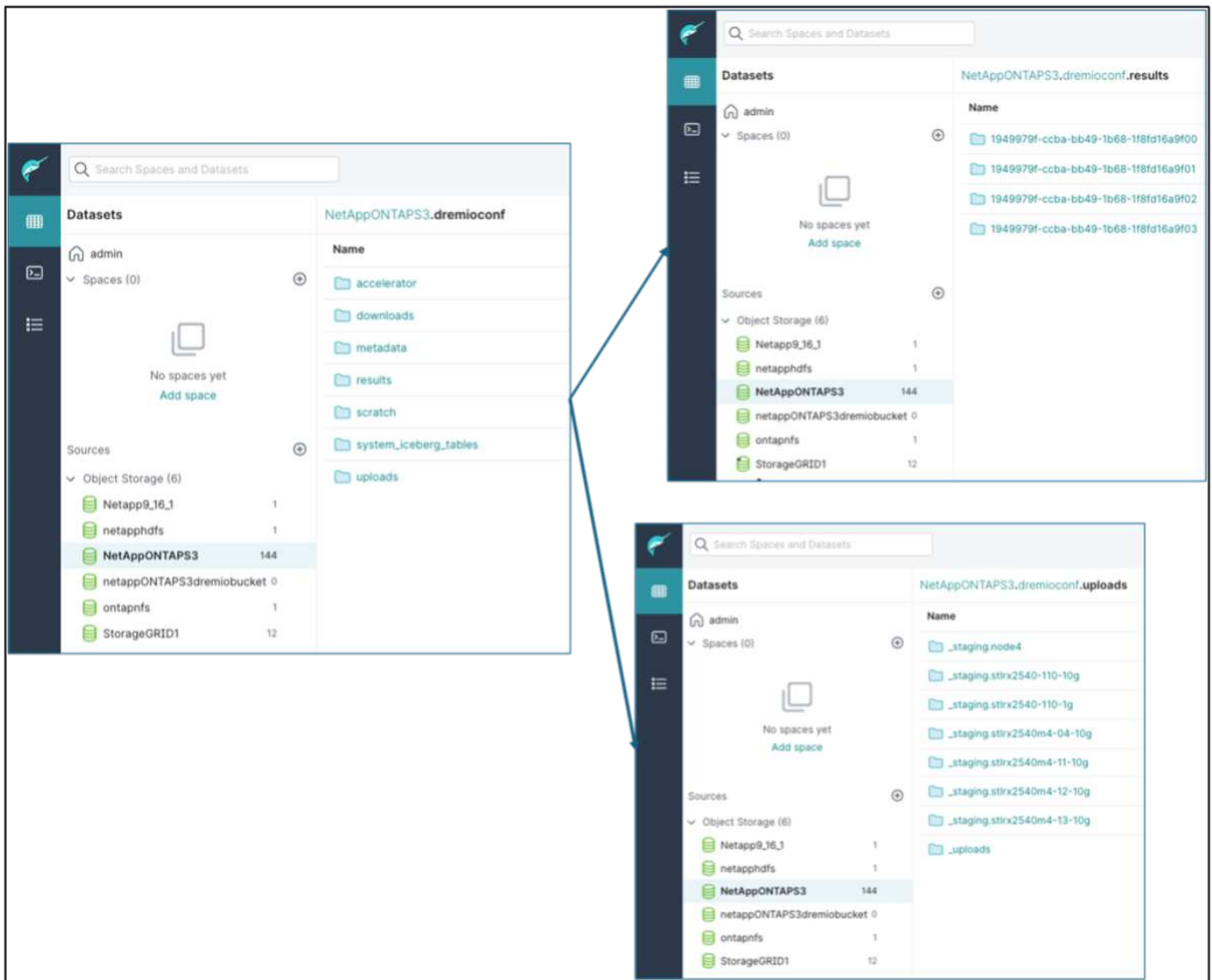
-->

<!-- 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. Dremio 組態儲存在 NetApp 物件儲存區中。在我們的驗證中、「dremioconf」貯體位於 ONTAP S3 貯體中。下圖顯示「dremioconf」S3 儲存區的「Scratch」和「uploads」資料夾中的一些詳細資料。



1. 在執行者上設定 Dremio 。在我們的設定中、我們有 3 位執行者。
 - 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 –與協調者組態相同。



NetApp 建議 StorageGRID 做為其主要目標儲存解決方案、以用於 Datalake 和 Lakehouse 環境。此外、NetApp ONTAP 也用於檔案 / 物件雙重性。在本文中、我們已針對客戶的要求、對 ONTAP S3 進行測試、並成功地做為資料來源。

多個來源設定

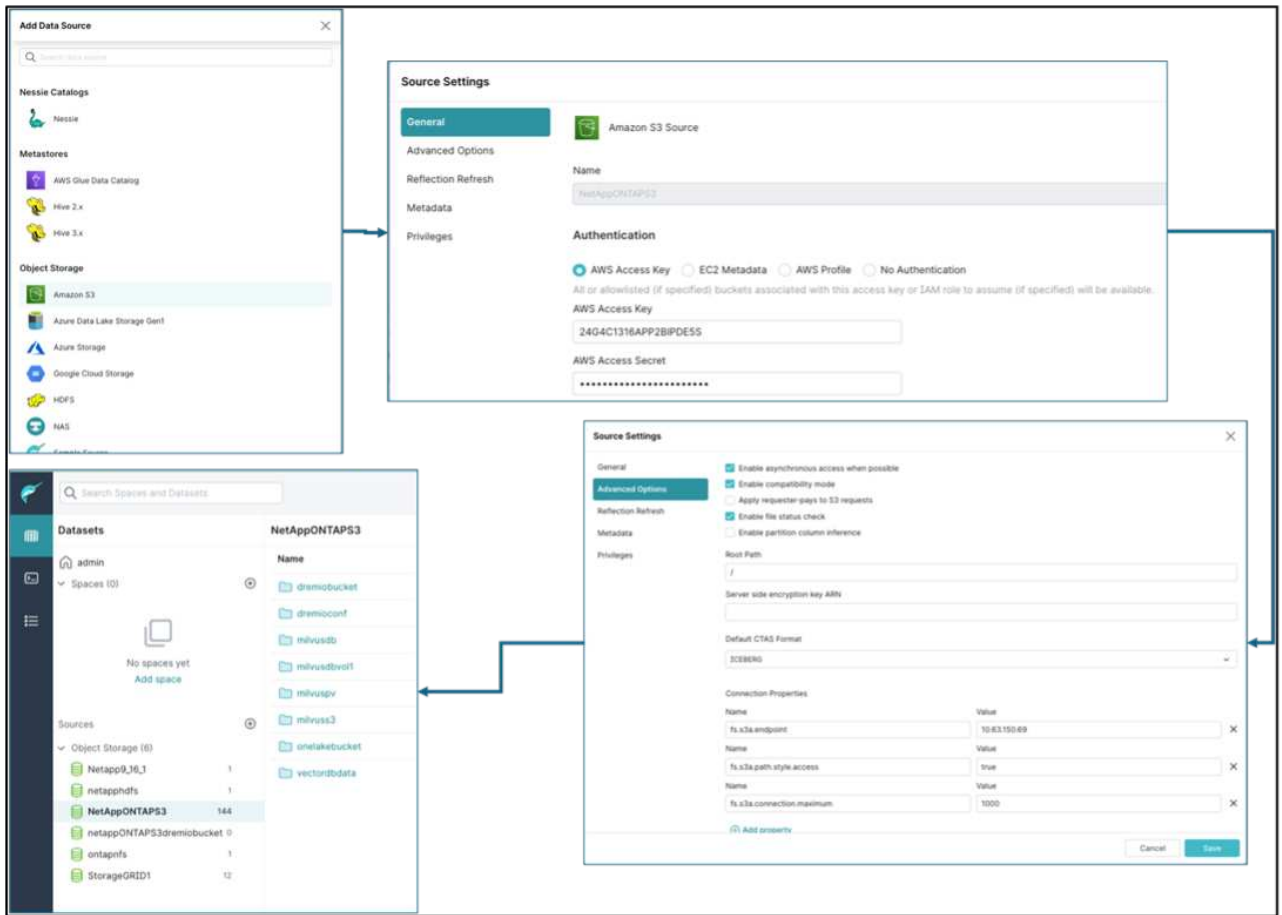
1. 在 Dremio 中將 ONTAP S3 和 StorageGRID 設定為 S3 來源。
 - a. Dremio 儀表板 → 資料集 → 來源 → 新增來源。
 - b. 請在一般章節中更新 AWS 存取和秘密金鑰
 - c. 在進階選項中、啟用相容模式、以下列詳細資料更新連線內容。來自 NetApp 儲存控制器的端點 IP/ 名稱、來自 ONTAP S3 或 StorageGRID。

```

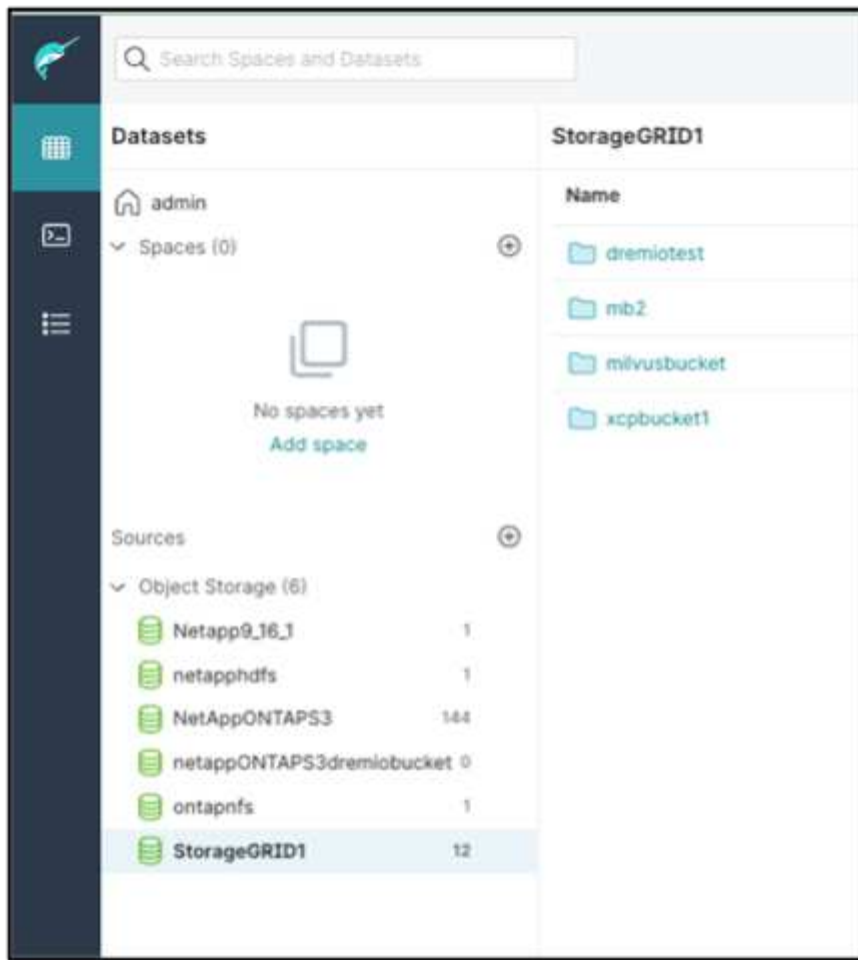
fs.s3a.endpoint = 10.63.150.69
fs.s3a.path.style.access = true
fs.s3a.connection.maximum=1000

```

- d. 盡可能啟用本機快取、盡可能使用的可用快取總容量百分比上限 = 100
- e. 然後檢視 NetApp 物件儲存區的貯體清單。



f. StorageGRID 貯體詳細資料的範例檢



視

2. 在 Dremio 中將 NAS（特別是 NFS）設定為來源。
 - a. Dremio 儀表板 → 資料集 → 來源 → 新增來源。
 - b. 在一般區段中、輸入名稱和 NFS 裝載路徑。請確定 NFS 裝載路徑已掛載於 Dremio 叢集中所有節點的同一個資料夾。

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	Count
Netapp9_16_1	1
netappphdfs	1
NetAppONTAPS3	144
netappONTAPS3dremiobucket	0
ontapnfs	1
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:~#

```

解決方案驗證總覽

在本節中、我們執行了來自多個來源的 SQL 測試查詢、以驗證功能、測試及驗證 NetApp 儲存設備的外溢問題。

物件儲存的 SQL 查詢

1. 在 `dremio.env` 中將每部伺服器的記憶體設為 250GB


```

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:~#

```

2. 檢查 `dremio.conf` 檔案和儲存詳細資料中的溢漏位置 (`$ {DREMIO_HOME} 「/dremiocache`) 。

```

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

```

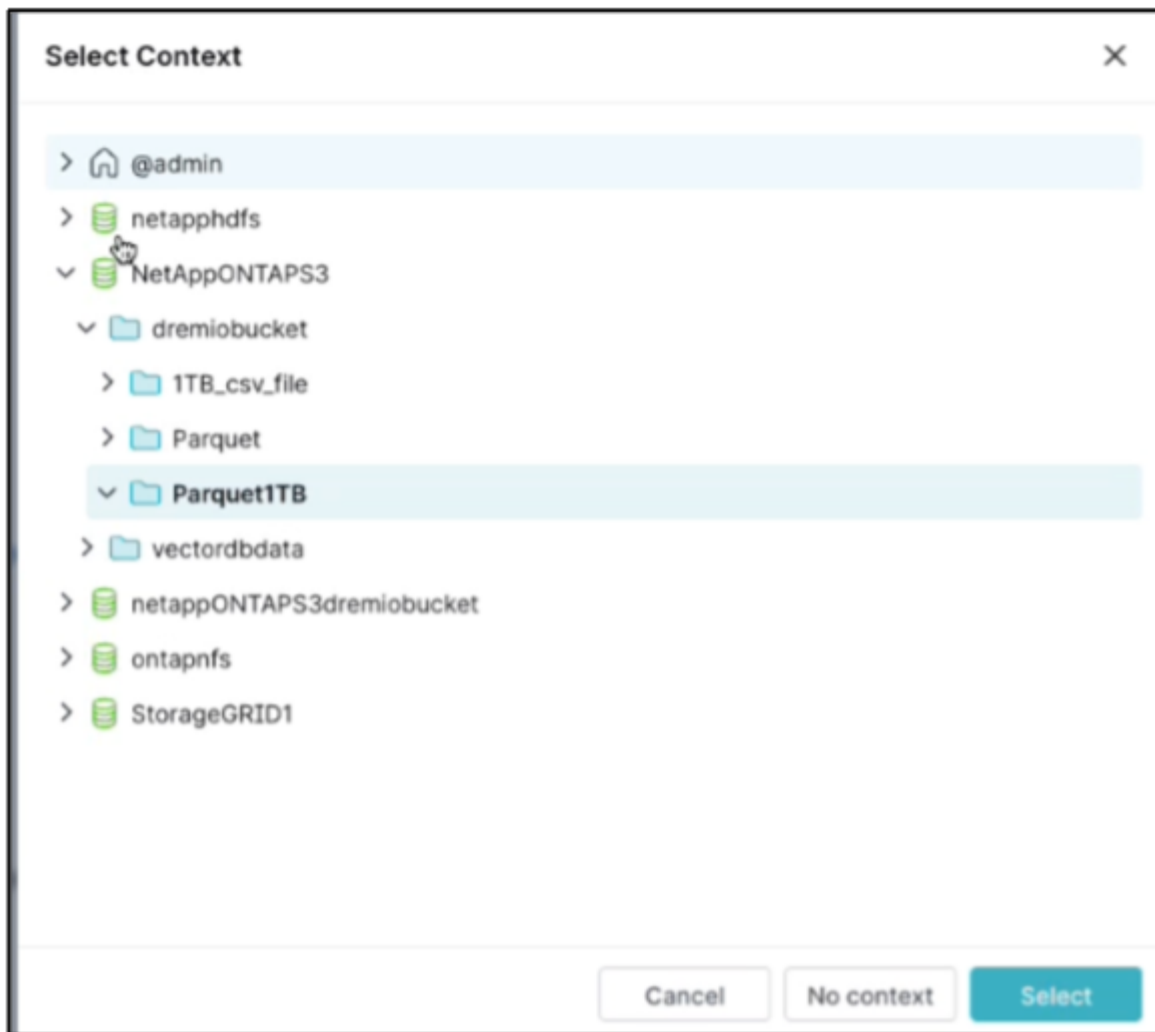
3. 將 Dremio 溢漏點指向 NetApp NFS 儲存設備

```

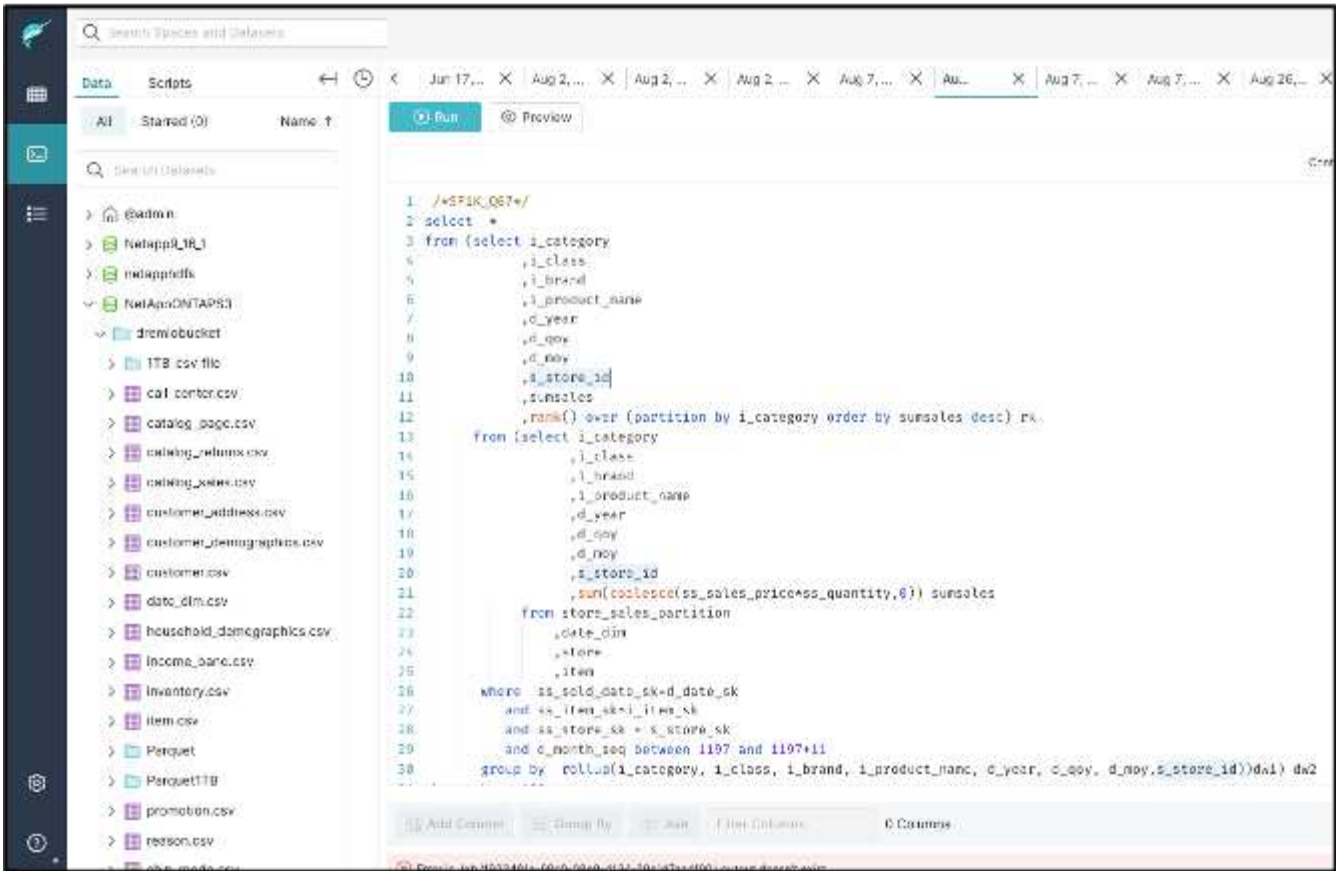
root@hadoopnode1:~# ls -ltrh /dremiocache
total 4.0K
drwx----- 3 nobody nogroup 4.0K Sep 13 16:00 spilling_stlrx2540m4-12-
10g_45678
root@hadoopnode1:~# ls -ltrh /opt/dremio/dremiocache/
total 8.0K
drwxr-xr-x 3 dremio dremio 4.0K Aug 22 18:19 spill_old
drwxr-xr-x 4 dremio dremio 4.0K Aug 22 18:19 cm
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. 選取內容。在我們的測試中、我們針對 ONTAP S3 中的 TPCDS 產生的硬地板檔案執行測試。Dremio 儀表板 → SQL runner → 內容 → NetAppONTAPS3→Parquet1TB



1. 從 Dremio 儀表板執行 TPC-DS 查詢 67



1. 检查工作是否在所有執行者上執行。Dremio 儀表板 → 工作 → <jobid> → 原始設定檔 → 選取 external sort → 主機名稱

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	0.004s	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	0.003s	0.000s	0.000s	0	0	128KB	str2540m4-13-10g	0	CAN_CONSUME	16:35:54

1. 執行 SQL 查詢時、您可以在 NetApp 儲存控制器中檢查分割資料夾的資料快取。

```

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 /dremiocache/spilling_stlrx2540m4-12-10g_45678/
total 4.0K
drwxr-xr-x 2 root daemon 4.0K Sep 13 16:23 1726243167416

```

2. SQL 查詢已完成、而且會溢出

Job ID	User	Dataset	Query Type	Queue	Start Time	Duration	SQL
19383301-5cd8-0a48-1e38-e2f5b4148f00	admin	store_sales_partition	JDBC Client	High Cost User Q...	08/22/2024, 19:42:54	00:08:25	/*SF1K_Q67*/ select + from (select i_category ,i_class ,i_brand ,i_product_name ,d_year ,d_qoy ,d_moy
19387983-2031-164f-cd9e-57c6c287bd00	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_qoy ,d_moy

3.

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 (0.00%)
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
- date_dim
- store

Show more >

Scans

- store_sales_partition
- date_dim
- store
- item

工作完成摘要。

4. 檢查溢出的資料大

EXTERNAL_SORT 04-06



Runtime	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

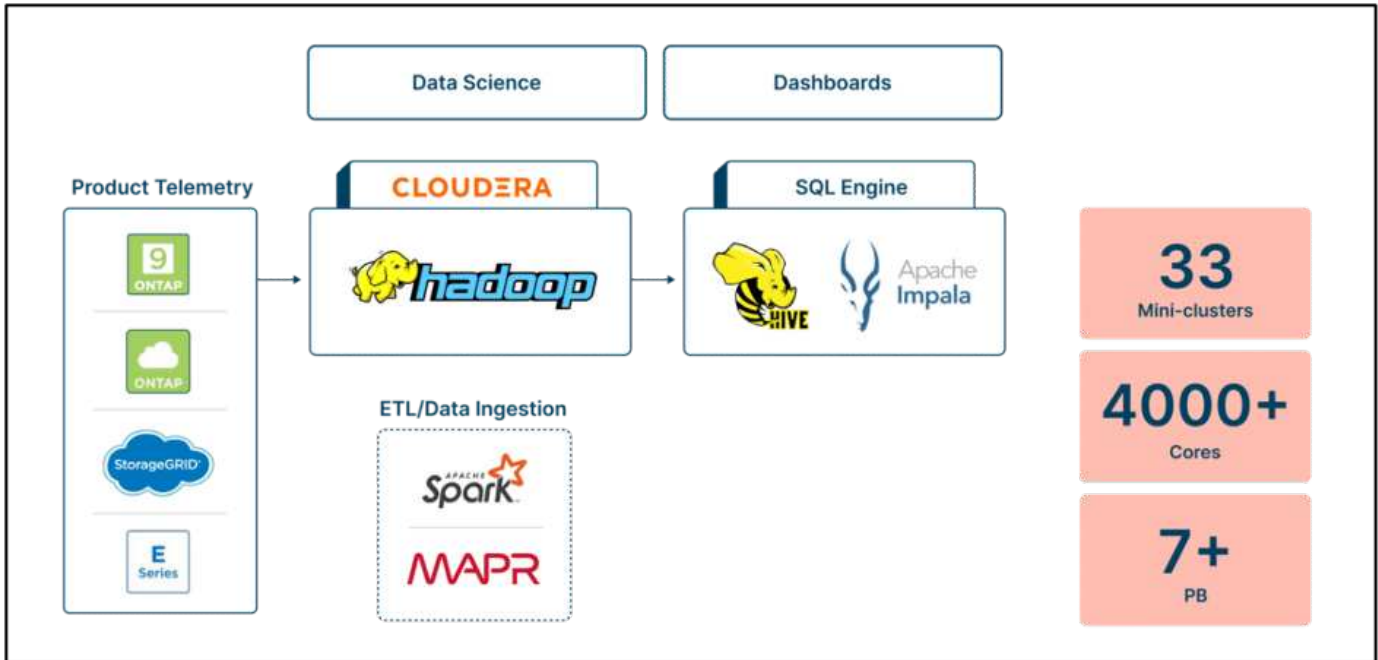
Merge Time Nanos:	0ns
Spill Count:	360
Spill Time Nanos:	37.68m
Total Spilled Data Size:	20,339,702,765
Batches Spilled:	97,854

小

適用於 NAS 和 StorageGRID 物件儲存的相同程序。

客戶使用案例

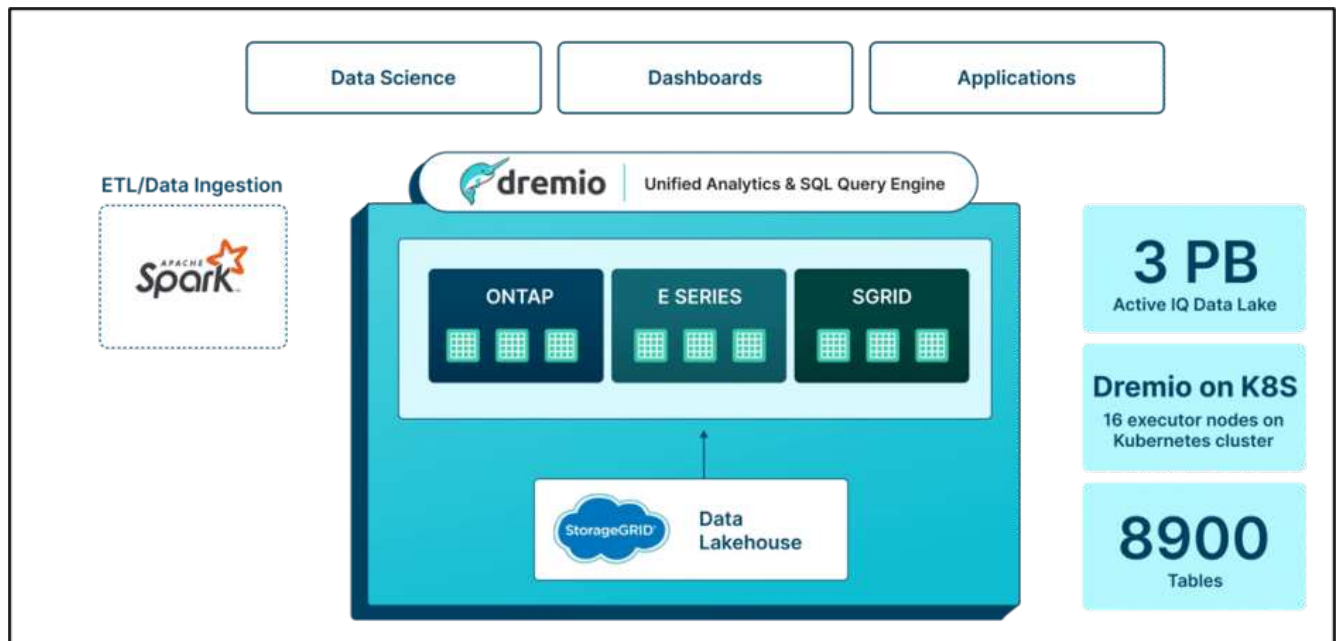
NetApp ActiveIQ 使用案例



- 挑戰 *：NetApp 本身的內部 Active IQ 解決方案、最初設計用於支援多種使用案例、已發展成適用於內部使用者和客戶的全方位產品。然而、由於資料快速成長、且需要有效率的資料存取、因此底層的 Hadoop / MapR 型後端基礎架構在成本和效能方面帶來挑戰。擴充儲存設備意味著增加不必要的運算資源、進而增加成本。

此外、管理 Hadoop 叢集非常耗時、而且需要專業的專業知識。資料效能和管理問題使情況更加複雜、因為查詢平均需要 45 分鐘、而且因為組態錯誤而導致資源不足。為了因應這些挑戰、NetApp 尋求替代現有的舊版 Hadoop 環境、並決定採用 Dremio 打造的全新現代化解決方案、可降低成本、分離儲存與運算、改善效能、簡化資料管理、提供精細的控制功能、並提供災難恢復功能。

- 解決方案 *：



Dremio 讓 NetApp 以分階段方式將其 Hadoop 型資料基礎架構現代化、為統一化分析提供藍圖。與其他需要對 Data Processing 進行重大變更的廠商不同、Dremio 與現有的管線無縫整合、可節省移轉期間的時間與費用。NetApp 改用完全容器化的環境、降低了管理成本、改善了安全性、並增強了恢復能力。Dremio 採用開放式生態系統、例如 Apache 冰山和 Arrow、確保了未來的保障、透明度和擴充性。

作為 Hadoop / Hive 基礎架構的替代方案、Dremio 透過語義層提供次要使用案例的功能。雖然現有的 Spark 型 ETL 和資料擷取機制仍然存在、但 Dremio 提供了統一的存取層、可更輕鬆地探索和探索資料、而不會產生重複資料。這種方法可大幅降低資料複寫因素、並將儲存與運算分離。

- 效益 * : 透過 Dremio、NetApp 可將資料環境的運算消耗和磁碟空間需求降至最低、大幅降低成本。新的 Active IQ 資料湖包含 8、900 個資料表、可容納 3 PB 的資料、相較於先前的基礎架構、容量超過 7 PB。移轉至 Dremio 也涉及從 33 個迷你叢集和 4、000 個核心移轉至 Kubernetes 叢集上的 16 個執行節點。即使運算資源大幅減少、NetApp 也獲得顯著的效能提升。透過 Dremio 直接存取資料、查詢執行時間從 45 分鐘縮短為 2 分鐘、因此可加快 95% 的時間、以利預測性維護與最佳化。移轉作業也能降低 60% 以上的運算成本、加快 20 倍以上的查詢速度、並節省 30% 以上的總持有成本 (TCO)。

汽車零件銷售客戶使用案例。

- 挑戰 * : 在這家全球汽車零件銷售公司內、執行與企業財務規劃與分析團隊無法取得銷售報告的整合檢視、被迫閱讀個別業務銷售指標報告、並嘗試將其整合。這導致客戶決定使用至少一天前的資料。取得新分析洞見的前置時間通常需要四週以上。疑難排解資料傳輸管道需要更多時間、在原本已過長的時間線中再增加三天或更長的時間。報告開發流程緩慢、報告效能也迫使分析師社群持續等待資料處理或載入、而非讓他們找到新的業務洞見、並推動新的業務行為。這些困難環境由許多不同的資料庫組成、適用於不同的業務單位、因此產生許多資料封閉環境。緩慢且零散的環境使資料管理變得複雜、因為分析師有太多方法可以自行找出真實的版本、而不是單一的真實來源。此方法的資料平台和人員成本超過 190 萬美元。維護舊版平台並滿足資料要求、每年需要七位現場技術工程師 (FTE)。隨著資料要求不斷增加、資料情報團隊無法擴充舊環境以滿足未來的需求
- 解決方案 * : 以符合成本效益的方式、在 NetApp 物件存放區中儲存及管理大型冰山表格。使用 Dremio 的語義層建置資料網域、讓企業使用者能夠輕鬆建立、搜尋及共用資料產品。
- 對客戶的好處 * : 改善並最佳化現有資料架構、並將深入分析時間從四週縮短為幾小時、將疑難排解時間從三天縮短為僅幾小時、資料平台與管理成本減少超過 38 萬美元、(2) 資料情報工作的 FTE 每年節省成本

結論

最後、本技術報告提供與 Dremio 搭配使用的 q Hybrid 冰山 Lakehouse 的完整部署詳細資料、以及 NetApp 儲存控制器的各種資料來源、包括 ONTAP S3、NAS 和 StorageGRID。部署程序已成功執行、並使用 TPC-DS 基準測試工具、在不同的資料來源中執行 99 個 SQL 查詢。本報告也探討了 NetApp 中的客戶使用案例、展現 Dremio 在滿足各種業務需求方面的靈活性與效率。此外、我們也針對涉及汽車零件銷售客戶的特定使用案例進行了審查、重點說明運用 Dremio 進行資料分析和深入分析的實際應用和效益。

總的來說、本文件是瞭解 Dremio 與 NetApp 儲存控制器的部署與使用的寶貴資源、展示其在不同產業中推動資料導向決策與最佳化的能力與潛力。

何處可找到其他資訊

若要深入瞭解本文所述資訊、請檢閱下列文件和 / 或網站：

- zookeeper 安裝

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

- 夢中

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

- 使用 StorageGRID 設定 Dremio

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

- NetApp 使用案例

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

版權資訊

Copyright © 2024 NetApp, Inc. 版權所有。台灣印製。非經版權所有人事先書面同意，不得將本受版權保護文件的任何部分以任何形式或任何方法（圖形、電子或機械）重製，包括影印、錄影、錄音或儲存至電子檢索系統中。

由 NetApp 版權資料衍伸之軟體必須遵守下列授權和免責聲明：

此軟體以 NETAPP「原樣」提供，不含任何明示或暗示的擔保，包括但不限於有關適售性或特定目的適用性之擔保，特此聲明。於任何情況下，就任何已造成或基於任何理論上責任之直接性、間接性、附隨性、特殊性、懲罰性或衍生性損害（包括但不限於替代商品或服務之採購；使用、資料或利潤上的損失；或企業營運中斷），無論是在使用此軟體時以任何方式所產生的契約、嚴格責任或侵權行為（包括疏忽或其他）等方面，NetApp 概不負責，即使已被告知有前述損害存在之可能性亦然。

NetApp 保留隨時變更本文所述之任何產品的權利，恕不另行通知。NetApp 不承擔因使用本文所述之產品而產生的責任或義務，除非明確經過 NetApp 書面同意。使用或購買此產品並不會在依據任何專利權、商標權或任何其他 NetApp 智慧財產權的情況下轉讓授權。

本手冊所述之產品受到一項（含）以上的美國專利、國外專利或申請中專利所保障。

有限權利說明：政府機關的使用、複製或公開揭露須受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中的「技術資料權利 - 非商業項目」條款 (b)(3) 小段所述之限制。

此處所含屬於商業產品和 / 或商業服務（如 FAR 2.101 所定義）的資料均為 NetApp, Inc. 所有。根據本協議提供的所有 NetApp 技術資料和電腦軟體皆屬於商業性質，並且完全由私人出資開發。美國政府對於該資料具有非專屬、非轉讓、非轉授權、全球性、有限且不可撤銷的使用權限，僅限於美國政府為傳輸此資料所訂合約所允許之範圍，並基於履行該合約之目的方可使用。除非本文另有規定，否則未經 NetApp Inc. 事前書面許可，不得逕行使用、揭露、重製、修改、履行或展示該資料。美國政府授予國防部之許可權利，僅適用於 DFARS 條款 252.227-7015(b)（2014 年 2 月）所述權利。

商標資訊

NETAPP、NETAPP 標誌及 <http://www.netapp.com/TM> 所列之標章均為 NetApp, Inc. 的商標。文中所涉及的所有其他公司或產品名稱，均為其各自所有者的商標，不得侵犯。