

解決方案簡介

# 使用 VSP One Object Storage 打造現代化資料湖倉環境

可因應未來需求的資料湖倉儲存空間：支援 Amazon S3 Table



## 簡化的資料湖倉架構

整合儲存和分析功能，搭配內部部署原生 S3 Table 支援，無需外部服務即可簡化架構和營運

## 加速提供符合成本效益的分析

使用 Iceberg 等開放式表格格式就地查詢資料，無需將資料移動或複製到外部引擎或資料倉儲

## 與現代分析工具相容

S3 Table 和 Iceberg 目錄支援可確保與現代分析工具緊密相容，以更快速度直接從物件儲存中獲得見解

資料湖倉的管理相當複雜，需要面對儲存碎片化、查詢緩慢和成本高昂的資料移動等問題，導致阻礙效能、擴充性和整合，因此需要採用現代儲存解決方案加以因應。

現今的資料導向組織正面臨巨大壓力，需要從不斷增加的非結構化和半結構化資料中擷取資料。然而許多組織卻難以維持高效、可擴充和分析就緒的資料湖倉環境。

首先面對的挑戰就是基礎架構碎片化。資料通常分散在多個系統中，需要複雜的管道和昂貴的資料移轉，才能讓資料可供分析工具存取使用。以上各種效率不彰的因素，會延長獲得見解所需的時間，並增加營運成本。由於缺乏對現代表格格式的原生支援，客戶通常依賴外部中繼資料服務和自訂整合層，導致建立脆弱不堪的架構，難以維護和擴充。

效能則是另一個關鍵問題。隨著資料集持續成長，分析工作負載也日益頻繁複雜，傳統物件儲存解決方案已無法滿足需求，也就是無法支援對大規模資料進行快速、可靠、一致和準確的查詢處理或互動式分析。

安全性和治理也受到影響。由於系統之間並未連結，再加上不一致的原則，導致在執行資料生命週期管理、存取控制或法規遵循等方面變得繁瑣且容易出錯。

客戶不禁會問：如何在不重建架構的情況下打造現代化資料湖倉？

答案在於重新思考基礎，並從儲存開始著手。Apache Iceberg 或 Delta Lake 等現代物件儲存系統可原生支援與 S3 相容的表格，能夠正面應對以上挑戰。無需資料移動或外部服務，即可直接以高效能存取結構化資料，同時還能與現有分析工具和管道無縫整合。

成果如何？簡化運作、加速獲得見解，以及能夠因應未來需求的資料基礎架構。



## 使用現代物件儲存解決方案應對各項挑戰

原生 Amazon S3 Table 支援的現代物件儲存解決方案，是專為滿足資料湖倉工作負載需求而建。可在儲存層直接支援 Apache Iceberg 等開放式表格格式，因此不需要外部中繼資料儲存庫和複雜的資料移轉。這樣就能對結構化資料進行高效能的就地查詢，成為實現快速及可擴充分析的重要關鍵。

符合 ACID 標準的運作 \* 確保可靠的並行讀寫，對於持續演變的資料集而言，是維持一致性的基本要素。

與 Apache Spark、Presto 和 Trino 等熱門分析引擎無縫整合，除了簡化資料工作流程，也加快獲得見解。隨著湖倉環境不斷成長以支援 AI、ML 和即時商業智慧 (BI)，原生 S3 Table 支援可確保基礎儲存能夠跟上腳步，提供現代資料團隊所需的規模、簡單性和效能。這種新一代物件儲存解決方案將奠定基礎，協助打造更高效、更具成本效益及更靈活的湖倉架構。

\* 遵循正確性、一致性、隔離性、耐久性 (ACID) 等原則，確保以可靠方式進行讀取、寫入、更新和刪除等作業。



圖 1：現代資料湖倉基礎架構結合 S3 相容性、原生 S3 Table 支援、開放式格式和智慧服務，提供統一的儲存和分析功能，進而實現可擴充、工具整合和大數據效能。

## 讓資料湖倉在 VSP One Object 執行運作

Hitachi Vantara Virtual Storage Platform One Object (VSP One Object) 是一款現代企業級解決方案，旨在為資料湖倉環境提供無與倫比的擴充性、效能和簡易性。VSP One Object 完全相容於 S3，支援內部部署原生 S3 Table，並可直接與 Apache Iceberg 目錄整合，進而對結構化資料進行快速且符合 ACID 標準的就地分析，無需使用外部中繼資料儲存庫或複雜的資料移動。

組織可以輕鬆地從 TB 擴展到 EB 規模，進而支援長期資料成長和 AI/ML 工作負載。VSP One Object 可獨立擴充運算和儲存資源，進而達到最佳化效能和成本效益。可與 Apache Spark、Trino 和 Presto 等頂尖分析工具原生整合，加速獲得見解，而內建支

援資料複製、安全控制和物件層級存取原則，則可確保達到企業級可靠性和治理。

VSP One Object 以智慧功能為核心建構而成，自動偵測個人識別資訊 (PII) 等敏感資料，協助實現法規遵循並簡化資料分類。可針對大數據進行最佳化，提供低延遲存取，並支援混合作業負載 (結構化、半結構化和非結構化資料) 使其成為現代湖倉架構的理想選擇。VSP One Object 具有高持久性、強大的中繼資料管理和智慧資料服務，除了提供出色的運作效率，也能協助您從資料中獲得深入見解。無論您是管理 PB 規模的物聯網遙測資料，還是建立即時 AI 管道，VSP One Object 都能為您的資料湖倉策略奠定基礎，協助您因應未來需求。



## 特色

### 內部部署原生 S3 Table 支援

使用 Apache Iceberg 等開放式表格格式，能夠對結構化資料就地進行符合 ACID 規範的作業。無需使用外部中繼資料儲存庫或重寫資料。

無需移動資料即可直接在物件儲存執行快速、可靠的查詢

### 與 Apache Iceberg 用戶端無縫整合

可立即搭配使用 Spark、Trino、Presto 和其他工具，協助您加速獲得見解，並簡化湖倉環境的資料工作流程。

使用 Spark、Trino 和 Presto 等熟悉工具加速獲得見解

### 支援與 S3 相容的 API

完全支援 S3 API，可輕鬆與現有工具、雲端原生應用程式和資料管道整合，進而減少衝擊，並在混合雲或多雲環境中加速運用。

輕鬆與現有工具和雲端原生應用程式整合

### 架構可從 TB 擴充到 EB 規模

輕鬆支援不斷成長的資料量，範圍涵蓋先導專案到企業規模湖倉環境，確保一致的效能和容量，無需進行破壞性遷移或重建架構。

您可充滿自信擴充湖倉環境，無需擔心效能或容量限制

### 獨立擴充運算和儲存

允許組織獨立擴充分析能力和儲存容量，針對動態及高容量的資料工作負載打造最佳化效能和成本。

根據需求擴充資源以達到最佳化效能和成本

### Apache Iceberg 目錄整合

支援跨查詢引擎無縫使用 Iceberg 表格，實現模式演進、時間旅行，以及大規模的最佳化中繼資料作業。

透過可擴充的開放式表格支援簡化結構化資料管理

### 企業級複製和資料保護

透過彈性的複製、進階儲存服務、強大的一致性和安全的物件級存取控制，提供高可用性和資料保護。

使用內建複製功能確保高可用性

### QLC 和 TLC 支援

支援基於 QLC/TLC 架構的 VSP One Block 目標，並保證所有區塊完整無缺。

獲得您應得的性價比

## 摘要

VSP One Object 專為因應現代資料湖倉環境的各項挑戰而設計，透過 S3 相容性、原生 S3 Table 支援和內建 Apache Iceberg 目錄整合，無需複雜管道或外部服務即可實現快速的就地分析。

VSP One Object 以統一的儲存和結構化資料存取簡化湖倉架構，進而實現與 Spark 和 Trino 等工具的無縫整合。其能夠從 TB 擴充到 EB 規模，協助您獨立擴充運算和儲存資源，並智慧地偵測 PII 資料，進而確保效能、治理和成本效益。

VSP One Object 為組織提供可擴充的智慧型基礎，協助擬定現代化且因應未來需求的資料湖倉策略。

## Hitachi 經驗談：20 年物件儲存創新歷程

Hitachi 憑藉 20 年來備受肯定的創新成果和客戶信賴，成為企業級物件儲存領域的全球領導廠商。我們的 VSP One Object 是專為現今資料湖倉環境而建置的儲存平台，可提供現代分析所需的效能、擴充性和彈性。Hitachi 憑藉強大的專家支援和技術合作夥伴生態系，準備好協助組織輕鬆自信地徹底發揮資料的全部潛能。

準備好使用 VSP One Object 打造現代化資料湖倉環境嗎？

請與我們聯絡 →

## 專業代理商

### TERRAONE 拓壹科技

拓壹科技股份有限公司

公司電話：+886-2-2656-5700

服務信箱：service@terraone.com.tw

公司地址：114台北市內湖區內湖路一段360巷8號6樓

公司網址：www.terraone.com.tw

