

# 重複資料刪除白皮書

基於 Active Backup for Business 2.3.0



# 目錄

簡介	3
資料減量技術	4
重複資料刪除程序	
異動區塊追蹤 (CBT)	
重複資料刪除比	
總結	8

## 相關資訊

Synology 發佈多種支援文件。

您可以在 [技術支援中心](#) 查看實用的說明文章及常見問題，或觀看逐步影片教學。您也可以在此找到使用手冊、解決方案指南、型錄、白皮書等文件。進階使用者及管理員可以在管理員指南及開發者指南查看有關疑難排解及技術指引的相關資訊。

若仍無法從官方文件中找到問題的解決方法，您可以至 [Synology Community](#) 查看相關文章，使用者或技術支援人員會在此分享各種問題的解決辦法。您也可以透過網頁表單、電子郵件、電話以取得 [Synology 技術支援](#)。



# 簡介

無論使用何種儲存裝置或系統，資料遺失或系統故障皆會帶來極大的不便。工作電腦或伺服器等重要裝置需要堅實且有效率的備份策略，以利在自然災害或其他技術問題發生時，能夠協助您避免無法預期的系統故障。

**Active Backup for Business (ABB)** 讓您可以透過 Synology NAS 上的單一集中化控制台管理所有 VMware、Hyper-V、Windows、Linux、檔案伺服器的備份。適用於 Windows 裝置、虛擬機器、個別檔案的即時還原選項，讓您可在災害發生時維持重要服務的運作。此外，Active Backup for Business 採用**重複資料刪除**及**增量備份**等最新技術，可提升備份效能並將備份資料所需的實體儲存空間降至最低。有許多備份服務已普遍使用這些技術，因此我們的使用者經常詢問 ABB 重複資料刪除比所採用的技術與計算方式。

本文件將介紹部分重複資料刪除所使用的技術，並定義 Synology Active Backup for Business 採用的重複資料刪除比。

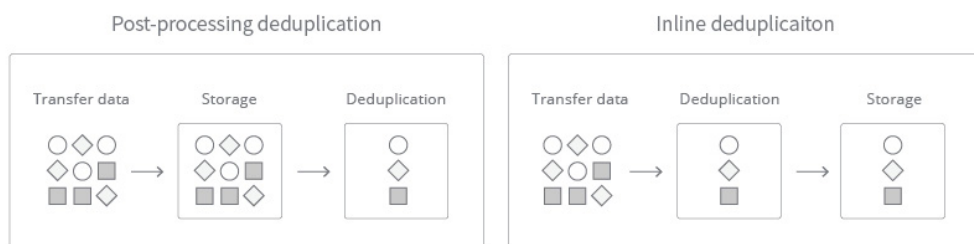
# 資料減量技術

資料減量技術可讓備份資料消耗較少的儲存空間，並最佳化儲存空間使用率，進而讓使用者能夠備份更多的裝置。Active Backup for Business 導入先進的資料減量技術以提升整體的備份效能。

## 重複資料刪除程序

重複資料刪除可移除重複的資料，藉此有效率地最佳化儲存空間的使用。重複資料刪除可區分為「在線處理」(in-line) 及「後處理」(post-process) 兩種類別。**在線重複資料刪除**會在資料寫入儲存空間之前便移除重複資料；**後處理重複資料刪除**僅會在備份程序完成後才移除重複資料。

您可以依據您的備份解決方案，在備份代理程式或備份目的地採用在線重複資料刪除。然而，透過代理程式執行重複資料刪除時，將會同時執行資料掃描、重複資料刪除、資料傳輸，因而導致備份過程消耗更多的時間並影響效能 (備份目的地)。



Synology **Active Backup for Business** 的設計加入了**在線重複資料刪除**的功能，且僅會在將資料寫入儲存空間之前，於目的地執行重複資料刪除。在備份來源上執行重複資料刪除通常會消耗過多的時間及 CPU 資源，而此方法即可避免此情形發生。此外，Active Backup for Business 讓使用者可將多種系統及裝置，備份至包含重複資料的單一目的地。

Active Backup for Business 同時採用**全域區塊重複刪除**技術，可移除多個目標之間的重複副本。全域區塊重複刪除讓 Active Backup for Business 可在單一共用資料夾內，刪除多個備份任務之間的重複資料，藉此可在不影響整體備份效能的情況下，最佳化實際的儲存空間消耗。

Active Backup for Business 的重複資料刪除機制會使用專屬識別碼來分類每個備份的檔案及檔案區塊。Active Backup for Business 的重複資料刪除引擎會將備份檔案分割為單一 **4 KB** 的個別檔案區塊，並為其指定專屬識別碼，再藉由重複資料刪除引擎所使用的 **SHA-256 演算法**來運算該識別碼，藉此辨識每個檔案區塊。

為檔案區塊生成識別碼後，系統會為每個已儲存的檔案區塊建立識別碼索引。此索引將會用來配對未來產生的檔案區塊與其對應的已儲存檔案區塊。成功辨識檔案區塊後，重複資料刪除引擎會移除重複的資料以最佳化裝置的儲存空間使用率。

## 異動區塊追蹤 (CBT)

**異動區塊追蹤 (CBT)** 是一種增量備份技術，常用於一些不同的備份解決方案。**增量備份**可僅備份自最近期的檔案版本以來，新增或曾經被修改的檔案。**CBT** 可透過增量備份僅備份有變更的資料區塊，而非每次執行完整備份，藉此可減少時間及資源的消耗。藉由此方法，可讓最終透過應用程式傳輸的資料量 (或稱**備份資料傳輸大小**) 小於來源的原始資料量。

Active Backup for Business 已在其架構導入此功能，可用於備份電腦、實體伺服器、虛擬機器。ABB 會在任務中執行永久增量備份以最大化可備份版本的數量，並在執行完整備份後，僅備份修改的檔案，藉此盡可能地減少檔案版本所使用的儲存空間容量。



## 重複資料刪除比

### 定義

**重複資料刪除比**為用來評估備份解決方案其重複資料刪除效能的數據。使用者可藉由此數據分析各供應商的相同功能及效能，為使用者選擇備份解決方案的決定性因素。然而，重複資料刪除比會依使用的刪除技術及資料本身的特性而異。

**重複資料刪除比或壓縮比**為：計算欲儲存的資料與其在移除重複的資料後，在備份目的地上實際用來儲存備份資料的儲存空間之比例。

例如：從裝置備份 100 GB 的資料，而備份儲存伺服器上實際使用的儲存空間為 20 GB。將備份資料 (100 GB) 除以備份儲存伺服器上使用的儲存空間 (20 GB)，則重複資料刪除比為 5:1，而重複資料刪除率為 80%。

重複資料刪除比被廣泛的使用在市場行銷以宣傳備份產品的能力，例如：廠商會以備份產品的重複資料刪除比優於市場上的其他產品 20 倍來作為宣傳。然而，重複資料刪除比會依據所使用的重複資料刪除技術、資料的一般特性、所使用的備份方法、重複資料刪除比的計算方式而異。

基於上述因素，透過重複資料刪除比來比較不同的產品可能並非最準確的方式。因此，需留意每個產品重複資料刪除比的計算方式，以及在備份目的地上實際使用的儲存量。

## 計算方式

您可以透過以下數值來計算重複資料刪除比：

- 移除重複資料前的備份資料總容量
- 傳輸的備份資料容量
- 備份目的地上實際使用的容量

一般而言，計算比例的較佳方法為將**傳輸的備份資料容量**除以**備份目的地上實際使用的容量**。  
Synology 即使用此方法。

部分備份供應商計算比例的方法為將**移除重複資料前的備份資料總容量**除以**備份目的地上實際使用的容量**，此方法所得的重複資料刪除比可能未必準確。

試想以下情境：

- 備份解決方案：A
- 移除重複資料前的備份資料總容量：6 TB
- 傳輸的備份資料容量：1.2 TB
- 備份目的地上實際使用的容量：0.83 TB
- 廣告中宣傳的重複資料刪除比：7.2:1

在此情境中，備份解決方案 A 的廣告中說明其重複資料刪除比為 **7.2:1** (6 TB / 0.83 TB)。然而，此計算方法未必完全正確。此範例僅傳輸 **1.2 TB 的備份資料**，透過 CBT 刪減重複的資料後，再將資料壓縮至 **0.83 TB** 以儲存在備份目的地。

**備份解決方案 A** 將備份資料的總容量除以備份目的地上實際使用的容量，以取得較高的重複資料刪除比。但實際上，只有**傳輸的備份資料容量**才能夠顯示經過重複資料刪除的實際資料量。

使用**傳輸的備份資料容量**除以**備份目的地上實際使用的容量**才能夠取得更加準確的計算結果。因此，可將重複資料刪除的計算方式調整如下：

### 重複資料刪除比

傳輸的備份資料 容量	÷	備份目的地上實際使用的 容量	=	重複資料刪除比
1.2 TB		0.83 TB		1.44:1

### 重複資料刪除率

(	傳輸的備份資 料容量	-	備份目的地上實際使用 的容量	)	÷	傳輸的備份資 料容量	x 100%	=	重複資料刪除 率
	1.2 TB		0.83 TB			1.2 TB			30.8%

## 備份方法

另一項可能會影響重複資料刪除比的因素為**備份方法**，特別是**完整備份**或**增量備份**。相較於增量備份，僅執行完整備份的備份系統其重複資料刪除比可能會較佳。這是因為增量備份僅傳輸經過變更的資料區塊，並在傳輸前便移除重複的資料，因而可降低在備份目的地上實際使用的儲存容量。反之，相較於原始或先前的備份版本，完整備份會包含較多的重複資料。

簡言之，即使增量備份解決方案比起完整備份可能無法呈現較佳的重複資料刪除比，但在備份目的地上仍可消耗較少的實際儲存容量。

## Active Backup for Business 重複資料刪除比

如前述，使用**移除重複資料前的備份資料總容量**來計算重複資料刪除比可得出較佳的重複資料刪除比結果，但可能無法正確反映經過重複資料刪除的實際資料量。Active Backup for Business 則採用**傳輸的備份資料容量**來進行計算，以呈現經過重複資料刪除的實際資料量。如下方表格所示，Active Backup for Business 採用的重複資料刪除技術，**在備份目的地上實際使用的儲存容量較少**，證明此方法整體而言可為使用者節省更多的儲存空間。

綜上所述，使用者可以更加著重於傳輸的備份資料容量實際占用的儲存容量，而非產品的重複資料刪除比，以更加了解產品能夠如何協助您節省儲存空間。

下方表格將提供 **Active Backup for Business** 的實際重複資料刪除比，並將其與其他備份服務供應商進行比較。

供應商	傳輸的備份資料容量	備份目的地上實際使用的容量	重複資料刪除比	重複資料刪除率
Synology Active Backup for Business	1.2 TB	637.2 GB	1.2 TB / 0.62 TB = 1.93:1	48.18%
供應商 A	1.2 TB	819.8 GB	1.2 TB / 0.8 TB = 1.5:1	33.33%
供應商 B	1.2 TB	847.7 GB	1.2 TB / 0.82 TB = 1.45:1	31.08%

備份來源包含：Windows 電腦 x 9、虛擬機器上的 Windows Server x 1



## 總結

重複資料刪除是一個複雜的過程，能夠協助使用者減少重複的資料量，以便更有效率地管理備份並最佳化儲存空間。重複資料刪除是許多備份解決方案常用的技術，因此許多使用者會希望了解重複資料刪除比以比較不同的產品。然而，有許多種可用的計算方法和不同的因素可能會影響重複資料刪除的評估，因此每個產品提供的比率數值可能並非可靠。據此，在比較不同產品的重複資料刪除比時，重要的是著重於經過重複資料刪除後，用來儲存備份資料的實際實體儲存容量。

**Synology Active Backup for Business** 不僅提供適用於多種平台的全方位備份解決方案，且其架構可實現快速、可靠、有效率的重複資料刪除，協助使用者節省時間並充分善用儲存空間。





**SYNOLOGY  
INC.**

新北市板橋區  
遠東路 1 號 9 樓  
台灣  
電話：+886 2 2955 1814

**SYNOLOGY  
AMERICA CORP.**

3535 Factoria Blvd SE, Suite #200,  
Bellevue, WA 98006  
USA  
電話：+1 425 818 1587

**SYNOLOGY  
UK LTD.**

Unit 5 Danbury Court, Linford  
Wood, Milton Keynes, MK14 6PL  
United Kingdom  
電話：+44 (0)1908048029

**SYNOLOGY  
FRANCE SARL**

102 Terrasse Boieldieu (TOUR  
W)  
92800 France France  
電話：+33 147 176288

**SYNOLOGY  
GMBH**

Grafenberger Allee 295  
40237 Düsseldorf  
Deutschland  
電話：+49 211 9666 9666

**SYNOLOGY  
SHANGHAI**

200070 中國上海市  
靜安區天目西路 511 號輔  
房 201 室  
中國

**SYNOLOGY  
JAPAN CO., LTD.**

4F, No. 3-1-2, Higashikanda,  
Chiyoda-ku, Tokyo, 101-0031  
Japan

**Synology®**



[synology.com](https://synology.com)

Synology 可能隨時修改產品規格與說明，恕不另行通知。Copyright © 2022 Synology Inc. 保留一切權利。®  
Synology 及其他群暉科技股份有限公司 (Synology Inc.) 所有產品之名稱，均為群暉科技股份有限公司所使用或註冊  
之商標或標章。本軟體產品所提及之產品及公司名稱可能為其他公司所有之商標。