



Microsoft  
**SQL Server** 2012



# SQL Server 2012 Parallel Data Warehouse

解決方案簡介

發佈日期：2013 年 2 月 22 日

# 目錄

介紹 .....	1
Microsoft 平台：Windows Server 和 SQL Server .....	2
SQL Server 2012 Parallel Data Warehouse .....	2
適用於海量資料 .....	3
隨規模增長的新一代效能 .....	4
取得最佳價值的工程設計 .....	4
結論 .....	5
開始行動 .....	5

## 重點提要

- 效能比舊式資料倉儲查詢快高達 100 倍
- 查詢效能比之前舊版 Microsoft SQL Server Parallel Data Warehouse 快高達 50 倍
- 資料壓縮高達 15 倍
- 存放裝置容量多高達 70%
- 資料載入速率比舊版快達兩倍
- 資料容量需求從最小 (1 TB) 到最大 (8 PB) 變動不同
- 使用最大密度存放裝置和改造過的架構，磁碟使用量減少 75%
- 業界最低每 TB 價格

## 客戶可享的重要優點

- 運用 xVelocity 可更新 Columnstore 技術比從前更快速深入了解
- 從簡單易用且熟悉的工具，如 Microsoft Excel，同時查詢兩種平台—Hadoop 的關聯式和非關聯式資料
- 輕鬆不費力地擴充、避免大規模升級，並在極少干擾的情況下順利擴展為多個 PB
- 從小規模開始，並以模組方式視需要成長，將風險降至最低並避免過度投資
- 在重要諮詢的協助及 Microsoft Services 的支援下，確保部署成功

## 介紹

今天，資料是企業的新通貨。這表示組織若知道如何讓資料發揮最大價值就能繁榮昌盛。Gartner<sup>1</sup> 斷言：「到 2015 年，組織將高價值、多樣化的新資訊類型和來源整合在連貫一致的資訊管理基礎架構中，可在財務表現上凌駕業界同行超過 20%。」McKinsey<sup>2</sup> 也同意 Gartner 的看法，確認利用資料和商務分析來推動決策的組織將會比其他組織獲得更高的生產力，並實現更高的淨值報酬率。

現在是資料世界前所未有的蓬勃發展時機。企業看著新資訊類型和來源匯整，讓產業從根本徹底轉型。例如，一般組織必須管理的資料量急劇暴增，大部分增長的資料都來自未結構化資料，完全超出傳統資料類型範疇之外。目前的軟體解決方案無法應付這種資料的複雜性，或者因資料而大量增加的工作負載。轉移至支援非傳統資料來源 (如海量資料) 的後端解決方案是能夠有效管理這種資料的唯一方式，但是由於設備侷限於特定廠商、部署成本高昂，以及轉移作業曠日費時，而使得組織無法進行這項轉移。

為了協助各方組織順利轉移至新的資料世界，Microsoft 推出 SQL Server 2012 Parallel Data Warehouse (PDW)，這是新一代向外擴充的大型平行處理 (MPP) 資料倉儲應用裝置。立基於業界標準資料庫和市售硬體，PDW 解決方案為客戶提供通用型資料平台，可發揮釋出其資料的價值。簡言之，SQL Server 2012 PDW 提供客戶高階解決方案，而沒有高階軟體價格的負擔。

<sup>1</sup>Casonato, Regina 等合著《21 世紀的資訊管理》(英文) Gartner，2011 年 9 月 2 日  
[http://download.microsoft.com/download/7/B/8/7B8AC938-2928-4B65-B1B3-0B523DDFCDC7/Big%20Data%20Gartner%20information\\_management\\_in\\_the\\_21st%20Century.pdf](http://download.microsoft.com/download/7/B/8/7B8AC938-2928-4B65-B1B3-0B523DDFCDC7/Big%20Data%20Gartner%20information_management_in_the_21st%20Century.pdf)

<sup>2</sup>《數位商務展望》(英文) McKinsey & Company，2012 年 1 月  
[http://www.mckinsey.com/~media/mckinsey/dotcom/client\\_service/BTO/PDF/MCBT\\_Compendium\\_Perspectives\\_on\\_Digital\\_Business.ashx](http://www.mckinsey.com/~media/mckinsey/dotcom/client_service/BTO/PDF/MCBT_Compendium_Perspectives_on_Digital_Business.ashx)

# Microsoft 平台： Windows Server 和 SQL Server

Microsoft 提供完整的解決方案技術套件，能啟用可互通的端對端基礎架構。這有助 IT 部門達到最大投資報酬率，並能以低擁有總成本 (TCO) 創造關鍵任務環境。

Microsoft 以 Windows Server 2012 提供已增強的 Hyper-V 虛擬化、儲存空間、執行階段記憶體設定及其他，徹底改變了伺服器作業系統的定義。

Microsoft 的最佳資料平台，SQL Server 2012 能夠創造關鍵任務環境，以低擁有總成本 (TCO) 提供可用性和效能。客戶也能因 SQL Server 2012 的其他企業級功能而受益，包括 xVelocity 和自動記憶體微調。

Microsoft 一直都是全心致力於提供同級最佳企業解決方案，從諮詢服務到修補到系統內整合，甚至使用不同工具，如 SQL Server Data Tools、SQL Server Integration Services 或其他商業智慧 (BI) 工具。無論何種解決方案，Microsoft 皆全權提供，且幾乎以內建功能提供，不需額外成本。

## SQL Server 2012 Parallel Data Warehouse

SQL Server 2012 PDW 已演進發展成可完全掌控運用新的資料世界。它可向外擴充以處理任何大小的虛擬需求、針對來自 Hadoop 的未結構化和半結構化資料集順暢無阻地執行分析，並提供新一代查詢效能改進，高達舊式資料倉儲部署的 100 倍。

部署簡易的 SQL Server 2012 PDW 是以預建的應用程式裝置提供，其中包含軟體、硬體和已預先安裝的網路元件。SQL Server 2012 PDW 也設計為提供最佳價值，透過提供今日市場上任何高階倉儲應用裝置每 TB 最低價格達成。

### 重要設計元素



- Microsoft Windows Server 2012 儲存空間和 Hyper-V 具有更高密度存放裝置和運算能力
- 提升速度和擴充性，所需運算資源更少，且 IT 所需投入精力更低
- 高達 15 倍的資料壓縮可節省 70% 的存放裝置需求
- 硬體成本減少 50% 且能源消耗也降低 50%

## 價值主張

**適用於海量資料：**使用 PolyBase 能夠跨 Hadoop 和關聯性資料達到整合式查詢，是資料處理引擎的根本大突破

**隨規模增長的新一代效能：**針對幾乎任何規模的資料集提供新一代效能，速度比從前快高達 50 倍，以包含新版可更新 xVelocity Columnstore 的主存放裝置引擎運作，並讓使用者能夠從小規模 TB 需求線性化向外擴充到 PB 規模資料

**取得最佳價值的工程設計：**透過軟體創新手法，包括 xVelocity Columnstore、PolyBase、Windows Server 2012 Hyper-V 和儲存空間，重新定義 PDW 應用裝置，同時提供監控套件，以利與現有 Microsoft System Center 部署進行整合

「我們擴充目前的 SQL Server 2012 商業智慧版實作納入 SQL Server PDW，增強了自助式資料入口網站，可容納大量資料集以及每日銷售點資料，提供來自客戶銷售量和行銷努力而隨需求變動的結果。SQL Server PDW 透過我們的自助式入口網站提供快速隨選資料存取，將提供客戶極大商機，可善加運用對購買人的深入了解，融入策略規劃及策略手法操作中。」

-Jim Norred -

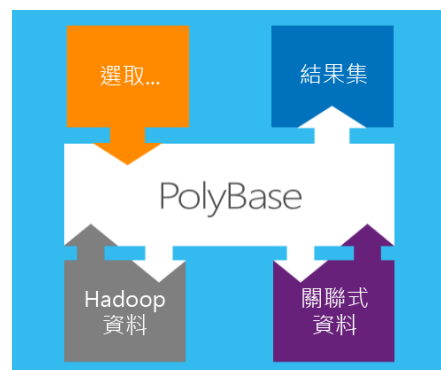
CROSSMARK 執行副總裁兼資訊長

<http://www.microsoft.com/en-us/news/Press/2012/Nov12/11-13CROSSMARKPR.aspx> (英文)

## 適用於海量資料

SQL Server 2012 PDW 使用 PolyBase，為客戶提供跨 Hadoop 和關聯性資料的整合式查詢，而且也能與 Microsoft BI 整合，是資料處理引擎的大突破。

**在單一查詢中順暢無間地整合 Hadoop 資料與資料倉儲：**PolyBase 可接受標準的 Transact-SQL 查詢，結合關聯式來源資料表與參考非關聯式來源的 Hadoop 叢集資料表，順暢無間地傳回結果給使用者。IT 不需要從 Hadoop 將資料事先載入倉儲之中，使用者也不需要學習 MapReduce 就能進行查詢。這樣能夠賦予組織能力，針對任何速度、複雜性和大小的資料，輕鬆簡單地執行互動式分析。PolyBase 中也提供匯出選項，讓組織能夠將歷史網路記錄檔儲存至低成本市售硬體的 Hadoop 叢集目錄中，然後在執行查詢時指定目錄位置，利用 PolyBase 進行處理。



**BI 整合：**原生 Microsoft BI 整合作業可讓使用者迅速創造吸引人的視覺效果，可使用如 Microsoft Excel 之類熟悉的工具，從幾乎是任何資料類型 (包括關聯性或非關聯性) 作成重要商業決策。Power View 是內建於 Excel 之中，可讓使用者迅速簡單地將資料視覺化，同時在熟悉的 Microsoft SharePoint Server 環境中，分享商務洞察力並進行協同合作。除此之外，Power View 也不需要建置



Cube，讓組織能夠查詢更多資料而不受任何表格式模型限制束縛。

## 隨規模增長的新一代效能

為了提供隨規模增長的新一代效能，SQL Server 2012 PDW 使用新的可更新版 xVelocity Columnstore 結合 MPP 引擎，以利順暢無間的線性向外擴充。

**xVelocity Columnstore：**SQL Server 2012 PDW 使用最新版 xVelocity 記憶體最佳化 Columnstore，兼具可更新和叢集化特性。將資料從傳統資料列存放區重新調整為欄，組織就能將查詢時間從小時減為秒，看到新一代的效能（多達 50 倍）及高度資料壓縮（多達 15 倍）。這是由於將資料集組成單欄格式（包含多個資料列的值），可同時讓存放裝置和傳回的資料集達到更高效率，最終達到降低成本並讓現有投資取得更高報酬。能夠以 xVelocity 的可更新 Columnstore 永久轉換傳統式資料存放區資料表是另一項新功能，可大幅縮小存放裝置（高達 70%），同時提供即時資料倉儲查詢效能。

**大型平行處理：**SQL Server 2012 PDW 使用 MPP 向外擴充引擎，提供強大的分散式運算和規模擴充。這項應用裝置使用分佈式資料庫架構配合多個實體節點，其中各個節點都以專屬 CPU、記憶體和存放裝置，各自執行專有的 SQL Server 執行個體。啟動查詢時，是跨節點平均分散並同時在各個實體節點上執行。MPP 架構的優勢在於透過平行執行，隨規模提供最高效能。使用者只需要增加更多容量，就能繼續不斷地向外擴充這項實作到高達以 PB 為單位的資料，並能處理複雜且同時進行的查詢，而不需要將現有資料模型簡維化，也不會危及業務關係。此外，實作也能納入高並行處理功能，並處理大規模的複雜查詢。SQL Server 2012 PDW 提供經過實

證的全方位擴充性，能達到更高的效能和更低的 TCO，而且資料載入速度更快可提供卓越的趨勢分析。客戶不需要因顧慮效能而限制資料，能夠隨時提出任何問題而不會損及效能或遭遇載入瓶頸。

**增加容量輕而易舉：**擴充傳統式資料倉儲時，一旦裝載資料倉儲的伺服器超出最大實體容量和效能，組織就必須重新建置所有伺服器。在 MPP 向外擴充架構下，SQL Server 2012 PDW 將可向外擴充至客戶所面臨任何或所有的未來資料需求，從最小的需求（0–15 TB）開始一直到最大的需求（8 PB）。有了 MPP 架構，組織只需購買所需，然後添加運算節點機架擴充規模，可避免大規模的升級作業。例如，從一個機架的基本系統開始。裝滿一個機架可容納 10 個節點，而另外再加一個 10 節點機架可增加到總共 40 個節點。

## 取得最佳價值的工程設計

**輕鬆設定：**SQL Server 2012 PDW 是預建的應用裝置，以完整封裝的應用程式解決方案直接運送到客戶門口。客戶不需要擔心建置最理想的硬體組合（就 CPU、存放裝置和記憶體而言），也不需要精確了解安裝及調整軟體的專門知識；只需要插上應用裝置，就能開始將特定資料整合至解決方案之中。

SQL Server 2012 PDW 的基礎架構結合自動化功能，降低了人為操作的需求。PDW 提供資料庫管理員簡單的資料庫和資料表定義程序，讓他們不再需要煞費苦心地仔細進行空間、記憶體或快取管理。最後，System Center 也納入套裝產品之中，簡化了整個應用裝置（包括硬體、作業系統、倉儲和應用程式）的 IT 管理作業，不需要額外的成本就能以簡單的工具進行管理。

**內建軟體功能使硬體完全發揮功效：**SQL Server 2012 PDW 透過軟體創新手法，重新改造應用裝置

所需的硬體規格，能夠提供最佳價值給客戶。PDW 透過 Windows Server 2012 儲存空間，其中包含內建的可靠性和效能功能，提供客戶極其優異的價值。儲存空間可確保應用裝置使用高密度直接連接磁碟的經濟實惠硬體選擇，有助大幅降低存放裝置成本。

Windows Server 2012 中提供的 Hyper-V 虛擬化功能，和立基於同質性節點的硬體設計，完全不需要舊有架構的「控制機架」，因而大幅降低硬體使用量和應用裝置的成本。

**業界每 TB 最低內部部署價格之一：**唯有 SQL Server 2012 PDW 這套解決方案，能夠繼續增加價值同時降低取得完整解決方案的入門成本。Microsoft 降低成本的手法包括透過虛擬化減少硬體使用量、利用軟體提供具彈性而可擴充的效能功能，以及運用壓縮 (高達 15 倍) 降低存放裝置使用量。這些功能全部都能讓 SQL Server 2012 PDW 達到每 TB 最低價格，而大幅領先其他所有公司。

**硬體廠商的選擇：**HP 和 Dell 以真正的四分之一機架、半機架和全機架解決方案，提供最能配合客戶基礎架構需求的硬體和支援選擇。

## Dell 平行資料倉儲裝置 HP 企業資料倉儲裝置



	Dell	HP
伺服器	Poweredge R620	ProLiant Gen8 DL360
運算節點	每個機架最多 9 個節點 (最少 3 個)	每個機架最多 8 個節點 (最少 2 個)
機架	¼ 到 6 個	¼ 到 7 個
原始磁碟容量 (未壓縮)	0TB-1.2PB	0TB-1.2PB

SQL Server 2012 PDW 讓您將是否可在數秒內查詢所有資料甚至更多的執行動作視覺化。

## 結論

SQL Server 2012 PDW 已有充分準備，能以低廉實惠的價格提供客戶革命性的效能改進，應付客戶所面臨的嚴苛挑戰，順暢無間地整合結構化和未結構化資料。有了 SQL Server 2012 PDW 及其增強的 PolyBase 和 xVelocity 技術，客戶能夠比從前更快速地將資料轉變成極具價值的洞察力，而只需投入最低限度的資本。

## 開始行動

如需 PDW 應用裝置的詳細資訊，請前往 <https://www.microsoft.com/sqlserver/en/us/solutions-technologies/data-warehousing/pdw.aspx> (英文)。