

Microsoft 超融合架構下軟體定義儲存服務效能評估之研究

Performance Evaluation of a Hyper Converged Architecture Based on Microsoft Storage Spaces

¹ 林聰武

² 呂嘉祐

¹Tsong-Wuu Lin

² Chia-Yu Lu

¹ 東吳大學資訊管理學系

¹ Department of Computer Science and Information Management
School of Business
Soochow University

² 東吳大學資訊管理學系碩專班

² Department of Computer Science and Information Management
School of Business
Soochow University

摘要

現行 IT 基礎架構所需的各種資源，幾乎能夠完全由伺服器搭配虛擬化平臺軟體來提供，不僅只是軟體定義運算(Software Defined Compute, SDC)相關虛擬處理器與記憶體應用，而且軟體定義儲存 (Software Defined Storage, SDS)也能夠架構在裸機伺服器 (bare metal)，或是整合在伺服器虛擬化平臺的環境、以超融合架構 (Hyper-Converged Infrastructure) 之姿提供儲存需求，企業為了降低建置儲存環境的成本，直接使用伺服器主機與內含磁碟資源，以軟體定義的方式，打造出具備高可用性與擴展性的軟體定義儲存架構，是近年來炙手可熱的一個儲存應用形式。

Microsoft Windows Server 2016 中提供了 Storage Spaces Direct (S2D)，這種直接內含於作業系統核心的超融合架構分散式儲存系統，因應效能的需求而加入企業等級非暫時性寫入快取 (Non-Volatile Write Cache, NVME) SSD 固態硬碟做為資料快取 I/O 交換。本論文將針對實際建置在不同功能應用面的儲存運作效率與效能進行分析與說明，提供給企業評估後續真實生產環境做規劃參考。

關鍵字：Storage Spaces Direct, 軟體定義儲存, 高可用性, 超融合架構

Abstract

The resources required by the current IT infrastructure can be almost completely provided by the server with virtualization platform software. Not only Software Defined Compute (SDC) related virtual processors and memory applications, but also Software Defined Storage (SDS) is an environment that can be configured on bare metal servers or be integrated into a server virtualization platform, providing storage requirement in the form of a Hyper-Converged Infrastructure. It is a hot topic for enterprises reducing the cost of building a storage environment by the server host and embedded disk resources, software-defined approach, to create a high-availability and extensibility of the software-defined storage architecture.

The Microsoft Windows Server 2016 provides the Storage Spaces Direct (S2D), a hyper-convergence architecture distributed storage system that is directly embedded in the core of the operating system and adds enterprise-level non-transient write caches in response to performance requirements. (Non-Volatile Write Cache, NVME) SSDs are used as data cache I/O exchanges. We will analyze and explain the storage operation efficiency and performance of different functional applications and provide enterprises with an assessment of the actual production environment and planning reference.

Keywords: Storage Spaces Direct, Software Defined Storage, High Availability, Hyper Convergence Architecture

1. 前言

近年來，IT 的成長與創新大都與行動裝置、雲端服務、社交網路與大資料這四個應用有相當的關係，IDC[1]將這些稱之為第三平臺（Third Platform），Gartner[2]則稱為力量連結（Nexus of Forces）。同時，24×7 不間斷的全球化業務進行，已經成形多年，而且，商業智慧與預測性資料分析（大資料）受到廣泛採用、應用深度大幅增加，再加上感應器與相關智慧型應用開始盛行，紛紛以物聯網的形態進入到現實世界，涵蓋範圍從後端的機房環境到個人端的穿戴式消費電子裝置。這種種因素，都驅使資料中心的維運和經營需有所突破，才能接下這些挑戰。

以資料中心的儲存資源管理而言，雖然伺服器虛擬化的應用在企業環境當中，已經相當普及，然而是否能發揮整合這類型環境的應有功能，仍有許多內外條件需要配合。就技術的角度來說，隨著接下來企業 IT 勢必要走向軟體定義儲存式環境的潮流，如何提升儲存效率與效能仍是重要議題。

現行資料中心對於儲存資源應用，陷入缺乏整合的困境，想要做到儲存資源集中管控，困難度很高，因此軟體定義式儲存（Software Defined Storage，SDS）的發展正是時候。

2. 主要內容

2.1 SDS 軟體定義儲存

根據 Gartner 的分析及統計顯示:在 2016 年時,市場中儲存設備產品僅有「15%」採用「軟體定義」的方式提供儲存服務、預估在 2019 年時比例將會提升至 50%。

另外根據 IDC 的分析及統計顯示:在 2016 年時,企業及組織內整合 SDS 儲存資源至 IT 基礎架構的比例約「10%」採用,預估在 2020 年時比例將會提升至 70-80%如下圖 1 所示。

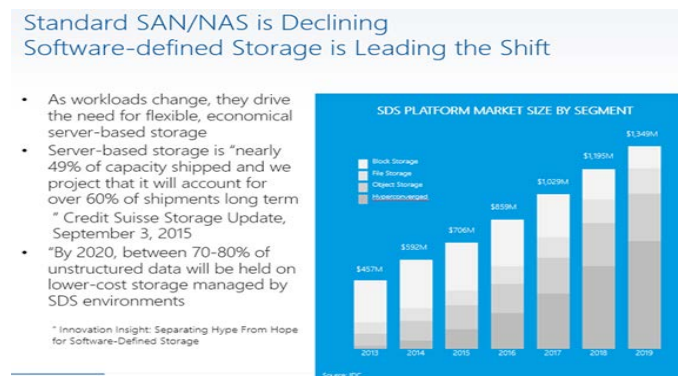


圖1 IDC SDS研究報告

軟體定義儲存(Software Defined Storage, SDS),為企業及組織帶來儲存資源的潛在好處,便是能夠提升零活性並降低整體維運成本(Total Cost of Ownership, TCO)。

因此 SDS 軟體定義儲存解決方案,應具備抽象化特定實體儲存設備或虛擬裝置的限制,提供更高的敏捷性、Qos 儲存資源管控機制、降低成本。

2.2 Microsoft SDS 軟體定義儲存技術

Windows Server 2016 Storage Space Direct (S2D)[3]新世代 SDS 軟體定義儲存解決方案,可以直接將叢集當中的叢集節點本地端儲存資源,透過網路互相連結後組成一個龐大的儲存資源池,達成更具彈性的 SDS 軟體定義儲存運作架構。

S2D 技術簡介,簡單來說,就在於 S2D 軟體定義儲存技術,可以將多台伺服器的「本機硬碟 (Local Disk)」結合成為一個大的儲存資源池如下圖 2 所示。

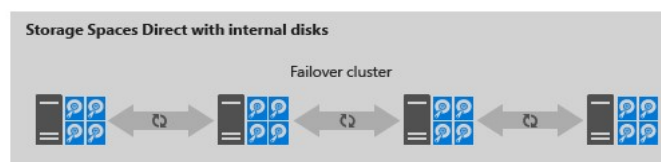


圖 2 S2D 軟體定義儲存技術支援本機硬碟運作架構示意圖

在 S2D 軟體定義儲存技術中，整個運作架構包含了 SOFS(Scale-Out File Server)、CSVFS(Clustered Shared Volume File System)、Storage Spaces、Failover Clustering 等技術。那麼，讓我們來看看在相關運作層級中，每個層級所專司的運作角色以及功能為何如下圖 3 所示。

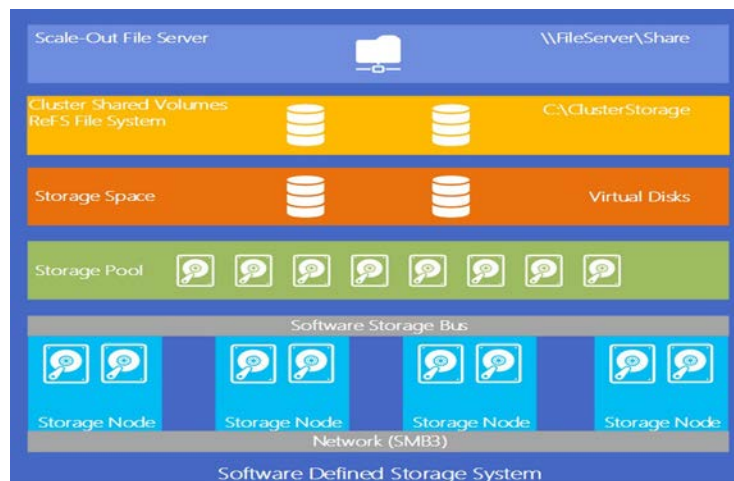


圖 3 Windows Server 2016 Storage Spaces Direct 運作架構示意圖

2.3 SMB Direct (RDMA)

SMB (Server Message Block) [4]為網路檔案共用通訊協定，允許電腦上的應用程式讀取和寫入檔案，SMB 通訊協定可以作用於 TCP/IP 通訊協定或其他網路通訊協定上。因此，允許應用程式讀取、建立、更新遠端伺服器上的檔案，簡單來說您習慣從網路上的芳鄰存取檔案伺服器上的檔案便是採用 SMB 通訊協定。

在微軟新世代 Windows Server 2016 雲端作業系統當中，可整合 SMB over RDMA (Remote Direct Memory Access) [5]特色功能，當實體主機採用支援 RDMA 技術 (iWARP、InfiniBand、RoCE) [6][7]的網路介面卡時、SMB Client 與 SMB Server 主機之間，將會採用記憶體到記憶體 (Memory to Memory) 方式進行資料傳輸，所以能夠最大程度降低伺服器 CPU 工作負載和延遲時間，達成存取遠端伺服器資料時類似存取本機資料一樣。

3. 系統設計

3.1 實驗環境建置

業界一般在評論儲存效能時，通常會以 IOPS (Input/Output Operations Per Second) 儲存效能數值表現。因此，本研究實驗環境 IOPS 效能測試工具則建議使用 Microsoft 原廠所提供的 VMFleet 儲存效能壓力測試工具。

總體來說，透過 VMFleet 儲存效能壓力測試工具，可以在 S2D 的 HCI 超融合部署架構中，透過已經撰寫好的 PowerShell 指令碼，呼叫大量的 VM 虛擬主機透過 Diskspd[8]同時執行 IOPS

工作負載壓力測試，並且在進行儲存效能壓力測試時同時觀察 IOPS 的效能表現情況，收集相關數據結果。

測試環境 (4 Nodes S2D Cluster)，4 Nodes S2D Cluster 測試環境中，每台 S2D 叢集節點主機的硬體配置如下列表 1：

QuantaGrid D51BP-1U		
SKU	Description	Qty per unit
Server Platform	D51BP-1U	1
CPU	Intel E5-2698 v4 2.2 GHz *2 (每顆 20 Cores / 40 執行緒)	2
Memory	32G DDR4 RDIMM 2133Mhz	16
Cache	Intel P3700 NVMe SSD 400GB	2
Storage	Samsung SM863 SATA SSD 960GB	8
Boot Drive	SATA-DOM 128GB	1
HBA Card	QCT LSI 3008	1
NIC Card	RoCE OCP Mezz CX3 Pro 40G Single Port (QSFP+)	1
OS	Windows Server 2016、S2D (Storage Spaces Direct)、High Performance Power Plan	1

表 1 S2D 叢集節點主機硬體規格配置

3.2 實驗測試流程

測試工具

(一)、 VMFleet(建立IOPS效能測試壓測環境)

每台S2D叢集節點主機需求建立20個IOPS效能測試壓測用途VM虛擬主機

(二)、 DISKSPD(執行IOPS儲存效能壓力測試)

企業應用程式服務壓測情境模擬

企業或組織所會用到的服務或應用程式類型通常為Random I/O存取類型，因此在實驗測試環境中，列舉三種企業最常應用的程式服務來進行IOPS效能壓力測試。

- A. 100% Random、100% Read、0% Write (EX.Web Server Application)
- B. 100% Random、90% Read、10% Write (EX.File Server Application)
- C. 100% Random、67% Read、33% Write (EX.Exchange Email Server Application)

每次實驗測試流程如下圖4所示。

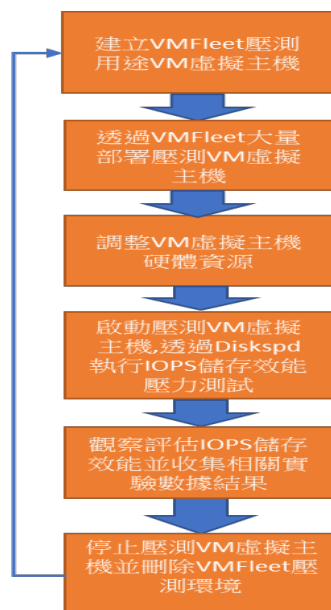


圖 4 實驗測試部署主要流程

3.2.1 RDMA 效能實驗環境測試

新一代 Windows Server 2016 雲端作業系統中，此次研究實驗需要透過 S2D (Storage Spaces Direct)技術建構 SDS 軟體定義儲存環境，並且強調支援「RDMA (Remote Direct Memory Access)」技術，能夠有效「降低 S2D 叢集節點 CPU 工作負載」同時「降低延遲時間」。

此次研究實驗採用同一批軟硬體環境，並且進行「RDMA Enabled」以及「RDMA Disabled」的測試，同時整理出 2 種運作環境的儲存效能，以便讓我們能夠了解這 2 種運作環境之間的效能差異。

3.2.2 多節點效能實驗環境測試

S2D 叢集建立在不同數量的叢集節點時(2 Nodes、3 Nodes、4 Nodes)，由於 S2D 俱備橫向擴充(Scale-Out)的能力，所以不同的叢集節點數量也會造成不同的 IOPS 儲存效能結果。因此，S2D 叢集能因為叢集節點的數量提升，帶來相對應的 IOPS 儲存效能表現提升。

此次研究實驗採用同一批軟硬體環境，並且進行 2 Nodes、3 Nodes、4 Nodes 的測試，同時整理出 3 種運作環境的儲存效能，以便讓我們能夠了解這 3 種運作環境之間的效能差異。

4 實驗測試結果結論與建議

4.1 RDMA 效能實驗環境測試

本次實驗結果測試三種企業常用應用程式服務(Web,File,Mail Service)

測試環境 (4 Nodes S2D Cluster)，透過 VMFleet 佈建 80 個 VM 虛擬壓測主機，接續使用 Diskspd 執行 IOPS 儲存效能壓力測試。實驗結果請參考下圖 5:

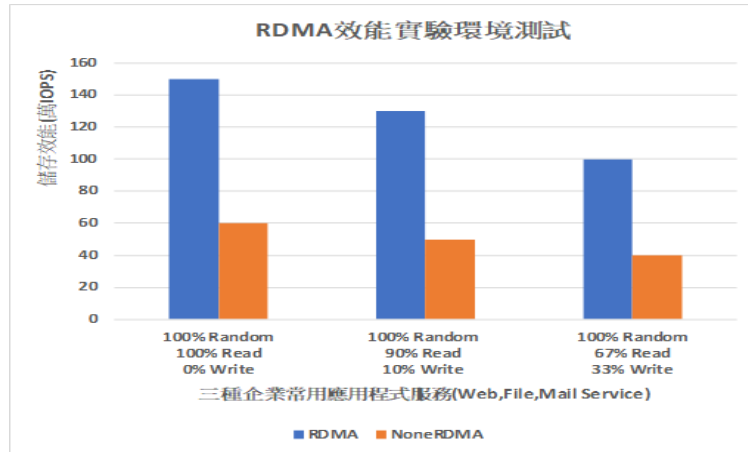


圖 5 RDMA 效能實驗環境測試結果

綜合實際實驗數據測試結果，RDMA 效能實驗環境測試在資料讀取與寫入方面相關，譬如 IOPS 效能表現、CPU 使用率、延遲時間等，S2D RDMA Enabled > S2D RDMA Disabled。

比較「RDMA Enabled」的功能啟用與「RDMA Disabled」的功能關閉，由實際實驗測試結果，可得知節省近 1/3 的 CPU 工作負載、更可達到提升至 2 倍的 IOPS 儲存效能表現以及更低的延遲時間。

由以上結果，我們建議：若企業希望 S2D 能夠有良好的儲存效能表現，則企業建置評估時應該要讓 S2D 維運執行在「RDMA Enabled」環境之中。

4.2 多節點 IOPS 儲存效能表現實驗環境測試

本次實驗結果測試三種企業常用應用程式服務(Web,File,Mail Service)

測試環境 (4 Nodes S2D Cluster)，透過 VMFleet 佈建 80 個 VM 虛擬壓測主機，接續使用 Diskspd 執行 IOPS 儲存效能壓力測試。結果請參考下圖 6：

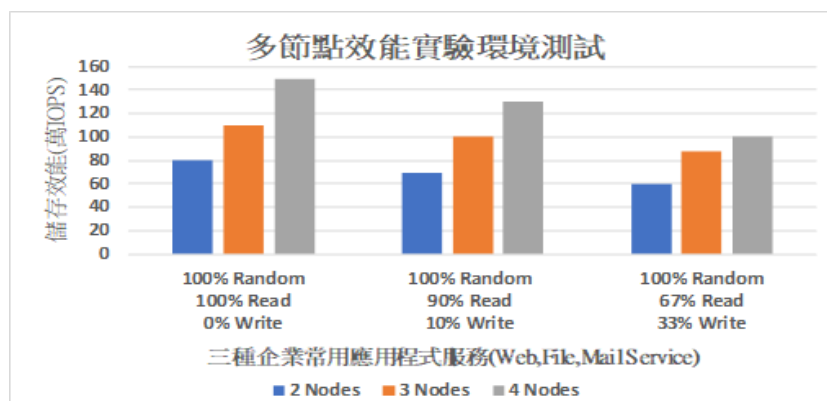


圖 6 多節點效能實驗環境測試結果

綜合實際實驗數據測試結果，多節點效能實驗環境測試在資料讀取與寫入方面相關，譬如 IOPS 效能表現，4 Nodes S2D Cluster > 3 Nodes S2D Cluster > 2 Nodes S2D Cluster。

建置 Microsoft S2D 叢集主機伺服器最少需求數量為 2 Nodes，最多可擴充至 16 Nodes。

實際實驗證明 Microsoft S2D 俱備橫向擴充(Scale-Out)的能力，所以不同的叢集節點數量也會造成不同的 IOPS 儲存效能表現結果。因此，S2D 叢集能因為叢集節點的數量提升，帶來相對應的 IOPS 儲存效能表現提升。

由以上結果，企業在評估需求除了資料高可用性之外，適當的儲存效能表現與相對的儲存空間效率並重與平衡時，建議 Microsoft S2D 叢集主機伺服器佈建數量為 4 Nodes S2D Cluster。

5. 參考文獻及書目

- [1] IDC.Four Pillars, <https://www.idc.com/promo/thirdplatform/fourpillars>
- [2] Gartner. Gartner Says Nexus of Forces Social, Mobile Cloud and Information, <https://www.gartner.com/newsroom/id/2097215>
- [3] Microsoft. Storage Spaces Direct in Windows Server 2016, <https://docs.microsoft.com/en-us/windows-server/storage/storage-spaces/storage-spaces-direct-overview>
- [4] T. Talpey. SMB remote file protocol. Global Education. Storage Networking Industry Association, SNIA, 2015
- [5] R. Recio, B. Metzler, P. Culley, J. Hilland, and D. Garcia. A Remote Direct Memory Access Protocol Speciation. IETF RFC 5040, October 2007
- [6] Infiniband. Released RoCE Specification, http://www.infinibandta.org/content/pages.php?pg=technology_public_specification
- [7] Infiniband. Released RoCEv2 Specification, http://www.infinibandta.org/content/pages.php?pg=technology_public_specification
- [8] Microsoft. Diskspd Utility, <https://gallery.technet.microsoft.com/DiskSpd-a-robust-storage-6cd2f223>