



(19)中華民國智慧財產局

(12)發明說明書公告本

(11)證書號數：TW I734895 B

(45)公告日：中華民國 110 (2021) 年 08 月 01 日

(21)申請案號：107107134

(22)申請日：中華民國 107 (2018) 年 03 月 02 日

(51)Int. Cl. : G06F13/00 (2006.01)

H04L29/06 (2006.01)

H04L12/12 (2006.01)

G06F15/16 (2006.01)

G06F15/173 (2006.01)

(30)優先權：2017/03/31 美國

62/480,113

2017/04/10 美國

62/483,913

2017/06/08 美國

15/618,081

(71)申請人：南韓商三星電子股份有限公司 (南韓) SAMSUNG ELECTRONICS CO., LTD. (KR)
南韓(72)發明人：卡佳瑞 瑞姆達斯 P.KACHARE, RAMDAS P. (IN)；歐拉利格 桑龐 保羅 OLARIG,
SOMPONG PAUL (US)；沃里 佛瑞德 WORLEY, FRED (US)

(74)代理人：林孟閔；盧珮君；陳怡如

(56)參考文獻：

TW 201212033A

TW 201510725A

US 2007/0214316A1

US 2011/0208922A1

審查人員：簡大翔

申請專利範圍項數：20 項 圖式數：12 共 57 頁

(54)名稱

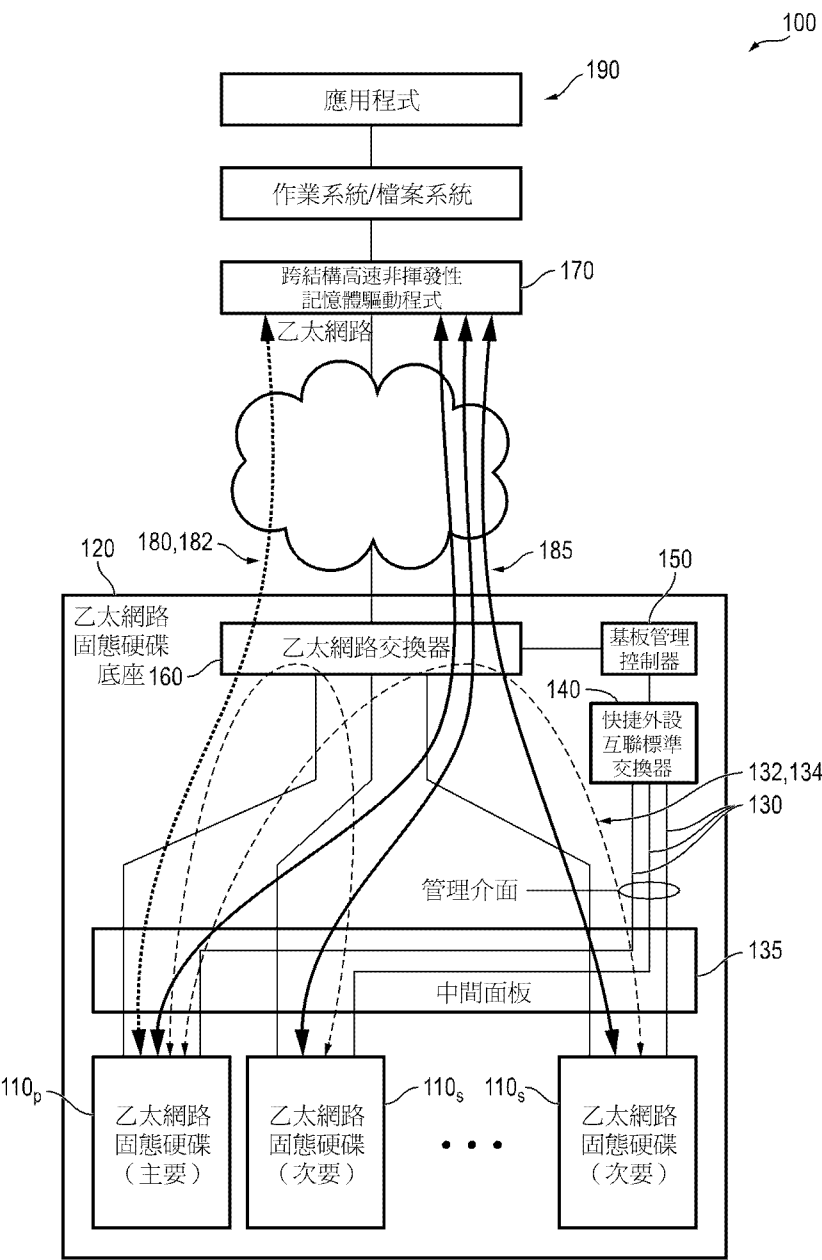
聚集儲存方法、跨結構高速非揮發性記憶體固態硬碟用的容量聚集方法及聚集的乙太網路固態硬碟群組

(57)摘要

一種跨結構高速非揮發性記憶體裝置用的儲存聚集方法，此方法包括：將聚集群組識別為包括多個跨結構高速非揮發性記憶體固態硬碟的聚集的乙太網路固態硬碟；選擇所述聚集群組的所述跨結構高速非揮發性記憶體固態硬碟的其中一個作為主要跨結構高速非揮發性記憶體固態硬碟；選擇所述聚集群組的跨結構高速非揮發性記憶體固態硬碟的其他作為次要跨結構高速非揮發性記憶體固態硬碟；以及以處理器將主要跨結構高速非揮發性記憶體固態硬碟中的對映配置表初始化，以管理跨結構高速非揮發性記憶體固態硬碟。

In a method of storage aggregation for NVMe Over Fabrics devices, the method includes: identifying an aggregation group as an aggregated Ethernet SSD comprising a plurality of NVMe-oF SSDs; selecting one of the NVMe-oF SSDs of the aggregation group as a primary NVMe-oF SSD; selecting others of the NVMe-oF SSDs of the aggregation group as secondary NVMe-oF SSDs; and initializing a Map Allocation Table in the primary NVMe-oF SSD with a processor for managing the NVMe-oF SSDs.

指定代表圖：



【圖1】



I734895

【發明摘要】

【中文發明名稱】聚集儲存方法、跨結構高速非揮發性記憶體固態硬碟用的容量聚集方法及聚集的乙太網路固態硬碟群組

【英文發明名稱】METHOD OF AGGREGATING STORAGE, METHOD OF NVME-OF SSD CAPACITY AGGREGATION AND AGGREGATED ETHERNET SSD GROUP

【中文】一種跨結構高速非揮發性記憶體裝置用的儲存聚集方法，此方法包括：將聚集群組識別為包括多個跨結構高速非揮發性記憶體固態硬碟的聚集的乙太網路固態硬碟；選擇所述聚集群組的所述跨結構高速非揮發性記憶體固態硬碟的其中一個作為主要跨結構高速非揮發性記憶體固態硬碟；選擇所述聚集群組的跨結構高速非揮發性記憶體固態硬碟的其他作為次要跨結構高速非揮發性記憶體固態硬碟；以及以處理器將主要跨結構高速非揮發性記憶體固態硬碟中的對映配置表初始化，以管理跨結構高速非揮發性記憶體固態硬碟。

【英文】In a method of storage aggregation for NVMe Over Fabrics devices, the method includes: identifying an aggregation group as an aggregated Ethernet SSD comprising a plurality of NVMe-oF SSDs; selecting one of the NVMe-oF SSDs of the aggregation group as a primary NVMe-oF SSD; selecting others of the NVMe-oF SSDs of

the aggregation group as secondary NVMe-oF SSDs; and initializing a Map Allocation Table in the primary NVMe-oF SSD with a processor for managing the NVMe-oF SSDs.

【指定代表圖】圖1。

【代表圖之符號簡單說明】

100：系統架構

110s：次要乙太網路固態硬碟

110p：主要乙太網路固態硬碟

120：乙太網路固態硬碟底座

130：私用通訊通道

132：子命令

134：完成項目

135：控制平面

140：快捷外設互聯標準匯流排

150：基板管理控制器

160：乙太網路交換器

170：機跨結構高速非揮發性記憶體驅動程式

180：命令

182：完成項目

185：資料

190：主機

【特徵化學式】

無

【發明說明書】

【中文發明名稱】聚集儲存方法、跨結構高速非揮發性記憶體固態硬碟用的容量聚集方法及聚集的乙太網路固態硬碟群組

【英文發明名稱】METHOD OF AGGREGATING STORAGE, METHOD OF NVME-OF SSD CAPACITY AGGREGATION AND AGGREGATED ETHERNET SSD GROUP

[相關申請案的交叉引用]

【0001】本申請案主張 2017 年 3 月 31 日於美國專利商標局申請的美國暫時專利申請案第 62/480,113 號與 2017 年 4 月 10 日於美國專利商標局申請的美國暫時專利申請案第 62/483,913 號的優先權，兩暫時專利申請案之全部揭露內容均併入本案供參考。

【技術領域】

【0002】本揭露部分實施例大體而言是有關於一種系統與方法，用於將多個記憶體硬碟（例如乙太網路固態硬碟）聚集為使主機視為單一的大型邏輯容量。

【先前技術】

【0003】固態硬碟（Solid State Drive，SSD）歷經迅速發展，目前已成為現代資訊科技基礎結構用的儲存元件首選，從而也取代了傳統硬碟（Hard Disk Drive，HDD）。固態硬碟的延遲極低，且

提供高資料讀取/寫入輸送量以及可靠的資料儲存能力。

【0004】 跨 結 構 高 速 非 揮 發 性 記 憶 體 (Non-Volatile-Memory-express-over-Fabric, NVMe-oF) 是一種新興技術，可使數百或數千個跨結構高速非揮發性記憶體裝置（例如跨結構高速非揮發性記憶體固態硬碟）透過網路結構（例如無限頻寬（IB）、光纖通道（FC）、乙太網路）相互連結。跨結構高速非揮發性記憶體通訊協定可用來執行遠端直連式儲存（Direct Attach Storage, rDAS）。如此一來即可將大量的固態硬碟連接遠端主機。跨結構高速非揮發性記憶體通訊協定使用了遠端直接記憶體存取（Remote Direct Memory Access, RDMA）通訊協定，來提供可靠的高速非揮發性記憶體命令、資料及回應。提供遠端直接記憶體存取服務的傳輸通訊協定包括 iWARP、RoCE v1、RoCE v2。

【0005】 跨結構高速非揮發性記憶體介面可使大量的固態硬碟連接遠端主機。在每個跨結構高速非揮發性記憶體固態硬碟中，驅動程式按例一般是在遠端主機上執行。對部分應用程式而言，單一個固態硬碟提供的儲存容量可能是不足夠的。

【發明內容】

【0006】 本揭露的部分實施例提供一聚集多個固態硬碟的方法，可使主機將該些固態硬碟視為單一大容量邏輯磁碟區，同時提供達成此方法的網路結構。

【0007】 根據部分實施例，在跨結構高速非揮發性記憶體裝置用的儲存聚集方法中，上述方法包括：將聚集群組識別為包括多個跨結構高速非揮發性記憶體固態硬碟的聚集的乙太網路固態硬碟；選擇上述聚集群組中的跨結構高速非揮發性記憶體固態硬碟的其中一個作為主要跨結構高速非揮發性記憶體固態硬碟；選擇上述聚集群組中其他的跨結構高速非揮發性記憶體固態硬碟作為次要跨結構高速非揮發性記憶體固態硬碟；以及以處理器將主要跨結構高速非揮發性記憶體固態硬碟中的對映配置表初始化，以管理跨結構高速非揮發性記憶體固態硬碟。

【0008】 根據部分範例實施例，以上述處理器初始化上述對映配置表是在與上述聚集群組連結的儲存管理員的指導下進行。

【0009】 根據部分範例實施例，對於上述聚集群組的各個跨結構高速非揮發性記憶體固態硬碟，上述對映配置表包括跨結構高速非揮發性記憶體固態硬碟的容量、跨結構高速非揮發性記憶體固態硬碟的位址及跨結構高速非揮發性記憶體固態硬碟的剩餘容量。

【0010】 根據部分範例實施例，上述方法更包括將主要跨結構高速非揮發性記憶體固態硬碟的上述位址提供給使用者應用程式，以使聚集群組與使用者應用程式之間能夠進行資料傳輸。

【0011】 根據部分範例實施例，上述方法更包括：在主要跨結構高速非揮發性記憶體固態硬碟上接收來自與聚集群組連結的主機的管理（Admin）命令；判斷與上述管理命令對應的資料是否僅儲

存在上述主要跨結構高速非揮發性記憶體固態硬碟上、或是儲存在一個或多個次要跨結構高速非揮發性記憶體固態硬碟上；當上述資料儲存在一個或多個次要跨結構高速非揮發性記憶體固態硬碟上時，將上述管理命令分割成分別對應一個或多個次要跨結構高速非揮發性記憶體固態硬碟的一個或多個管理子命令；將上述資料傳輸至主機；接收來自一個或多個次要跨結構高速非揮發性記憶體固態硬碟的子命令完成項目；以及從上述主要跨結構高速非揮發性記憶體固態硬碟建立並傳送完成項目至主機。

【0012】 根據部分範例實施例，上述方法更包括：在一個或多個次要跨結構高速非揮發性記憶體固態硬碟的一個對應的次要跨結構高速非揮發性記憶體固態硬碟上接收一個或多個管理子命令的一個管理子命令；根據上述管理子命令判斷是否要將上述資料從對應的次要跨結構高速非揮發性記憶體固態硬碟傳輸到上述主要跨結構高速非揮發性記憶體固態硬碟；建立完成項目；以及將上述完成項目傳送到上述主要跨結構高速非揮發性記憶體固態硬碟。

【0013】 根據部分範例實施例，上述方法更包括：在上述主要跨結構高速非揮發性記憶體固態硬碟上接收建立命名空間（Namespace）的命令或刪除命名空間的命令；參考上述主要跨結構高速非揮發性記憶體固態硬碟的上述對映配置表；當上述命令是建立命名空間時，在上述主要跨結構高速非揮發性記憶體固態硬碟中及/或在一個或多個次要跨結構高速非揮發性記憶體固態硬

碟中配置容量，或是當上述命令是刪除命名空間時，收回主要跨結構高速非揮發性記憶體固態硬碟及/或一個或多個次要跨結構高速非揮發性記憶體固態硬碟的對應一者；以及更新上述對映配置表。

【0014】 根據部分實施例，上述方法更包括：在上述主要跨結構高速非揮發性記憶體固態硬碟上接收讀取/寫入（Read/Write）命令；查詢上述主要跨結構高速非揮發性記憶體固態硬碟的對映配置表；建立一個或多個讀取/寫入子命令；將上述一個或多個讀取/寫入子命令分別傳送到一個或多個次要跨結構高速非揮發性記憶體固態硬碟；根據上述讀取/寫入命令，在主機與上述主要跨結構高速非揮發性記憶體固態硬碟及/或一個或多個次要跨結構高速非揮發性記憶體固態硬碟之間傳輸資料；以及傳輸上述資料後傳送完成至主機。

【0015】 根據部分實施例，上述方法更包括：在一個或多個次要跨結構高速非揮發性記憶體固態硬碟的一個對應的次要跨結構高速非揮發性記憶體固態硬碟上接收上述一個或多個讀取/寫入子命令中的一個讀取/寫入子命令；擷取對應讀取/寫入子命令的輸送資訊；從上述對應的次要跨結構高速非揮發性記憶體固態硬碟發出讀取/寫入要求至主機；以及在對應上述讀取/寫入子命令的資料傳輸完成之後，將完成項目傳送到主要跨結構高速非揮發性記憶體固態硬碟。

【0016】 根據部分範例實施例，一種跨結構高速非揮發性記憶體

乙太網路固態硬碟群組中的跨結構高速非揮發性記憶體固態硬碟用的容量聚集方法，上述方法包括：識別聚集群組的多個跨結構高速非揮發性記憶體固態硬碟；將跨結構高速非揮發性記憶體固態硬碟的其中一個指派為主要跨結構高速非揮發性記憶體固態硬碟；以及將剩餘的跨結構高速非揮發性記憶體固態硬碟指派為次要跨結構高速非揮發性記憶體固態硬碟，其中主機的主機驅動程式唯一可見的跨結構高速非揮發性記憶體固態硬碟為主要跨結構高速非揮發性記憶體固態硬碟。

【0017】 根據部分實施例，上述方法更包括根據主要跨結構高速非揮發性記憶體固態硬碟從主機接收的命令，維護主要跨結構高速非揮發性記憶體固態硬碟的對映配置表，其中上述對映配置表指示在聚集群組的主要跨結構高速非揮發性記憶體固態硬碟與一個或多個次要跨結構高速非揮發性記憶體固態硬碟分割的邏輯區塊位址（logical block address，LBA）空間。

【0018】 根據部分實施例，上述方法更包括以處理器將對映配置表初始化，並依據指派為主要跨結構高速非揮發性記憶體固態硬碟的跨結構高速非揮發性記憶體固態硬碟的其中一個來設定聚集群組。

【0019】 根據部分實施例，上述方法更包括依據主要跨結構高速非揮發性記憶體固態硬碟從主機接收到的命令來聚集次要跨結構高速非揮發性記憶體固態硬碟與上述主要跨結構高速非揮發性記憶體固態硬碟的容量，使得上述聚集群組的所述多個跨結構高速

非揮發性記憶體固態硬碟對主機而言顯現為單一的聚集邏輯容量。

【0020】 根據部分實施例，上述方法更包括：使用上述主要跨結構高速非揮發性記憶體固態硬碟來將容量配置給一個或多個次要跨結構高速非揮發性記憶體固態硬碟；以及使用上述主要跨結構高速非揮發性記憶體固態硬碟來記錄已配置的容量與對映配置表中相關聯且對映的邏輯區塊位址範圍。

【0021】 根據部分實施例，上述方法更包括利用上述主要跨結構高速非揮發性記憶體固態硬碟過度配置次要跨結構高速非揮發性記憶體固態硬碟與上述主要跨結構高速非揮發性記憶體固態硬碟的聚集容量。

【0022】 根據部分實施例，上述方法更包括：在上述主要跨結構高速非揮發性記憶體固態硬碟上接收來自上述主機的命令；將上述命令分割成多個子命令且每一子命令對應至對應的次要跨結構高速非揮發性記憶體固態硬碟的各別一者；以及將子命令從上述主要跨結構高速非揮發性記憶體固態硬碟傳送到對應的次要跨結構高速非揮發性記憶體固態硬碟。

【0023】 根據部分實施例，上述方法更包括基於個別的子命令，從對應的次要跨結構高速非揮發性記憶體固態硬碟直接傳送資料到主機。

【0024】 根據部分實施例，上述方法更包括：在對應的次要跨結構高速非揮發性記憶體固態硬碟上接收來自上述主要跨結構高速

非揮發性記憶體固態硬碟的個別子命令；執行對應個別子命令的任務；以及任務完成後，從對應的次要跨結構高速非揮發性記憶體固態硬碟傳送個別子命令完成項目到上述主要跨結構高速非揮發性記憶體固態硬碟。

【0025】 根據部分實施例，上述方法更包括：使用上述主要跨結構高速非揮發性記憶體固態硬碟來維護子命令內容表；在上述主要跨結構高速非揮發性記憶體固態硬碟接收來自次要跨結構高速非揮發性記憶體固態硬碟的子命令完成項目；以及根據接收到的子命令完成項目，以上述主要跨結構高速非揮發性記憶體固態硬碟來追蹤子命令的執行。

【0026】 根據部分範例實施例，提供一種聚集的乙太網路固態硬碟群組，包括：乙太網路固態硬碟底座；乙太網路固態硬碟底座上的乙太網路交換器，以與主機驅動程式通訊；處理器，耦接至乙太網路交換器；快捷外設互聯標準（Peripheral Component Interconnect Express，PCIe）交換器，耦接至基板管理控制器；多個跨結構高速非揮發性記憶體固態硬碟，包括：主要跨結構高速非揮發性記憶體固態硬碟；以及多個次要跨結構高速非揮發性記憶體固態硬碟，透過包括乙太網路交換器與快捷外設互聯標準交換器的私用通訊通道，來與主要跨結構高速非揮發性記憶體固態硬碟連結，其中只有上述主要跨結構高速非揮發性記憶體固態硬碟對主機而言是可見的，且其中上述基板管理控制器用於在開始決定跨結構高速非揮發性記憶體固態硬碟中何者包括上述主要跨

結構高速非揮發性記憶體固態硬碟。

【0027】 基於上述，因為只有單一主要乙太網路固態硬碟進行所有跨結構高速非揮發性記憶體通訊協定的處理作業，同時追蹤所有透過次要乙太網路固態硬碟的相關子命令是否已完成，並確保次要乙太網路固態硬碟對主機而言是不可見的，因此乙太網路固態硬碟的聚集群組對主機是顯示為單一的大型邏輯容量。

【圖式簡單說明】

【0028】 為讓本發明能更明顯易懂，下文特舉實施例，並配合所附圖式作詳細說明如下，在所附圖式中：

【0029】

圖 1 為繪示根據本揭露一實施例的用於包括單一乙太網路固態硬碟底座內的數個聚集的乙太網路固態硬碟的跨結構高速非揮發性記憶體乙太網路固態硬碟（Ethernet SSD，eSSD）儲存的系統架構 100 方塊圖。

圖 2 為繪示根據本揭露一實施例的圖 1 所示的數個乙太網路固態硬碟底座在乙太網路固態硬碟機架上彼此相連的方塊圖。

圖 3 為根據本揭露一實施例的主要乙太網路固態硬碟所維護的對映配置表（Map Allocation Table）的範例。

圖 4 為繪示根據本揭露的一實施例的對映配置表初始化流程圖。

圖 5 為繪示根據本揭露一實施例包括資料往返於分別位於多個機架中的多個乙太網路固態硬碟底座中多個聚集的乙太網路固

態硬碟的用於跨結構高速非揮發性記憶體乙太路固態硬碟儲存的系統架構方塊圖。

圖 6 繪示一表格的範例命令內容。

圖 7 為繪示根據本揭露一實施例的主要乙太網路固態硬碟處理管理（Admin）命令的流程圖。

圖 8 為繪示根據本揭露一實施例的執行命名空間建立與刪除（Namespace Create and Delete）命令的流程圖。

圖 9 為繪示根據本揭露一實施例所在自主要乙太網路固態硬碟控制下執行讀取/寫入（Read/Write）命令的流程圖。

圖 10 為繪示根據本揭露一實施例的次要乙太網路固態硬碟執行管理子命令的流程圖。

圖 11 為繪示根據本揭露一實施例的次要乙太網路固態硬碟執行讀取/寫入子命令的流程圖。

圖 12 為繪示根據本揭露一實施例的在次要乙太網路固態硬碟中的資料傳送與子命令完成同步的流程圖。

【實施方式】

【0030】 通過參考下方對實施例與所附圖式的詳細描述，可更容易地理解本發明的發明概念特徵與其實現方法。本發明之實施例將於下文搭配所附圖式進行說明，且在全文中相同的標號代表相同的構件。然而，本發明可被例示為諸多不同形式且不應被視為僅限於本文所述的實施例。確切而言，提供該些實施例是作為範例，為了使此揭露內容更為透徹及完整，並向本領域具通常知識

者充分傳達本揭露的觀點與特徵。據此，對本領域具通常知識者而言，在完整理解本發明的觀點與特徵方面非必要的程序、元件、技術可能不會加以說明。除非另有說明，在所附圖式與所述說明中，利用相同符號指代相似的元件，因此不再重複說明。在圖式中，為清晰起見，可放大元件、層、區域的相關尺寸。

【0031】 在接下來的描述中，出於解釋上的目的，列舉出數個具體細節以提供對於各項實施例的通盤理解。然而顯而易見的是，可以在沒有這些具體細節或提供一個或多個相等配置的情況下實作各項實施例。在其他情況下，以方塊圖形式顯示已知的結構與裝置，以避免不必要地混淆各項實施例。

【0032】 應可理解的是，儘管本文中可能使用「第一」、「第二」、「第三」等用語來闡述各種元件、構件、區域、層及/或區段，然而該些元件、構件、區域、層及/或區段不應受限於該些用語。該些用語僅用於區分各個元件、構件、區域、層或區段。因此，在不脫離本發明的精神與範圍內，以下所論述的第一元件、構件、區域、層或區段可被稱為第二元件、構件、區域、層或區段。

【0033】 在本文中，為便於解釋，可使用例如「向...下面(beneath)」、「在...之下(below)」、「下方的(lower)」、「在...下方(under)」、「在...之上(above)」、「上方的(upper)」等空間相對性用語來闡述圖式中所示一個元件或特徵相對於另一(其他)元件或特徵的關係。應理解，該些空間相對性用語旨在除了圖式中所示定向以外，亦囊括裝置在使用或操作中的不同定向。舉例而言，若翻轉

圖式中的裝置，則描述為在其他元件「之下」或「下方」或「下方的」元件，此時將被定向為在其他元件或特徵「之上」。因此，用語「在...之下」與「在...下方」可包含上方及下方兩種定向。所述裝置亦可具有其他定向（例如旋轉 90 度或其他定向），且本文中使用的空間相對性描述語應可相應地進行解釋。

【0034】 應理解，當稱一元件、層、區域或構件位於另一元件、層、區域或構件「上」、「連接至」或「耦合於」另一元件、層、區域或構件時，所述元件、層、區域或構件可直接位於所述另一元件、層、區域或構件「上」、直接「連接至」或直接「耦合於」所述另一元件、層、區域或構件或其間可存在其他一個或多個元件、層、區域或構件。此外，同時應理解，當稱一元件或層位於兩個元件或層的「之間」時，所述元件或層是唯一位於兩元件或層之間的元件或層，或可存在其他一個或多個元件或層。

【0035】 為了本揭露的目的，「X、Y、Z 中至少其中之一」與「從 X、Y、Z 組成的群組中選擇至少一個」應理解為僅限於 X、僅限於 Y、僅限於 Z，或 X、Y、Z 其中兩者或更多的任意組合，例如 XYZ、XYY、YZ、ZZ。在通篇中相同的編號指代相同的元件。本文中所使用的用語「及/或」包含相關列出項其中一個或多個項的任意及所有組合。

【0036】 在以下範例中，x 軸，y 軸與 z 軸不限於直角坐標系的三個軸，可以從更廣泛的意義上解釋。舉例而言，x 軸，y 軸和 z 軸可以彼此垂直，或者也可代表彼此不互垂直的不同方向。

【0037】 本文所用術語僅用於闡述特定實施例，並非用以限定本發明。除非上下文中清楚地另外指明，否則本文所用的單數形式「一（a、an）」旨在亦包括複數形式。更應理解，當在本說明書中使用用語「包括（comprises、comprising、includes、including）」時，是用於指明所述特徵、整數、步驟、操作、元件及/或構件但不排除一個或多個其他特徵、整數、步驟、操作、元件及/或構件的存在或添加。本文中所使用的用語「及/或」包含相關列出項其中一個或多個項的任意及所有組合。當在元件列表之前出現「至少其中之一」之時，該表達方式旨在修飾整個元件列表說明，並且非用於修飾列表中的單一個元件。

【0038】 如本文所用術語「大致上」、「大約」與其他類似術語被用作近似術語而不是用來表達程度的術語，並且這些術語旨在說明本領域具通常知識者所悉知的測量值或計算值中的固有偏差。更進一步而言，當描述本發明的實施例時使用「可」是指「本發明的一個或多個實施例」。如本文所用術語，「使用（use）」、「使用中（using）」與「使用過（used）」可分別被視為與術語「利用（utilize）」、「利用中（utilizing）」與「利用過（utilized）」同義。而術語「示例性的（exemplary）」旨在指代範例或說明。

【0039】 當某個實施例可以不同方式實施時，可執行與所描述的順序不同的特定程序順序。舉例而言，兩個前後接連說明的程序大致上可在同一時間執行，或者可也以與所述順序相反的順序來執行。

【0040】 同時，任何於此揭露及/或引述的數值範圍指在涵括引述範圍內所包含的相同精確數值的子範圍。舉例而言，「1.0 到 10.0」的範圍指在涵括（且包括）所有所引述的最小數值 1.0 與所引述的最大數值 10.0 之間的所有子範圍。也就是說，大於等於最小數值 1.0 與小於等於最大數值 10.0 的數值之間任何的範圍，例如 2.4 到 7.6。任何於此引述的最大數值限制指在包括所有涵蓋的較低數值限制，且任何文中所引述的最小數值限制指在包括所涵蓋的所有較高數值限制。據此，本申請者保留本申請案修改的權利，包括本申請專利範圍，以可以明確的引述涵蓋在本文明確引述的範圍內的子範圍。

【0041】 於此說明各項實施例，搭配參考用來示意性的說明實施例及/或中間結構的剖面示意圖。因此，可預期將出現由例如製造技術及/或公差而導致的圖示形狀變化的產生。因此，揭露的實施例不應被視為僅限於本文所示的特定形狀，而是例如包括由製造而引起的形狀改變。例如，描述為矩形的植入區域通常在其邊緣具有圓形或彎曲的特徵及/或植入濃度梯度，而不是僅是從植入區域到非植入區域的二元變化。同樣的，通過佈植形成的掩埋區可造成在掩埋區與植入產生的表面之間的區域中產生佈值。因此，圖式中示出的區域本質上是示意性的，並且其形狀並非用於顯示出裝置區域的實際形狀，而且不用於限制內容。

【0042】 根據本發明實施例於此所述的電子裝置及/或任何其他相關裝置或構件，可利用任何適合的硬體、軟體（例如特殊應用積

體電路)、軟體或硬體、韌體及軟體的組合來實作。舉例而言，這些裝置的各項構件可在一個積體電路(Integrated Circuit, IC)晶片或在不同積體電路晶片上形成。更進一步而言，這些裝置的各項構件，可在柔性印刷電路薄膜、捲帶式晶片載體封裝(Tape Carrier Package, TCP)、印刷電路板(Printed Circuit Board, PCB)上執行，或形成於基板上。此外，這些裝置的各項組件可以是程序或執行緒，在一個或多個處理器、一個或多個運算裝置上執行，並執行電腦程式指令，還可以與其他系統構件整合，以執行本文所述的各項功能。電腦程式指令存儲在記憶體中，可以使用標準記憶體裝置(例如隨機存取記憶體(RAM))在運算裝置中執行。電腦程式指令也可以儲存在其他非暫時性電腦可讀媒體中，例如光碟機、快閃磁碟機等等。在不脫離本發明例示性實施例的精神和範圍下，此項技術領域中具有通常知識者應可知不同運算設備的功能可以相結合或整合成單一的運算裝置，或者特定運算裝置的功能可分配到一個或多個其他運算裝置上。

【0043】 除非另行定義，於此使用的所有術語(包括技術和科學術語)具有與本發明所屬領域的具有通常知識者通常理解的相同的含義。應進一步理解的是，諸如在通用字典中定義的術語，應該被解釋為具有與其在相關技術及/或本說明書的上下文中的含義一致的含義，並且不應被解釋為理想化或過度形式化的意義，除非另行明確定義。

【0044】 圖 1 為繪示根據本揭露一實施例的用於包括單一乙太網

路固態硬碟底座 120 內的數個聚集的乙太網路固態硬碟 110 的跨結構高速非揮發性記憶體乙太網路固態硬碟(Ethernet SSD, eSSD)儲存的系統架構方塊圖。圖 2 為繪示根據本揭露一實施例的圖 1 所示的數個乙太路固態硬碟底座 120 在一乙太網路固態硬碟機架彼此相連的方法塊。

【0045】 如上所述，跨結構高速非揮發性記憶體介面可使大量的固態硬碟 111 連接遠端主機 190。在每個跨結構高速非揮發性記憶體固態硬碟 110 中，驅動程式例如是在遠端主機 190 上執行。然而，對部分應用程式而言，單一個固態硬碟 110 提供的儲存容量是不足夠的。這類應用程式較為適用於具有數百兆位元組(Terabyte, TB)的單一邏輯磁碟區。據此，這類應用程式可能較為適用於本揭露所提供的實施例。因為本揭露實施例提供大量的單獨固態硬碟 110，這些固態硬碟 110 共同聚集成「聚集群組(Aggregation Group)」，且對應用程式而言顯示為單一邏輯磁碟區。

【0046】 舉例而言，24 個 16 TB 的固態硬碟可顯示為單一邏輯 384 TB 硬碟。需要大量聚集的固態硬碟 110 應用程式的一些範例包括大型資料採礦與分析、石油化學、氣體與能源探勘，實驗粒子物理學，以及製藥藥物開發。這些範例需要高效能運算(High Performance Computing, HPC)，高效能運算同時需要大量儲存容量以及高效能。

【0047】 雖然可以具有系統軟體層聚集底下的固態硬碟 110，並且

提供單一邏輯、擴充性磁碟區，但是這樣的系統軟體通常非常複雜且精密。這樣的軟體可能需要在例如主機 190 上執行的大量跨結構高速非揮發性記憶體驅動程式，因此會消耗系統資源，例如記憶體、中央處理器（Central Processing Unit，CPU）週期，以及電力。目標端解決方案可能使用 x86 伺服器或獨立磁碟冗餘陣列晶片（Redundant Array of Independent Disk-on-Chip，ROC）系統，來提供大容量的單一邏輯磁碟區。然而，這類的解決方案通常很複雜、昂貴，且對效能與能源使用上有負面影響。舉例而言，根據本發明實施例，使用中央處理器來接收與傳送資料所消耗的能源，是使用直接記憶體存取（Direct Memory Access，DMA）引擎、特殊應用積體電路（Application-Specific Integrated Circuit，ASIC）等等所消耗能源的好幾倍。

【0048】 據此，本發明實施例提供可在乙太網路跨結構高速非揮發性記憶體固態硬碟中，以有效率且符合成本效益的方式聚集多個乙太網路固態硬碟 110 的方法與架構。

【0049】 請參照圖 1，乙太網路固態硬碟 110 被指派了二個角色當中的其中一個角色（例如，由儲存管理員指派），因此每一個乙太網路固態硬碟 110 可作為主要乙太網路固態硬碟（P-eSSD）110p 或是次要乙太網路固態硬碟（S-eSSD）110s。在單一個底座 120（或在給定機架中的多個底座 120 中，或在分佈在廣大範圍內的多個機架 230 中的多個底座 120）中的單一個主要乙太網路固態硬碟 110p 與數個次要乙太網路固態硬碟 110s 的集合共同地提供遠端

主機 190 所需的快閃記憶體容量作為單一個邏輯硬碟。乙太網路固態硬碟底座 120 包括乙太網路固態硬碟 110 以及處理器（例如基板管理控制器（BMC）裝置 150）、與用於連接外部裝置的乙太網路交換器 160。儘管以下實施例中所描述乙太網路固態硬碟底座 120 是用來指代一個跨結構高速非揮發性記憶體裝置群組，但本發明的其他實施例也可同樣應用於其他任何的多個跨結構高速非揮發性記憶體裝置，不論該些裝置的實體外殼（例如底座、機架或箱式外殼）為何。更進一步而言，雖然乙太網路固態硬碟 110 是用來描述以下實施例的跨結構高速非揮發性記憶體裝置，其他的跨結構高速非揮發性記憶體裝置也可同樣適用於本發明的其實施例。

【0050】 據此，透過對應的乙太網路交換器 160，即可在一個機架 230 中橫跨多個底座 120，以及跨越包含了多個底座 120 的多個機架 230 來聚集乙太網路固態硬 110。

【0051】 主要乙太網路固態硬碟 110p 是遠端主機跨結構高速非揮發性記憶體驅動程式 170 唯一可見的乙太網路固態硬 110，因此終端跨結構高速非揮發性記憶體通協定。主要乙太網路固態硬碟 110p 在相同的聚集群組中代表其自身以及所有剩餘的次要乙太網路固態硬碟 110s，對遠端主機 190 顯示為單一的大型聚集邏輯容量。主要乙太網路固態硬碟 110p 接收來自遠端主機跨結構高速非揮發性記憶體驅動程式 170 的所有輸入/輸出（input/output，I/O）命令 180，並提供命令回應（例如完成項目）182 給遠端主機 190。

【0052】 主要乙太網路固態硬碟 110p 也維護對映配置表（Map Allocation Table，MAT），對映配置表指示了在相同的乙太網路固態硬碟 110 聚集群組中，在主要乙太網路固態硬碟 110p 以及部分或全部的次要乙太網路固態硬碟 110s 之間分割的邏輯區塊位址（logical block address，LBA）空間。當主要乙太網路固態硬碟 110p 接收到輸入/輸出命令 180 時，主要乙太網路固態硬碟 110p 先查詢對映配置表（例如下文所述的圖 3 的對映配置表 300），來決定那一個乙太網路固態硬碟 110（主要乙太網路固態硬碟 110p、一個或多個次要乙太網路固態硬碟 110s 或兩者的組合）能夠滿足輸入/輸出命令 180。根據對映配置表，主要乙太網路固態硬碟 110p 接下來傳送經適當應修改的跨結構高速非揮發性記憶體輸入/輸出子命令 132 給合適的次要乙太網路固態硬碟 110s 的集合。

【0053】 為了要傳送子命令 132，在接上電源後，主要乙太網路固態硬碟 110p 還會在快捷外設互聯標準（Peripheral Component Interconnect Express，PCIe）匯流排 140 及帶有各個次要乙太網路固態硬碟 110s 的控制平面 135 上建立私用乙太網路-遠端直接記憶體存取（Ethernet-RDMA）連線（或是建立適當的專屬通訊通道）130。私用佇列對（Queue-Pair，QP）通訊通道 130 被主要乙太網路固態硬碟 110p 用來將輸入/輸出命令（例如子命令 132）傳送給次要乙太網路固態硬碟 110s，以及用來從次要乙太網路固態硬碟 110s 接收完成項目 134。私用通訊通道 130 可以是乙太網路，而且能夠讓資料透過乙太網路交換器 160 傳輸。然而，私用通訊通

道 130 也可以是基於快捷外設互聯標準建立的通道，因而能讓資料透過快捷外設互聯標準交換器傳輸。亦即，所有的乙太網路固態硬碟 110 都有可能使用二個或二個以上的通訊模式來相互通訊。舉例而言，乙太網路通道通常可用來進行資料傳輸，快捷外設互聯標準通道可用來進行管理，而其中任一個通道皆可作為私用通訊通道 130 使用。

【0054】 次要乙太網路固態硬碟 110s 即為一般的跨結構高速非揮發性記憶體固態硬碟 110，僅用來透過跨結構高速非揮發性記憶體通訊協定，將資料傳送至遠端主機 190 或自遠端主機 190 接收資料。與遠端主機 190 的資料傳輸是透過使用遠端直接記憶體存取讀取服務與遠端直接記憶體存取寫入服務來直接完成。次要乙太網路固態硬碟 110s 接收來自主要乙太網路固態硬碟 110p 的命令（例如子命令 132），但是不直接從遠端主機跨結構高速非揮發性記憶體驅動程式 170 接收命令。次要乙太網路固態硬碟 110s 傳送子命令完成項目 134 給主要乙太網路固態硬碟 110p 來指示子命令 132 已完成，而是不傳送給遠端主機 190。

【0055】 主要乙太網路固態硬碟 110p 處理所有的跨結構高速非揮發性記憶體通訊協定終端，處理所有主機命令與完成佇列（例如提交佇列/完成佇列（Submission Queue/Completion Queue，SQ/CQ）），而且對在遠端主機啟動器執行的遠端主機跨結構高速非揮發性記憶體驅動程式 170 是可見的。當遠端主機驅動程式 170 發送跨結構高速非揮發性記憶體管理（Admin）命令 180 或是輸入

/輸出命令 180，命令 180 是發送給主要乙太網路固態硬碟 110p，而所有的管理命令 180 是由主要乙太網路固態硬碟 110p 所執行。然而輸入/輸出命令 180 可以在多個乙太網路固態硬碟 110 之間傳送。

【0056】 主要乙太網路固態硬碟 110p 也可以根據輸入/輸出命令 180 來執行本身的資料傳輸共享。主要乙太網路固態硬碟 110p 接下來等待所有子命令完成項目 134 都抵達（例如出自於次要乙太網路固態硬碟 110s 的集合）私用通訊通道 130 後，才會傳送對應原始命令 180 的命令完成項目 182 給遠端主機 190。

【0057】 主要乙太網路固態硬碟 110p 也維護命令內容表（例如請參見圖 6）中執行的各個命令的「命令內容」。主要乙太網路固態硬碟 110p 使用此命令內容來追蹤子命令 132 的執行狀態、資料傳輸狀態，以及任意錯誤狀態。當所有子命令 132 都完成後，命令回應/完成項目 182 將會提供給遠端主機 190，而命令內容表將會解除配置。

【0058】 請參照圖 2，多個乙太網路固態硬碟底座 120 可以在乙太網路固態硬碟機架 230 中連結在一起，其中機架頂端（Top-Of-Rack，TOR）交換器 240 是用來使共同機架 230 中多個底座 120 之間的連結性。同樣的，分別位於不同地理位置的多個機架 230 也能透過個別的機架頂端交換器 240 來相互連結，其是直接相互連結，或是透過外部交換器而相互連結。這些乙太網路機架 230 可以是位於同一棟資料中心建築物內，或也可以是分佈在不同的地理區

域中。

【0059】 如上所述，本揭露實施例提供的機制，能夠聚集多個乙太網路跨結構高速非揮發性記憶體固態硬碟（乙太網路固態硬碟）110以顯示為單一大容量的跨結構高速非揮發性記憶體固態硬碟。乙太網路固態硬碟 110 可以位於單一底座 120 內，位於單一機架 230 的多個底座 120 內，或甚至也可以散佈在多個乙太網路機架 230 上，且各個乙太網路機架 230 皆具有多個底座 120。乙太網路固態硬碟 110 的其中一個被指派為主要乙太網路固態硬碟（P-eSSD）110p 的角色。其他的乙太網路固態硬碟 110 被指派為次要乙太網路固態硬碟（S-eSSD）110s 的角色。儘管次要乙太網路固態硬碟 110s 與遠端主機啟動器執行直接傳輸資料，不過次要乙太網路固態硬碟 110s 接收來自主要乙太網路固態硬碟 110p 的子命令 132、完成子命令 132、並且回傳子命令 132 的完成項目 134 給主要乙太網路固態硬碟 110p。據此，在本實施例中，在不犧牲任何儲存頻寬的情況下，容量得以聚集而作動為有效的單一乙太網路固態硬碟。

【0060】 圖 3 為根據本揭露一實施例一主要乙太網路固態硬碟 110p 所維護的「對映配置表」300 範例。圖 4 為繪示根據本揭露的一實施例的對映配置表 300 初始化流程圖 400。

【0061】 請參照圖 3 與圖 4，如上所述，本實施例利用了兩種乙太網路固態硬碟 110 類型（例如主要乙太網路固態硬碟 110p 與次要乙太網路固態硬碟 110s）。主要乙太網路固態硬碟 110p 與次要乙

太網路固態硬碟 110s 兩者皆使用跨結構高速非揮發性記憶體通訊協定，來提供儲存服務給主機 190。主要乙太網路固態硬碟 110p 維護表格（例如對映配置表（MAT）300），此表格包含乙太網路固態硬碟 110 的聚集群組中的次要乙太網路固態硬碟 110s 的細節，而乙太網路固態硬碟 110 的聚集群組對主機 190 而言為單一邏輯磁碟區。

【0062】對映配置表 300 可經由基板管理控制器 150 在相同底座 120 中初始化為主要乙太網路固態硬碟 110p。基板管理控制器 150 可管理乙太網路底座 120 與構件，像是乙太網路交換器 160 與乙太網路固態硬碟 110。基板管理控制器 150 具有用於系統管理目的的快捷外設互聯標準介面與系統管理匯流排（System Management Bus，SMBus）介面。同時，基板管理控制器 150 決定要聚集那些乙太網路固態硬碟 110（S410）（例如在儲存管理員的指示下）。當要聚集的乙太網路固態硬碟 110 決定好時，基板管理控制器 150 即可設定乙太網路交換器 160。

【0063】基板管理控制器 150 在儲存管理員的指引下初始化對映配置表 300 的左側三欄。基板管理控制器 150 以及儲存管理員看得見也知道存在聚集群組/儲存系統中的所有乙太網路固態硬碟 110。所知內容包括每一個乙太網路固態硬碟 110 的容量 311 與位址位置 312。儲存管理員可決定那些次要乙太網路固態硬碟 110s 是形成「聚集的乙太網路固態硬碟」（例如聚集群組）所必需的。基板管理控制器 150 與儲存管理員會通告或提供主要乙太網路固

態硬碟 110p 的網路位址給使用者，如此一來，對應於遠端主機跨結構高速非揮發性記憶體驅動程式 170 的使用者應用程式，將知道在那裡可以找到聚集的乙太網路固態硬碟。基板管理控制器 150 與儲存管理員也可以選擇或指定乙太網路固態硬碟 110 的其中一個作為主要乙太網路固態硬碟 110p (S420)，也可以在做出原始指派動作後，依據一個或一個以上的各種不同理由，變更作為主要乙太網路固態硬碟 110p 的乙太網路固態硬碟 110。接下來，基板管理控制器 150 可程式化聚集群組的主要與次要模式 (S430)。

【0064】 主要乙太網路固態硬碟 110p 可在基板管理控制器 150 上保留一份對映配置表 300 的複本，而且可以定期更新儲存在基板管理控制器 150 上的對映配置表 300。在部分實施例中，只有主要乙太網路固態硬碟 110p 包含正式的對映配置表 300，而「0」的乙太網路固態硬碟索引 313 代表主要乙太網路固態硬碟 110p，而剩餘的乙太網路固態硬碟索引值則相對於各別的次要乙太網路固態硬碟 110s。

【0065】 主要乙太網路固態硬碟 110p 終端主機驅動程式 170 的跨結構高速非揮發性記憶體通訊協定，並執行主機驅動程式 170 發出的所有命令 180。當主機 190 的命令 180 完成後，主要乙太網路固態硬碟 110p 利用「完成項目」(Completion Entry) 182 形式將完成項目 182 回傳給主機驅動程式 170。關於主機命令 180，遠端主機跨結構高速非揮發性記憶體驅動程式 170 則完全不知道次要乙太網路固態硬碟 110s 的存在。主要乙太網路固態硬碟 110p 也維

護提交佇列（SQ），並且提交命令完成給完成佇列（CQ）。

【0066】 主要乙太網路固態硬碟 110p 更新並維護對映配置表 300 的右側三欄。當遠端主機跨結構高速非揮發性記憶體驅動程式 170 建立「命名空間」（Namespace）時，某些快閃容量會配置給該命名空間。命名空間邏輯區塊位址範圍 314 對映到乙太網路固態硬碟 110 的集合，並且紀錄在由主要乙太網路固態硬碟 110p 所維護的對映配置表 300 中。以下將參考圖 8，對此程序的詳細內容進行如下闡述。

【0067】 主要乙太網路固態硬碟 110p 也可以執行某些初始化程序。一旦對映配置表 300 在主要乙太網路固態硬碟 110p 中初始化後（S440），主要乙太網路固態硬碟 110p 即得知那些乙太網路固態硬碟 110 是為對應的次要乙太網路固態硬碟 110s。主要乙太網路固態硬碟 110p 接下來與聚集群組中的每一個次要乙太網路固態硬碟 110s 設定通訊通道 130。通訊通道 130 能通過底座 120 內的乙太網路交換器 160 而經過乙太網路介面，或者通過底座 120 內的快捷外設互聯標準交換器而經過快捷外設互聯標準介面。如果次要乙太網路固態硬碟 110s 的其中一個是位於相同機架 230 中的不同底座 120 內，則通訊通道 130 是透過機架頂端交換器 240 來建立的。通訊通道 130 在給定的底座 120 內也可以是經過快捷外設互聯標準匯流排 140。

【0068】 圖 5 為繪示根據本揭露一實施例包括資料往返於分別位於多個機架 230 中的多個乙太網路固態硬碟底座 120 中多個聚集

的乙太網路固態硬碟 110 的用於跨結構高速非揮發性記憶體乙太網路固態硬碟（eSSD）儲存的系統架構 500 方塊圖。

【0069】 請參照圖 5，主要乙太網路固態硬碟 110p 可以透過外部網路交換器與路由器，與位於廣域網路(Wide Area Network, WAN) 的次要乙太網路固態硬碟 110s 來建立乙太網路通訊通道 530。此私用的乙太網路通訊通道 530 可以是遠端直接記憶體存取佇列對（QP）或可以是專屬方法。乙太網路通訊通道 530 是用來交換子命令 132 與相關聯的完成。

【0070】 圖 6 繪示一表格 600 的範例命令內容。

【0071】 請參照圖 6，每一個字命令 132 具有命令識別(ID) 640 並攜帶「命令標記」(Command Tag) 610，因此當主要乙太網路固態硬碟 110p 接收到完成項目 134 時，可以透過這些完成項目 134 來追溯原始命令 180。在追溯子命令 132 的完成項目 134 時，「子命令#」(# of Sub-commands) 欄位遞減，而接收到的錯誤狀態與當前狀態鎖存。當「子命令#」欄位為零時，對應的命令 180（亦即子命令 132 的上層命令）完成，而主要乙太網路固態硬碟 110p 可將完成項目 182 傳送回遠端主機 190。這時主要乙太網路固態硬碟 110p 利用累積的錯誤狀態 630 來產生一完成項目，並將此完成項目放入相關聯的完成佇列中。

【0072】 圖 7 為繪示根據本揭露一實施例的主要乙太網路固態硬碟 110p 處理管理（Admin）命令 180 的流程圖 700。

【0073】 請參照圖 1 與圖 7，如上所述，各個主要乙太網路固態硬

碟 110p 維護命令提交佇列。當主要乙太網路固態硬碟 110p 接收到可執行的命令 180 (S710) 時，主要乙太網路固態硬碟 110p 仲裁提交佇列，然後選擇要執行的命令 180。主要乙太網路固態硬碟 110p 執行所有高速非揮發性記憶體命令 (例如管理命令與輸入/輸出命令) 180。也就是說，雖然次要乙太網路固態硬碟 110s 可直接傳送資料給主機 190，次要乙太網路固態硬碟 110s 並不會直接接收來自主機 190 的命令，也不會直接傳送完成給主機 190。由主要乙太網路固態硬碟 110p 執行的管理命令可不需要與次要乙太網路固態硬碟 110s 進行任何通訊。

【0074】 在接收命令 180 後，主要乙太網路固態硬碟 110P 判斷資料的位置，以及判斷是否已存取所有資料 (S720)。如果主要乙太網路固態硬碟 110p 具備所有資料，主要乙太網路固態硬碟 110p 則將資料傳輸到主機 190 (S770)。

【0075】 如果主要乙太網路固態硬碟 110p 判斷沒有具備所有資料 (S720)，則主要乙太網路固態硬碟 110p 接下來諮詢對映配置表 300 來判斷所要求資料的位置 (S730)。一旦識別出相關的乙太網路固態硬碟 110 的集合之後，主要乙太網路固態硬碟 110p 即進行到命令 180 的執行。當主要乙太網路固態硬碟 110p 本身具備了所有相關要求的資料後，主要乙太網路固態硬碟 110p 傳輸資料 185。然而當要求的資料是分散在主要乙太網路固態硬碟 110p 及/或次要乙太網路固態硬碟 110s 的集合，則主要乙太網路固態硬碟 110p 將原始命令 180 分割成適當數目的子命令 132(S740)。子命令 132

的數目對應於要求的資料所分散其上的乙太網路固態硬碟 110 的數目。各個子命令 132 對應各個乙太網路固態硬碟 110 擁有的要求的資料部分。

【0076】 主要乙太網路固態硬碟 110p 將合適的開始邏輯區塊位址 (Start LBA, SLBA)、區塊數目 (Number of Blocks, NLB) 以及多個遠端分散/回收清單 (Scatter/Gather Lists, SGLs) 放入子命令 132 中。遠端分散/回收清單包括位址、金鑰、以及遠端主機 190 上的傳輸緩衝區大小。主要乙太網路固態硬碟 110p 接下來透過私用佇列對通訊通道 130, 在命令分割程序中將那些子命令 132 傳送到各別的次要乙太網路固態硬碟 110s (S750), 並且等待接收來自每一各別次要乙太網路固態硬碟 110s 的完成項目 134 (S760)。據此, 原始命令 180 被分割為子命令 132, 適當的乙太網路固態硬碟 110 即可平行執行資料傳輸, 因而使資料得以傳送到主機 190 (S770)。

【0077】 主要乙太網路固態硬碟 110p 針對正在執行中的命令 180 建立命令內容, 請參照圖 6 的說明。命令內容是用來追蹤子命令 132 的執行狀態, 以及子命令 132 的任何中間錯誤狀態 (S780)。一旦主要乙太網路固態硬碟 110p 確認管理命令完成, 主要乙太網路固態硬碟 110p 傳送完成項目給主機 190 (S790)。

【0078】 據此, 主要乙太網路固態硬碟 110p 接收 (S710) 並執行遠端主機跨結構高速非揮發性記憶體驅動程式 170 發出的所有管理命令 180。主要乙太網路固態硬碟 110p 可能獨自完成多項或全

部的管理命令 180(例如當主要乙太網路固態硬碟 110p 判斷(S720)獨自具有用來完成管理命令的所有所需資訊)。在部分案例中，在完成管理命令 180 之前，主要乙太網路固態硬碟 110p 可從次要乙太網路固態硬碟 110s 擷取某部分的資訊。如果主要乙太網路固態硬碟 110p 要從次要乙太網路固態硬碟 110s 尋找某些非使用者資料，主要乙太網路固態硬碟 110p 會建立並傳送管理子命令 132 給各別的次要乙太網路固態硬碟 110s。次要乙太網路固態硬碟 110s 可利用兩者之間的私用通訊通道 130，回傳任何所需的資料與子命令完成項目 134 給主要乙太網路固態硬碟 110p。

【0079】圖 8 為繪示根據本揭露一實施例的執行命名空間建立與刪除 (Namespace Create and Delete) 命令的流程圖 800。

【0080】請參照圖 8，主要乙太網路固態硬碟 110p 可接收與執行命名空間建立 (Namespace Create) 命令 (S810) 及/或刪除管理 (Delete Admin) 命令 (S811)。當主要乙太網路固態硬碟 110p 接收命名空間建立命令 (S810) 時，主要乙太網路固態硬碟 110p 可查詢對映配置表 300 (S820)，並且從總合可用集區配置適當的容量 (S830)。新建立的命名空間可具有完全來自主要乙太網路固態硬碟 110p 或完全來自某些次要乙太網路固態硬碟 110s 的快閃容量，或者新建立的命名空間可具有來自主要乙太網路固態硬碟 110p 與次要乙太網路固態硬碟 110s 任意組合的快閃容量。主要乙太網路固態硬碟 110p 接下來可在對映配置表 300 中記錄所配置的容量與相關的對應邏輯區塊位址範圍 (S840)。

【0081】 當主要乙太網路固態硬碟 110p 接收刪除命名空間的命名空間刪除（Namespace Delete）命令（S811）時，主要乙太網路固態硬碟 110p 查詢對映配置表 300（S821）、擷取對應的乙太網路固態硬碟（S831）、解除配置相關容量（S841），並據此更新對映配置表 300。

【0082】 關於由次要乙太網路固態硬碟 110s 執行的命名空間建立/刪除命令，次要乙太網路固態硬碟 110s 不會直接接收命名空間建立/刪除命令。在正常情況下，次要乙太網路固態硬碟 110s 會包含代表總容量的單一命名空間。在適當時間，主要乙太網路固態硬碟 110p 可發送命名空間建立命令或命名空間刪除命令給次要乙太網路固態硬碟 110s，以作為子命令 132。次要乙太網路固態硬碟 110s 接下來分別執行那些命令，並且回傳對應的完成項目 134 給主要乙太網路固態硬碟 110p。這樣的流程與任何接收來自主要乙太網路固態硬碟 110p 的管理子命令的流程是相同的。

【0083】 圖 9 為繪示根據本揭露一實施例所在自主要乙太網路固態硬碟 110p 控制下執行讀取/寫入（Read/Write）命令的流程圖 900。

【0084】 請參照圖 9，主要乙太網路固態硬碟 110p 可接收與執行所有輸入/輸出命令 180，包括讀取與寫入命令 180。當主要乙太網路固態硬碟 110p 接收到讀取/寫入命令 180 時（S910），主要乙太網路固態硬碟 110p 先查詢對映配置表 300（S920）。從對映配置表 300 中，主要乙太網路固態硬碟 110p 識別出相關使用者資料所在

的乙太網路固態硬碟 110 的集合。

【0085】 如圖 6 所述，主要乙太網路固態硬碟 110p 接下來針對原始命令建立命令內容 (S930)，如此一來，主要乙太網路固態硬碟 110p 便能追蹤子命令 132 的執行狀態。主要乙太網路固態硬碟 110p 接下來建立對應的讀取/寫入子命令 132 (S940)，並且傳送適當的子命令 132 給適當的次要乙太網路固態硬碟 110s (S950)。主要乙太網路固態硬碟 110p 也提供所有所需的傳輸網路相關資訊 (例如位址) 給次要乙太網路固態硬碟 110s。作為子命令 132 的一部分，次要乙太網路固態硬碟 110s 接收含有遠端緩衝位址、大小、安全性金鑰的遠端主機 190 分散/回收清單。

【0086】 子命令 132 中的資料傳輸欄位適當的修改為正確的位移。次要乙太網路固態硬碟 110s 直接傳送資料給遠端主機 190 的緩衝器 (S960)，當完成後，次要乙太網路固態硬碟 110s 傳送完成給主要乙太網路固態硬碟 110p (而不是直接傳送給主機 190)。如有必要，主要乙太網路固態硬碟 110p 會執行自己這部分的資料傳輸共享至遠端主機 190 (S960)。

【0087】 更進一步來說，各個次要乙太網路固態硬碟 110s 從主要乙太網路固態硬碟 110p 接收足夠的資訊，來執行任何直接資料傳輸至遠端主機 190 (S960)。在跨結構高速非揮發性記憶體通訊協定中，遠端直接記憶體存取傳輸服務 (遠端直接記憶體存取讀取與遠端直接記憶體存取寫入) 是用來將資料從次要乙太網路固態硬碟 110s 傳輸至遠端主機 190。遠端主機 190 可能需要支援遠端

直接記憶體存取通訊協定的共用接收佇列(Shared Receive Queue , SRQ) 功能 , 如此才能讓多個次要乙太網路固態硬碟 110s 傳輸資料至遠端主機 190 (S960) 。遠端直接記憶體存取通訊協定可在 Ethernet/IP/TCP (iWARP) 、 Ethernet/InfiniBand (RoCE v1) 或 Ethernet/IP/UDP (RoCE v2) 上執行。在次要乙太網路固態硬碟 110s 與遠端主機 190 之間的通訊方面 , 次要乙太網路固態硬碟 110s 嚴格執行僅與遠端主機 190 進行資料傳輸 (S960) 。亦即次要乙太網路固態硬碟 110s 與遠端主機 190 之間僅進行遠端直接記憶體存取-讀取與遠端直接記憶體存取-寫入作業 (S960) 。子命令完成項目 134 與任何非使用者資料傳輸都是使用遠端直接記憶體存取傳送作業或其他專屬通訊協定與主要乙太網路固態硬碟 110p 執行。

【0088】 當完成所有子命令後 (S970) , 如指示當主要乙太網路固態硬碟 110p 接收到所有的子命令完成項目時 , 主要乙太網路固態硬碟 110p 針對原始命令 180 建立完成項目 182 , 並且將完成項目 182 傳送給在主機 190 上的適當完成佇列 (S980) 。主要乙太網路固態硬碟 110p 接下來解除命令內容的配置 (S990) 。

【0089】 圖 10 為繪示根據本揭露一實施例的次要乙太網路固態硬碟 110s 執行管理子命令的流程圖 1000 。

【0090】 請參照圖 10 , 在本實施例中 , 沒有任何次要乙太網路固態硬碟 110s 曾經直接接收到來自主機 190 跨結構高速非揮發性記憶體驅動程式 170 的管理命令 180 或任何命令。反而是主要乙太網路固態硬碟 110p , 僅在必要時 , 傳送管理子命令 132 給次要乙

太網路固態硬碟 110s (S1010)。次要乙太網路固態硬碟 110s 接下來決定是否需要進行任何的資料傳輸 (S1020)，並且與主要乙太網路固態硬碟 110p 在私用通訊通道 130 上進行所需的任何資料傳輸 (S1030)。次要乙太網路固態硬碟 110s 接下來建立 (S1040) 並傳送完成項目 134 給主要乙太網路固態硬碟 110p (S1050)。在其他實施例中，次要乙太網路固態硬碟 110s 可使用遠端直接記憶體存取傳送作業來傳送資料與完成項目 134 給主要乙太網路固態硬碟 110p。

【0091】 圖 11 為根據本揭露一實施例所繪示的一次要乙太網路固態硬碟 110s 執行讀取/寫入命令的流程圖 1100。

【0092】 請參照圖 11，次要乙太網路固態硬碟 110s 的主要工作是執行讀取或寫入字命令 132(例如而不是管理字命令 132)。亦即，次要乙太網路固態硬碟 110s 主要是執行將資料移動至或移動自遠端主機 190，而不會進行通訊協定處理的其他方面。當次要乙太網路固態硬碟 110s 接收到讀取/寫入字命令時 (S1110)，次要乙太網路固態硬碟 110s 使用接收到的傳輸網路資訊 (S1120) 來發送遠端直接記憶體存取讀取或遠端直接記憶體存取寫入要求給遠端主機 190 (S1130)。作為子命令 132 的一部分，次要乙太網路固態硬碟 110s 接收遠端緩衝位址/位移、大小、安全性金鑰的細節。當必要資料傳輸完成後 (S1140)，次要乙太網路固態硬碟 110s 傳送含有適當錯誤狀態的完成項目 134 給主要乙太網路固態硬碟 110p 後 (S1150)。

【0093】 圖 12 為繪示根據本揭露一實施例的在次要乙太網路固態硬碟 110s 中的資料傳送與子命令完成同步的流程圖 1200。

【0094】 請參照圖 12，在給定的主機命令 180 中，雖然主要乙太網路固態硬碟 110p 將完成項目 182 給主機 192，但次要乙太網路固態硬碟 110s 的集合可執行傳送資料至主機 190。所有資料都傳送給遠端主機 190 之後，在給定的命令 180 的完成項目 182 才能提供給主機，這是因為命令 180 的完成項目 182 如果比相關資料早一步傳送到主機 190，可能會造成未定義的行為/錯誤。

【0095】 如上所述，主要乙太網路固態硬碟 110p 接收來自主機 190 的讀取/寫入命令 180 (S1210)，並且將命令 180 分割成數個子命令 132 (S1220)。接下來，分別接收來自主要乙太網路固態硬碟 110p 的子命令 132 的各個次要乙太網路固態硬碟 110s 發送資料傳輸給主機 190 (S1250)。在決定資料傳輸完成後 (S1260)，次要乙太網路固態硬碟 110s 發送完成項目 134 給主要乙太網路固態硬碟 110p (S1270)。接下來，接收到來自每個相關次要乙太網路固態硬碟 110s 的所有子命令完成項目 134 (S1230) 後，主要乙太網路固態硬碟 110p 發送完成項目 182 給主機 190 (S1240)。

【0096】 因為聚集的乙太網路固態硬碟資料傳輸與完成項目公佈是分散在乙太網路固態硬碟 110 的集合上，且應要達成資料與完成同步。一般而言，當單一乙太網路固態硬碟 110 進行命令執行的兩階段 (資料傳輸+完成公佈) 時，不會發生這些問題。然而，聚集的乙太網路固態硬碟不是這樣運作的。因此，主要乙太網路

固態硬碟 110p 必須等待來自各別次要乙太網路固態硬碟 110s 的所有子命令完成項目 134，才能公佈命令 180 的完成項目 182。更進一步而言，次要乙太網路固態硬碟 110s 必須確保所有的資料傳輸已全部且確實地完成，才能將子命令 132 的完成項目 134 傳送給主要乙太網路固態硬碟 110p。這樣的兩階段同步程序確保在任何時候，跨結構高速非揮發性記憶體通訊協定的完整性在聚集的乙太網路固態硬碟中是可達成的。

【0097】 綜上所述，因為只有單一主要乙太網路固態硬碟對主機而言是可見的，且此一主要乙太網路固態硬碟進行所有跨結構高速非揮發性記憶體通訊協定的處理作業，同時要追蹤所有次要乙太網路固態硬碟發出的相關子命令是否已完成，因此乙太網路固態硬碟的聚集群組對主機是顯示為單一的大型邏輯容量。

【符號說明】

【0098】

100、500：系統架構

110：乙太網路固態硬碟

110s：次要乙太網路固態硬碟

110p：主要乙太網路固態硬碟

120：乙太網路固態硬碟底座

130：私用通訊通道

132：子命令

134：完成項目

135：控制平面

140：快捷外設互聯標準匯流排

150：基板管理控制器

160：乙太網路交換器

170：機跨結構高速非揮發性記憶體驅動程式

180：命令

182：完成項目

185：資料

190：主機

230：乙太網路機架

240：機架頂端交換器

300：對映配置表

311：容量（百兆位元組）

312：傳輸位址（MAC/IP）

313：乙太網路固態硬碟索引

314：對映位址命名空間邏輯區塊位址範圍

400、700、800、900、1000、1100、1200：流程圖

530：乙太網路通訊通道

600：表格

610：命令標記

620：子命令#

630：累積的錯誤狀態

640：命令 ID

S410、S420、S430、S440：步驟

S710、S720、S730、S740、S750、S760、S770、S780、S790：

步驟

S810、S820、S830、S840、S811、S821、S831、S841：步驟

S910、S920、S930、S940、S950、S960、S970、S980、S990：

步驟

S1010、S1020、S1030、S1040、S1050：步驟

S1110、S1120、S1130、S1140、S1150：步驟

S1210、S1220、S1230、S1240、S1250、S1260、S1270：步

驟

【發明申請專利範圍】

【第1項】 一種跨結構高速非揮發性記憶體裝置用的儲存聚集方法，所述方法包括：

將聚集群組識別為包括多個跨結構高速非揮發性記憶體固態硬碟的聚集的乙太網路固態硬碟；

選擇所述聚集群組中的所述跨結構高速非揮發性記憶體固態硬碟的其中一個作為主要跨結構高速非揮發性記憶體固態硬碟；

選擇所述聚集群組中的所述跨結構高速非揮發性記憶體固態硬碟的其他個作為次要跨結構高速非揮發性記憶體固態硬碟；以及

以處理器將所述主要跨結構高速非揮發性記憶體固態硬碟中的對映配置表（Map Allocation Table）初始化，以管理所述跨結構高速非揮發性記憶體固態硬碟；

其中只有所述主要跨結構高速非揮發性記憶體固態硬碟對主機而言是可見的，且

其中基板管理控制器用於在開始時決定所述跨結構高速非揮發性記憶體固態硬碟中，何者為所述主要跨結構高速非揮發性記憶體固態硬碟。

【第2項】 如申請專利範圍第1項所述的方法，其中以所述處理器初始化所述對映配置表是在與所述聚集群組連結的儲存管理員的指導下進行。

【第3項】 如申請專利範圍第1項所述的方法，其中針對所述聚集群組的各個跨結構高速非揮發性記憶體固態硬碟，所述對映配置表包括所述跨結構高速非揮發性記憶體固態硬碟的容量、所述跨結構高速非揮發性記憶體固態硬碟的位址、所述跨結構高速非揮發性記憶體固態硬碟的剩餘容量。

【第4項】 如申請專利範圍第3項所述的方法，所述方法更包括將所述主要跨結構高速非揮發性記憶體固態硬碟的所述位址提供給使用者應用程式，以使所述聚集群組與所述使用者應用程式之間能夠進行資料傳輸。

【第5項】 如申請專利範圍第1項所述的方法，所述方法更包括：
在所述主要跨結構高速非揮發性記憶體固態硬碟上接收來自與所述聚集群組連結的所述主機的管理（Admin）命令；

判斷與所述管理命令對應的資料是否僅儲存在所述主要跨結構高速非揮發性記憶體固態硬碟上，或是儲存在一個或多個的所述次要跨結構高速非揮發性記憶體固態硬碟上；

當所述資料儲存在所述一個或多個的所述次要跨結構高速非揮發性記憶體固態硬碟上時，將所述管理命令分割成分別對應所述一個或多個的所述次要跨結構高速非揮發性記憶體固態硬碟的一個或多個管理子命令；

將所述資料傳輸至所述主機；

接收來自所述一個或多個的所述次要跨結構高速非揮發性記憶體固態硬碟的子命令完成項目；以及

從所述主要跨結構高速非揮發性記憶體固態硬碟建立並傳送完成項目至所述主機。

【第6項】如申請專利範圍第5項所述的方法，所述方法更包括：
在所述一個或多個的所述次要跨結構高速非揮發性記憶體固態硬碟的對應的次要跨結構高速非揮發性記憶體固態硬碟上接收所述一個或多個管理子命令的一管理子命令；

根據所述管理子命令判斷是否要將所述資料從所述對應的次要跨結構高速非揮發性記憶體固態硬碟傳輸到所述主要跨結構高速非揮發性記憶體固態硬碟；

建立完成項目；以及

將所述完成項目傳送到所述主要跨結構高速非揮發性記憶體固態硬碟。

【第7項】如申請專利範圍第1項所述的方法，方法更包括：
在所述主要跨結構高速非揮發性記憶體固態硬碟上接收建立命名空間（Namespace）的命令或刪除命名空間的命令；

參考所述主要跨結構高速非揮發性記憶體固態硬碟的所述對映配置表；

當所述命令是建立所述命名空間時，在所述主要跨結構高速非揮發性記憶體固態硬碟中及/或在一個或多個的所述次要跨結構高速非揮發性記憶體固態硬碟中配置容量，或是當所述命令是刪除所述命名空間時，收回所述主要跨結構高速非揮發性記憶體固態硬碟及/或所述一個或多個的所述次要跨結構高速非揮發性記憶

體固態硬碟的對應一者；以及

更新所述對映配置表。

【第8項】 如申請專利範圍第1項所述的方法，所述方法更包括：
在所述主要跨結構高速非揮發性記憶體固態硬碟上接收讀取/寫入（Read/Write）命令；

查詢所述主要跨結構高速非揮發性記憶體固態硬碟的所述對映配置表；

建立一個或多個讀取/寫入子命令；

將所述一個或多個讀取/寫入子命令分別傳送到一個或多個的所述次要跨結構高速非揮發性記憶體固態硬碟；

根據所述讀取/寫入命令，在所述主機與所述主要跨結構高速非揮發性記憶體固態硬碟及/或所述一個或多個的所述次要跨結構高速非揮發性記憶體固態硬碟之間傳輸資料；以及

傳輸所述資料後傳送完成至所述主機。

【第9項】 如申請專利範圍第8項所述的方法，所述方法更包括：
在一個或多個的所述次要跨結構高速非揮發性記憶體固態硬碟的對應的次要跨結構高速非揮發性記憶體固態硬碟上接收所述一個或多個讀取/寫入子命令的一個讀取/寫入子命令；

擷取對應所述一個讀取/寫入子命令的輸送資訊；

從所述對應的次要跨結構高速非揮發性記憶體固態硬碟發出讀取/寫入要求至所述主機；以及

在對應所述一個讀取/寫入子命令的資料傳輸完成之後，將完

成項目傳送到所述主要跨結構高速非揮發性記憶體固態硬碟。

【第10項】 一種跨結構高速非揮發性記憶體乙太網路固態硬碟群組中的跨結構高速非揮發性記憶體固態硬碟用的容量聚集方法，所述方法包括：

識別聚集群組的多個跨結構高速非揮發性記憶體固態硬碟；

將所述跨結構高速非揮發性記憶體固態硬碟的其中一個指派為主要跨結構高速非揮發性記憶體固態硬碟；以及

將所述跨結構高速非揮發性記憶體固態硬碟的剩餘者指派為次要跨結構高速非揮發性記憶體固態硬碟，

其中主機的主機驅動程式唯一可見的跨結構高速非揮發性記憶體固態硬碟為所述主要跨結構高速非揮發性記憶體固態硬碟，

其中只有所述主要跨結構高速非揮發性記憶體固態硬碟對所述主機而言是可見的，且

其中基板管理控制器用於在開始時決定所述跨結構高速非揮發性記憶體固態硬碟中，何者為所述主要跨結構高速非揮發性記憶體固態硬碟。

【第11項】 如申請專利範圍第10項所述的方法，所述方法更包括根據所述主要跨結構高速非揮發性記憶體固態硬碟從所述主機接收的命令，維護所述主要跨結構高速非揮發性記憶體固態硬碟的對映配置表，

其中所述對映配置表指示在所述聚集群組的所述主要跨結構高速非揮發性記憶體固態硬碟與一個或多個的所述次要跨結構高

速非揮發性記憶體固態硬碟之間分割邏輯區塊位址（logical block address，LBA）空間。

【第12項】 如申請專利範圍第10項所述的方法，所述方法更包括以處理器將對映配置表初始化，以依據所述跨結構高速非揮發性記憶體固態硬碟中作為所述主要跨結構高速非揮發性記憶體固態硬碟的指派一者來設定所述聚集群組。

【第13項】 如申請專利範圍第10項所述的方法，所述方法更包括依據所述主要跨結構高速非揮發性記憶體固態硬碟從所述主機接收到的命令來聚集所述次要跨結構高速非揮發性記憶體固態硬碟與所述主要跨結構高速非揮發性記憶體固態硬碟的容量，使得所述聚集群組的多個所述跨結構高速非揮發性記憶體固態硬碟對所述主機而言顯現為單一的聚集邏輯容量。

【第14項】 如申請專利範圍第10項所述的方法，所述方法更包括：
使用所述主要跨結構高速非揮發性記憶體固態硬碟將容量配置給一個或多個的所述次要跨結構高速非揮發性記憶體固態硬碟；
以及

使用所述主要跨結構高速非揮發性記憶體固態硬碟來記錄所配置的所述容量與相關聯的對映邏輯區塊位址範圍於對映配置表中。

【第15項】 如申請專利範圍第10項所述的方法，所述方法更包括以所述主要跨結構高速非揮發性記憶體固態硬碟過度配置所述次要跨結構高速非揮發性記憶體固態硬碟與所述主要跨結構高速非揮發性記憶體固態硬碟的聚集容量。

【第16項】 如申請專利範圍第10項所述的方法，所述方法更包括：
在所述主要跨結構高速非揮發性記憶體固態硬碟上接收來自
所述主機的命令；

將所述命令分割成多個子命令且每一子命令分別對應至對應
的所述次要跨結構高速非揮發性記憶體固態硬碟的個別一者；以
及

將所述子命令從所述主要跨結構高速非揮發性記憶體固態硬
碟傳送到對應的所述次要跨結構高速非揮發性記憶體固態硬碟。

【第17項】 如申請專利範圍第16項所述的方法，所述方法更包括
基於個別的所述子命令，從對應的次要跨結構高速非揮發性記憶
體固態硬碟直接傳送資料到所述主機。

【第18項】 如申請專利範圍第16項所述的方法，所述方法更包括：
在對應的所述次要跨結構高速非揮發性記憶體固態硬碟上接
收來自所述主要跨結構高速非揮發性記憶體固態硬碟的個別的所
述子命令；

執行對應個別的所述子命令的任務；以及

所述任務完成後，從對應的所述次要跨結構高速非揮發性記
憶體固態硬碟傳送個別的子命令完成項目到所述主要跨結構高速
非揮發性記憶體固態硬碟。

【第19項】 如申請專利範圍第18項所述的方法，所述方法更包括：
使用所述主要跨結構高速非揮發性記憶體固態硬碟來維護子
命令內容表；

在所述主要跨結構高速非揮發性記憶體固態硬碟接收來自所述次要跨結構高速非揮發性記憶體固態硬碟的所述子命令完成項目；以及

根據接收到的所述子命令完成項目，以所述主要跨結構高速非揮發性記憶體固態硬碟來追蹤所述子命令的執行情況。

【第20項】一種聚集的乙太網路固態硬碟群組，包括：

乙太網路固態硬碟底座；

乙太網路交換器，位於所述乙太網路固態硬碟底座上，用以與主機的主機驅動程式通訊；

處理器，耦接至所述乙太網路交換器；

快捷外設互聯標準（Peripheral Component Interconnect Express，PCIe）交換器，耦接至基板管理控制器；以及

多個跨結構高速非揮發性記憶體固態硬碟，包括：

主要跨結構高速非揮發性記憶體固態硬碟；以及

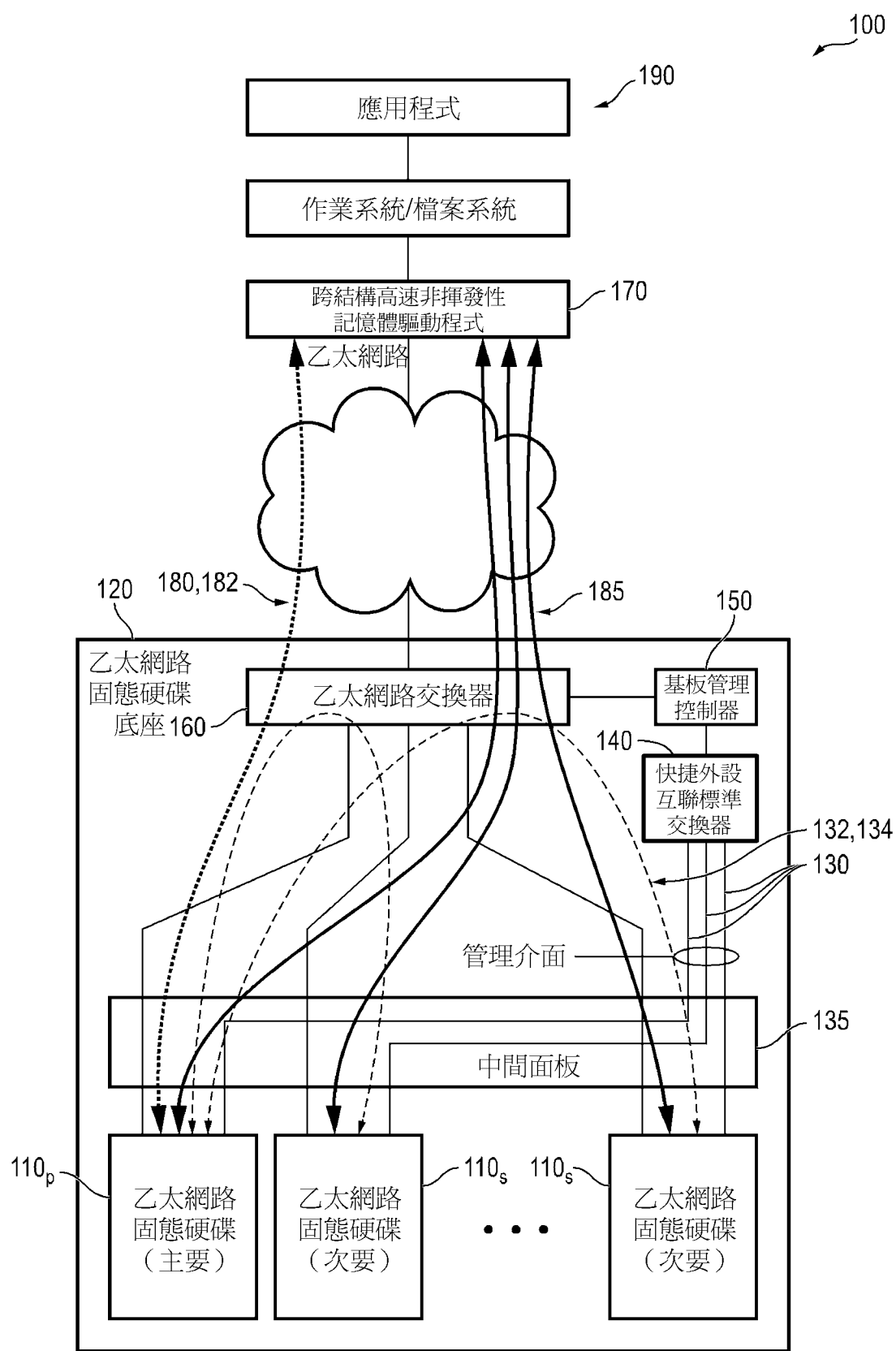
多個次要跨結構高速非揮發性記憶體固態硬碟，透過包括所述乙太網路交換器與所述快捷外設互聯標準交換器的私用通訊通道與所述主要跨結構高速非揮發性記憶體固態硬碟連結，

其中只有所述主要跨結構高速非揮發性記憶體固態硬碟對所述主機而言是可見的，且

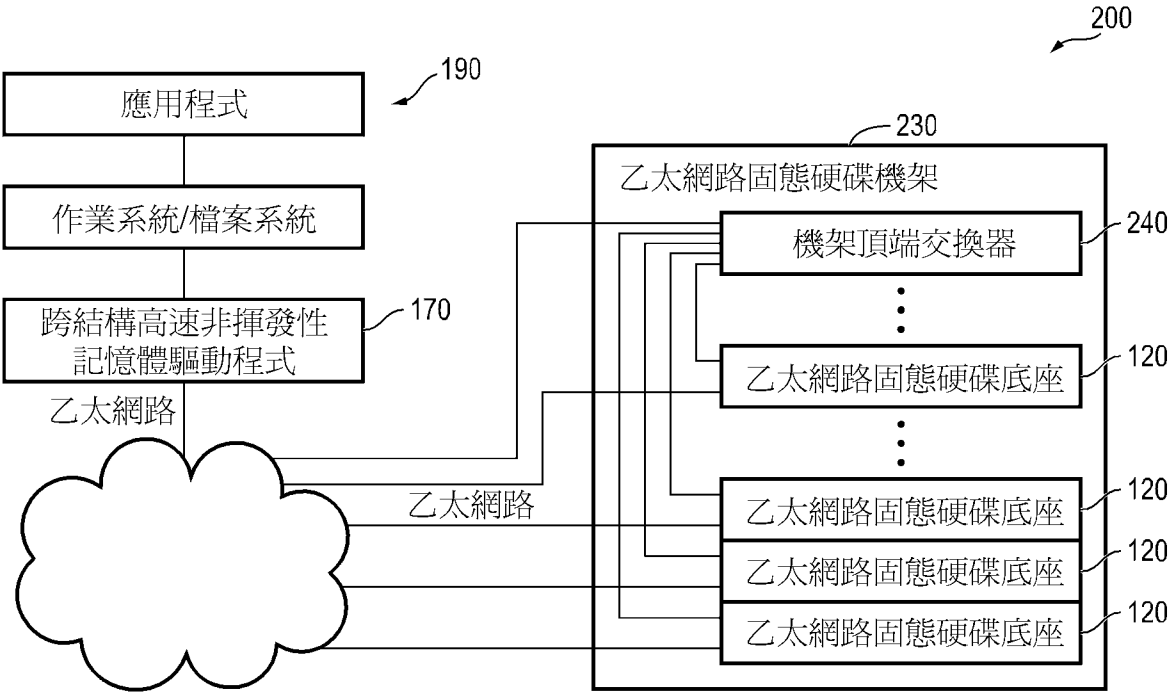
其中所述基板管理控制器用於在開始時決定所述跨結構高速非揮發性記憶體固態硬碟中，何者為所述主要跨結構高速非揮發

性記憶體固態硬碟。

【發明圖式】



【圖1】

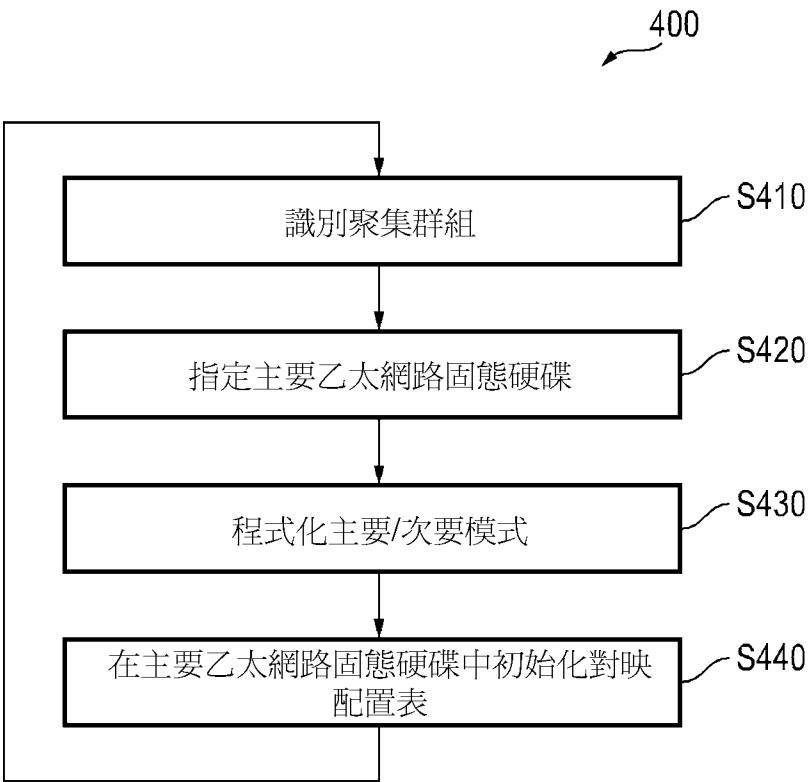


【圖2】

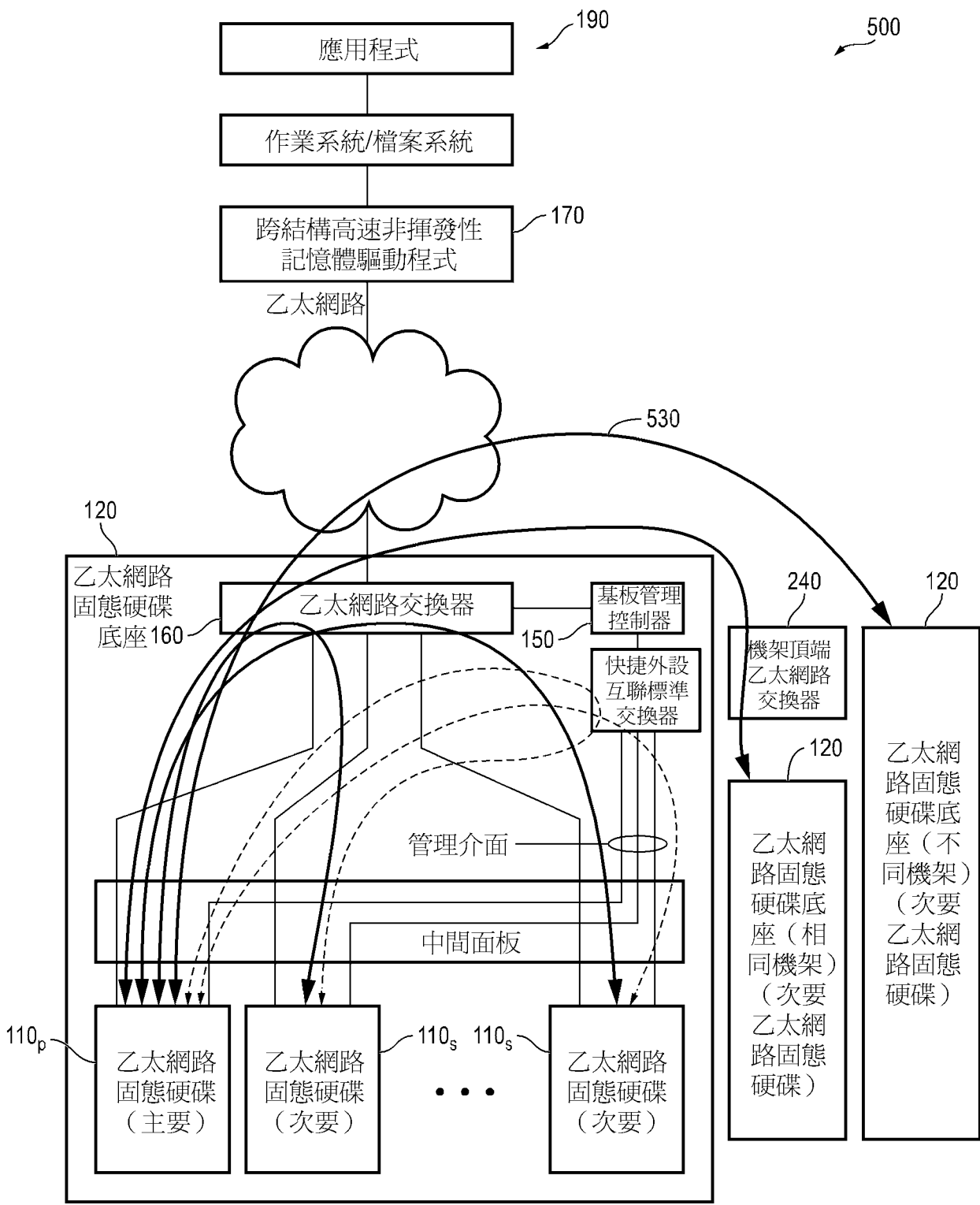
300

313 乙太網路 固態硬碟 索引	311 容量 (百兆位元組)	312 傳輸位址 (MAC/IP)	314 對映位址命名空間 邏輯區塊位址範圍	可用容量 (百兆位元組)	其他
0	16				
1	8				
2	32				
3	16				

【圖3】



【圖4】

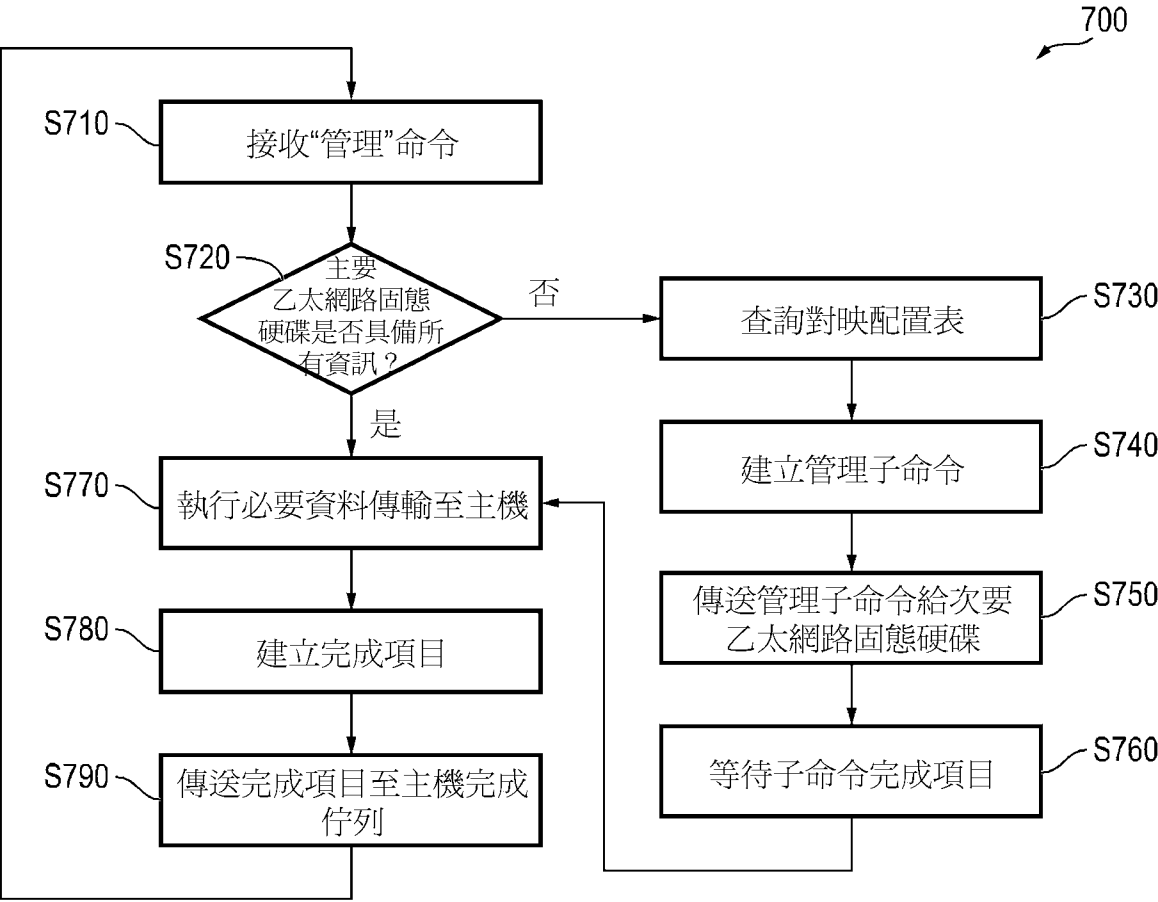


【圖5】

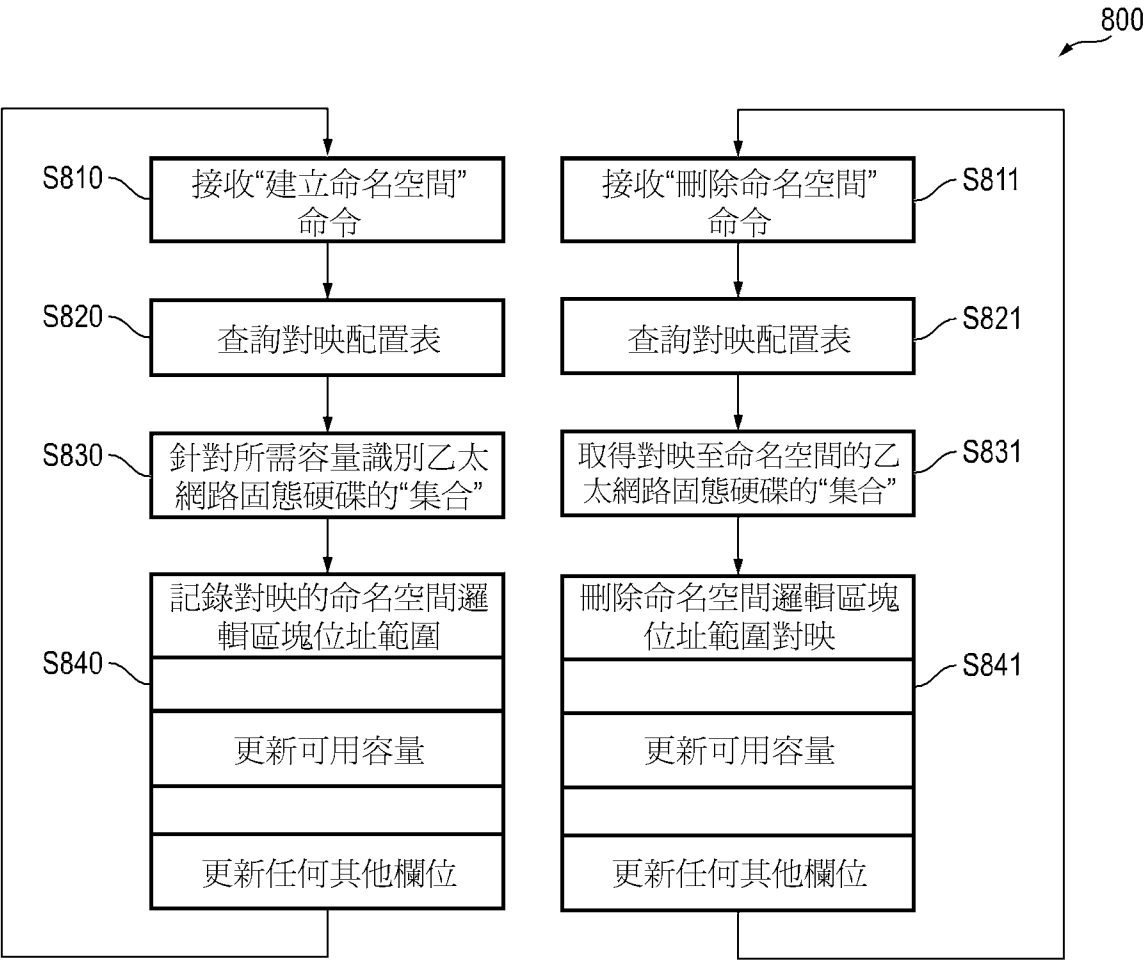
600

640 命令ID	610 命令標記	620 子命令#	630 累積的錯誤狀態
123			
15			
39			
3			

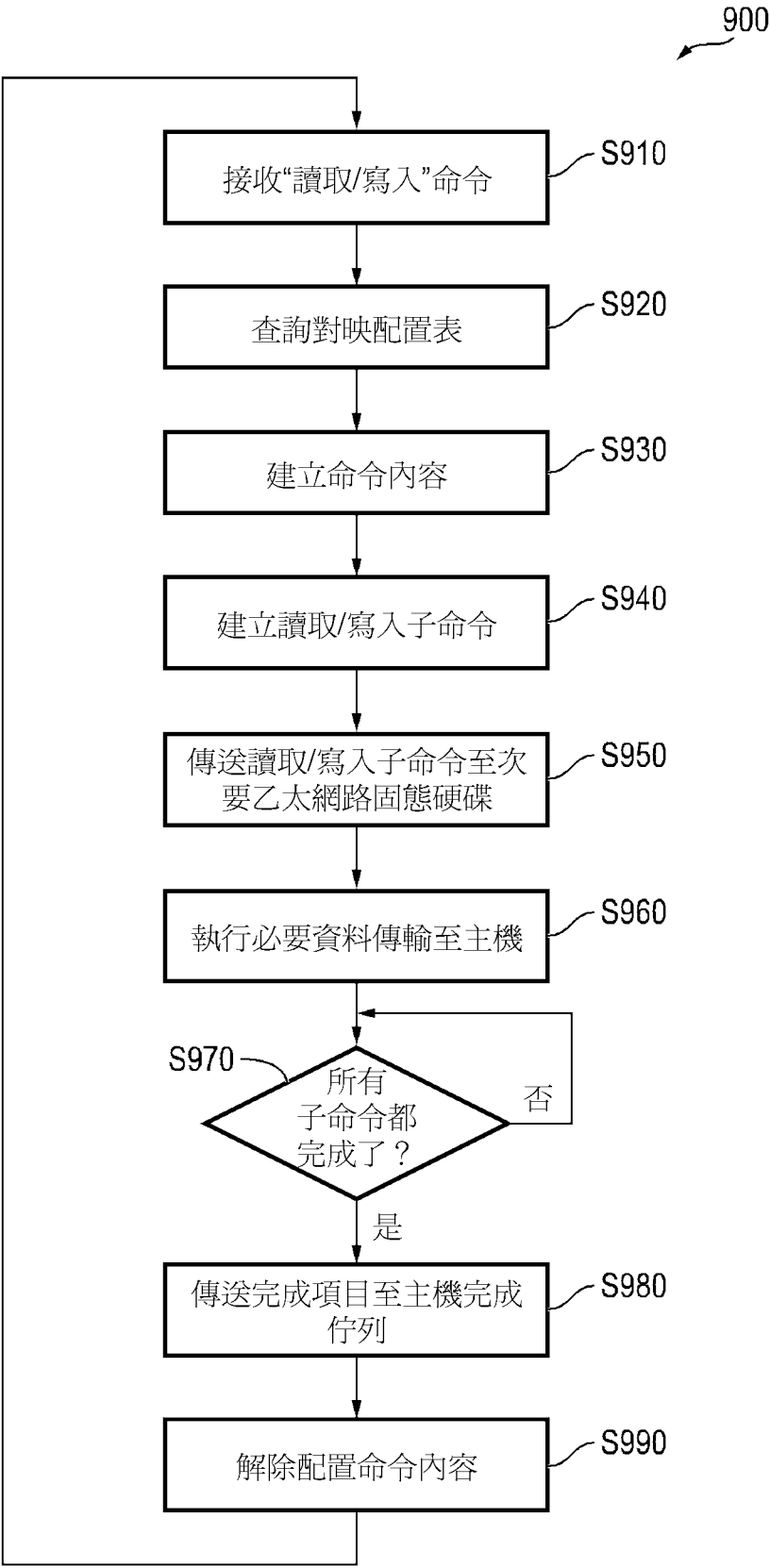
【圖6】



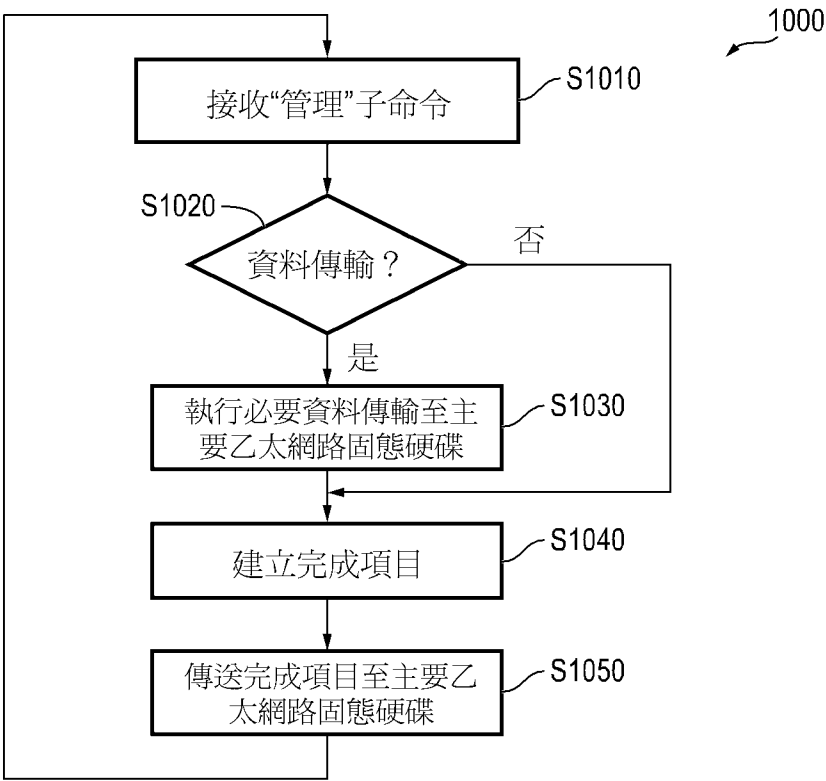
【圖7】



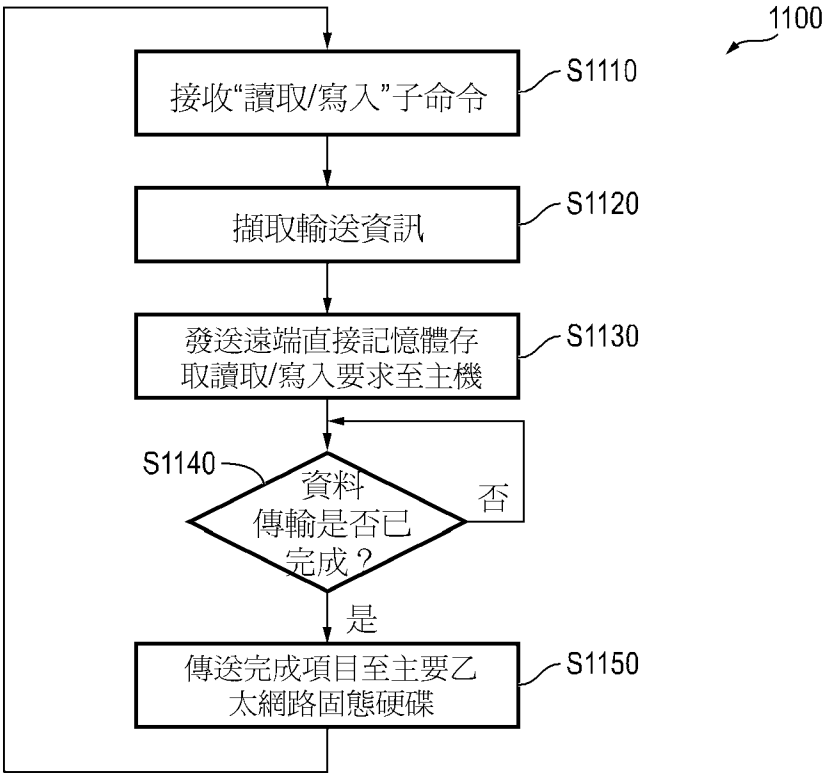
【圖8】



【圖9】

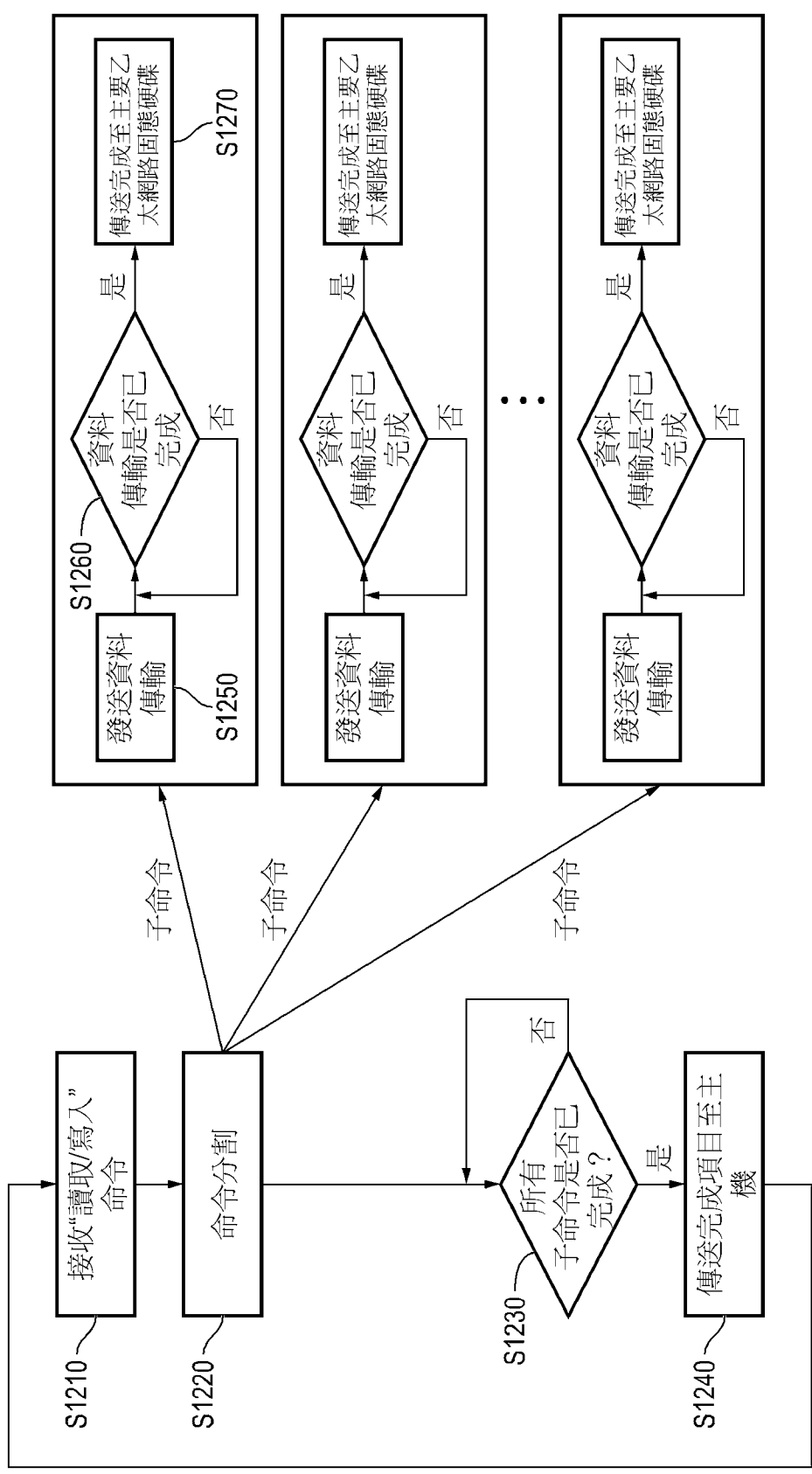


【圖10】



【圖11】

1200



【圖12】